ALGORITHMIC METHODS IN DIFFRACTION
MICROSCOPY

A Dissertation
Presented to the Faculty of the Graduate School
of Cornell University
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

by
Pierre Thibault
August 2007
Recent diffraction imaging techniques use properties of coherent sources (most notably x-rays and electrons) to transfer a portion of the imaging task to computer algorithms. “Diffraction microscopy” is a method which consists in reconstructing the image of a specimen from its diffraction pattern. Because only the amplitude of a wavefield incident on a detector is measured, reconstruction of the image entails recovering the lost phases. This extension of the “phase problem” commonly met in crystallography is solved only if additional information is available.

The main topic of this thesis is the development of algorithmic techniques in diffraction microscopy. In addition to introducing new methods, it is meant to be a review of the algorithmic aspects of the field of diffractive imaging.

An overview of the scattering approximations used in the interpretation of diffraction datasets is first given, as well as a numerical propagation tool useful in conditions where known approximations fail. Concepts central to diffraction microscopy – such as oversampling – are then introduced and other similar imaging techniques described.

A complete description of iterative reconstruction algorithms follows, with a special emphasis on the difference map, the algorithm used in this thesis. The formalism, based on constraint sets and projection onto these sets, is then defined and explained. Simple projections commonly used in diffraction imaging are then
described.

The various ways experimental realities can affect reconstruction methods will then be enumerated. Among the diverse sources of algorithmic difficulties, one finds that noise, missing data and partial coherence are typically the most important. Other related difficulties discussed are the detrimental effects of crystalline domains in a specimen, and the convergence problems occurring when the support of a complex-valued specimen is not well known.

The last part of this thesis presents reconstruction results; an account is given on real specimen reconstruction from x-ray data (a yeast cell, and a quasicrystal electron density) and from visible light experiments (insect wings). Simulated data in various multiple-measurements experiments are also reconstructed using a new iterative approach.
BIOGRAPHICAL SKETCH

Pierre Thibault was born in 1978, in Rivière-du-Loup, in the province of Québec, Canada. He lived there until he was 18 years old. After a 6 month travel in Europe, he attended Université de Montréal to obtain his bachelor degree in mathematics and physics. In 2002, he received his M.Sc. degree at the same institution under the direction of Laurent Lewis. During the next five years, he lived in Ithaca to complete his doctoral studies in Cornell. Pierre is interested in many things: working wood, drawing, singing, sailing, traveling, baking bread, ...probably too many for him to ever settle on one and become good at it.
ACKNOWLEDGEMENTS

I want first to express my deepest gratitude to Veit Elser, my advisor for this thesis work. Working with Veit has been a great pleasure from the beginning to the end. After I arrived at Cornell, I chose to work with Veit because the research he was involved in seemed very fun, but mostly because I had the feeling that working with him would be a very enjoyable experience. I did not know how right I would be. Veit will remain a model for me in many aspects of my work and my life. I especially admire his sharp intelligence, his very strong sense of ethics, and his generosity. Veit was very supportive and helping when and after my son Éloi was born. I must also thank him and his wife for lending me their car for almost a month, after mine was broken in an accident.

I also want to thank my colleagues Ivan Rankenburg, Duane Loh and Simon Gravel. Working in their company was both fun and stimulating. All were always ready to help and discuss.

I am indebted to David Sayre, Janos Kirz, Chris Jacobsen and David Shapiro, from Stonybrook University, who generously shared both their experiments and their experience. I especially thank David Sayre for his continued encouragements and his advice, as well as for inviting me to the Erice summer school. I also thank Andrew Stewart for many fruitful discussions. Thanks to Tobias Beetz for inviting me to Brookhaven National Laboratory while he was postdoc there. Tobi is among the most rigorous and thorough young scientists I have met.

The optical diffraction experiments at Cornell were made possible with the help of Georg Hoffstätter and Nick Szabo, who both provided advice and access to essential equipment. Itai Cohen and his student Leif Gibbens also helped making a second setup for additional measurements.
Marc de Boissieu kindly shared his x-ray diffraction data used in the quasicrystal reconstructions.

I take this opportunity to thank people who have been an inspiration for me and whose contact impacted on my life and on my work. High-school professors Jacques Viel and Gilles Lebel helped me shape the way I think and how I see science. Friends from my B.Sc. program, Jean-François Brière, Louis-Pierre Arguin, Simon Gravel and many others, also had a positive and durable influence on the way I am today.

I thank my mother, Lise, my father, Claude, and my sister, Danielle, for their support and their love. Thank you also to Louise, François, Thérèse and Guy for their help, advice and encouragements.

Finally, I thank Amélie for being there always, and for giving me the inspiration and the balance I need to continue. Being with you and Éloi is the most beautiful adventure.

Parts of this work have been supported by the Natural Sciences and Engineering Research Council of Canada and by the US Department of Energy (Grant No. DoE-FG02-05ER46198).
# Table of Contents

Biographical Sketch .................................................. iii  
Acknowledgements ....................................................... iv  
Table of Contents ...................................................... vi  
List of Tables ........................................................... viii  
List of Figures ........................................................... ix  

1 Introduction ................................................................ 1  

2 Theory ......................................................................... 6  
2.1 Preliminaries: definitions of useful tools ......................... 6  
2.2 Scattering of light and electrons by matter ....................... 10  
2.2.1 The wave equation for electromagnetic fields ............... 10  
2.2.2 The wave equation for electrons ............................... 12  
2.2.3 Solutions of the wave equation ............................... 12  
2.2.4 Numerical simulations ........................................... 25  
2.3 The phase problem .................................................... 28  
2.4 Diffraction-based imaging techniques ............................. 29  
2.4.1 Diffraction microscopy .......................................... 30  
2.4.2 Other techniques ................................................ 38  

3 Iterative algorithms ..................................................... 43  
3.1 Formalism .................................................................. 44  
3.1.1 Constraint sets .................................................... 44  
3.1.2 Projections ........................................................ 45  
3.1.3 Iterative maps ..................................................... 46  
3.2 The difference map .................................................... 47  
3.2.1 Discussion ........................................................ 49  
3.3 Other algorithms ....................................................... 56  
3.3.1 Algorithms for convex sets ..................................... 56  
3.3.2 Common phase retrieval algorithms ......................... 58  
3.4 Projections .............................................................. 63  
3.4.1 Diffraction projections .......................................... 64  
3.4.2 Direct space projections ........................................ 73  

4 Working with experimental data .................................... 81  
4.1 Experiment limitations ............................................... 81  
4.1.1 Coherence ........................................................ 82  
4.1.2 Noise .............................................................. 90  
4.1.3 Missing data ..................................................... 96  
4.2 Algorithm limitations ................................................ 107  
4.2.1 Data centering .................................................... 107  
4.2.2 Support shape determination ................................. 108  
4.2.3 Complex-valued reconstructions .............................. 111
5 Reconstructions and simulations

5.1 The yeast cell reconstruction
5.1.1 Reconstruction procedure
5.1.2 Reconstruction resolution
5.1.3 Dust
5.1.4 Discussion

5.2 Optical diffraction microscopy
5.2.1 Experiment description
5.2.2 Reconstructions

5.3 Quasicrystal reconstruction
5.3.1 Reconstruction details
5.3.2 Algorithmic test of the centrosymmetry assumption

5.4 Simulations
5.4.1 Propagation through a yeast cell model
5.4.2 Reconstructions involving many datasets

6 Conclusion

A Retrieved phases for the AlMnPd quasicrystal x-ray dataset

B The effect of averaging on the resolution of a reconstruction
B.1 Averaging over translation in the transverse plane
B.2 Averaging over defocus planes
LIST OF TABLES

5.1 Index of refraction of the two constituents of a model yeast cell . . 168
5.2 Physical parameters in the aperture scanning simulation . . . . . . 178

A.1 Retrieved phases for the AlMnPd quasicrystal . . . . . . . . . . 187
# LIST OF FIGURES

1.1 The colorwheel used to encode complex-valued images ................................ 5

2.1 Geometry of free-space wave propagation ........................................... 15
2.2 An oversampled diffraction pattern ...................................................... 35
2.3 Example of support shapes and their constraint ratios. ......................... 36
2.4 Fourier holography simulation .............................................................. 40

3.1 A typical difference map error plot ......................................................... 53
3.2 Illustration of the threshold projection .................................................. 76

4.1 Illustration of the effect of partial transverse coherence ....................... 87
4.2 Illustration of the effect of partial longitudinal coherence ................. 89
4.3 One-dimensional example of unconstrained degrees of freedom caused by missing data .......................................................... 98
4.4 Examples of unconstrained modes with a circular support and a square missing data region ......................................................... 105
4.5 Examples of unconstrained modes with an L-shaped support and a square missing data region ......................................................... 106
4.6 Three support shapes having the same autocorrelation support ......... 109
4.7 Example of the triple-intersection method for the determination of a tighter support .......................................................... 110
4.8 Regions of validity of various wave scattering approximations ......... 113
4.9 Illustration of different cases of complex-valued images ............... 115
4.10 Illustration of the diffraction pattern from a crystalline domain .... 120
4.11 Simulated density of a double-walled nanotube and its diffraction pattern .......................................................... 122
4.12 Demonstration of the interference problem with multiple crystalline domains .......................................................... 124
4.13 Possible crystalline domains reconstructions with a noisy diffraction pattern .......................................................... 126

5.1 Soft x-ray diffraction pattern of a yeast cell ...................................... 130
5.2 Real part of the yeast cell autocorrelation .......................................... 132
5.3 Early reconstruction using a tentative value constraint ..................... 134
5.4 Unconstrained modes for the yeast cell reconstruction .................... 136
5.5 Yeast cell reconstruction prior to mode replacement ................. 137
5.6 Difference map error for the yeast cell reconstruction .................... 138
5.7 Yeast cell reconstruction after the mode replacement ...................... 139
5.8 Algorithmic transfer function for the yeast cell reconstruction ....... 140
5.9 Reconstruction of the small scatterers surrounding the cell ........ 142
5.10 Collimator part of the optics experiment ........................................... 145
5.11 Detector part of the optics experiment .............................................. 146
5.12 Diffraction pattern and autocorrelation of an insect wing ............... 149
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.13</td>
<td>Insect wing reconstruction</td>
<td>151</td>
</tr>
<tr>
<td>5.14</td>
<td>Algorithmic transfer function and fidelity map of the insect wing reconstruction</td>
<td>153</td>
</tr>
<tr>
<td>5.15</td>
<td>Diffraction pattern and autocorrelation of a <em>Drosophila</em> wing</td>
<td>155</td>
</tr>
<tr>
<td>5.16</td>
<td><em>Drosophila</em> wing reconstruction</td>
<td>156</td>
</tr>
<tr>
<td>5.17</td>
<td>Pixel value distribution of the <em>Drosophila</em> reconstruction</td>
<td>157</td>
</tr>
<tr>
<td>5.18</td>
<td>Reconstructed atomic surfaces of an AlMnPd quasicrystal</td>
<td>164</td>
</tr>
<tr>
<td>5.19</td>
<td>Slice, in parallel space, of the reconstructed quasicrystal electron density</td>
<td>165</td>
</tr>
<tr>
<td>5.20</td>
<td>Histogram distribution of the phases of six bright reflections</td>
<td>167</td>
</tr>
<tr>
<td>5.21</td>
<td>One slice of the yeast cell model</td>
<td>169</td>
</tr>
<tr>
<td>5.22</td>
<td>Exit wave and pixel value distribution for the model yeast cell</td>
<td>170</td>
</tr>
<tr>
<td>5.23</td>
<td>Simulated wave propagation through the model cell</td>
<td>171</td>
</tr>
<tr>
<td>5.24</td>
<td>Forward and backward propagation of the exit-wave from the model cell</td>
<td>171</td>
</tr>
<tr>
<td>5.25</td>
<td>Simulation of an aperture scanning experiment</td>
<td>177</td>
</tr>
<tr>
<td>5.26</td>
<td>Reconstruction from the diffraction patterns shown on Figures 5.25(b)–(c)</td>
<td>179</td>
</tr>
<tr>
<td>5.27</td>
<td>Simulated focal series measurements of a pure phase object</td>
<td>181</td>
</tr>
<tr>
<td>5.28</td>
<td>Focal series reconstruction</td>
<td>182</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

This thesis gives an account of recent advances in the algorithmic methods used in various diffractive imaging techniques. Recent developments in various scientific disciplines make this work timely, but also, in a sense, obvious: it stands as a natural research direction for the further advancement of diffraction imaging.

Almost exactly 60 years ago, synchrotron radiation was observed for the first time in an accelerator (a betatron) built at General Electric (Elder et al., 1947). The following decades have seen a tremendous development of synchrotron accelerators, first favored for the high energy at which they can accelerate particles and later built solely for the x-rays they produce. Synchrotron sources provide nowadays the most powerful and most coherent x-ray beams available, two characteristics essential to modern applications of diffraction imaging techniques.

Diffractive imaging relies on many other technological innovations, including: charge-coupled devices (CCD), diverse microfabrication techniques, and of course modern computers. The latter has certainly deeply affected the way modern science is practiced in general. Modern computers offer far more than a mere increase in calculation power, and have caused an important, paradigmatic shift in how recent problems are approached. The iterative algorithms described in this thesis are an example of problem solving methods whose efficiency is based on the availability of computing power. The development of these iterative algorithms is also essential to modern diffractive imaging methods.

In this work, diffractive imaging methods will be defined as the ensemble of imaging techniques requiring, in the data analysis, that the wave properties of the
scattered radiation be taken into account. Most of the time, retrieval of information (such as the reconstruction of a scattering specimen) necessitates coherent illumination (both transverse and longitudinal). For this reason, the field is sometimes called “coherent diffractive imaging”. Detection of a scattered wave involves always, fundamentally, the measurement of an intensity, proportional to the squared amplitude of the wave. The task of algorithms generally entails recovering the phase part of the wave that was lost in the measurement. In this work, we will consider that “reconstruction” “phasing” and “phase retrieval” algorithms are essentially synonymous terms for the computational (iterative) methods that retrieve information from diffraction data.

The central imaging technique studied in this work is diffraction microscopy, which involves the measurement of the far-field diffraction pattern from a specimen, and the subsequent reconstruction of the specimen’s image from this diffraction pattern. The origins of diffraction microscopy can be traced back to Sayre (1952a) who, inspired by Shannon’s sampling theorem (Shannon, 1949), mentioned that the phase problem would be solvable in principle if a crystal diffraction pattern was sufficiently oversampled. Meanwhile, the development of iterative algorithms to solve the phase problem in optics was pioneered by Gerchberg and Saxton (1972) and Fienup (1978, 1982). The first x-ray measurement of a non-crystalline sample was made by Yun et al. (1987), but no reconstruction was attempted. Experimental demonstrations of full diffraction microscopy came only in the late 90s. The first successful two-dimensional reconstruction of an experimental data set was achieved in 1999 by Miao et al. (1999). Subsequently, other successful reconstructions of metallic structures (Miao et al., 2002; Williams et al., 2003; He et al., 2003; Chapman et al., 2006b) and of biological specimens (Miao et al., 2003;
Shapiro et al., 2005) were reported.

This thesis is organized as a review of the theory and algorithmic techniques used in diffraction microscopy and other imaging methods. The original work is not solely concentrated in a “results” chapter, and is somewhat more scattered. The overall effort of joining many concepts together in a single framework could be, in itself, considered as an original contribution. An outline of the structure of the thesis is given below.

Chapter 2 introduces the theory on the physical phenomena occurring in diffraction imaging techniques. A classical description, with the scalar wave equation will suffice to explain most of the needed concepts. The main derivations in this chapter are well-known approximations to the solution of the wave equation. However, these results are not merely stated but carefully reexpressed in a language appropriate for the type of geometry generally present in imaging experiments. After the analytical results, we will present a numerical multislice formula that can be used to simulate wavefield propagation through a specimen. In the last part of the chapter, diffraction microscopy will be defined in detail, along with the important concepts of oversampling, overdetermination and uniqueness of the solution.

The general formalism of the iterative reconstruction algorithm (called the difference map) is given in chapter 3. This algorithm is based on constraint sets and projections onto these constraint sets. The discussion will be kept general as this algorithm can be applied to a wide variety of problems. After a description of other reconstruction algorithms, we will provide a list of projections especially useful in phase retrieval applications.

Chapter 4 is a discussion of situations recurrent in practical applications. It lists various ways experimental realities can lead to a departure from the ideal model
presented in the two previous chapters. Noise, missing data and partial coherence, to name a few, are sources of difficulty that can compromise reconstructions.

Chapter 5 presents various applications of the methods and concepts introduced in the previous chapters. It includes both simulations and reconstructions from real data. The reconstructed objects are a yeast cell from its x-ray diffraction pattern, insect wings from visible-wavelength data taken in a setup built at Cornell, and the electron density of a quasicrystal from x-ray data. An application of the wave propagation algorithm introduced in chapter 2 is also shown. Finally, possible extensions of the difference map to techniques involving multiple datasets is discussed and illustrated with simulated data.

**Conventions**

The terminology used in this thesis is standard for most physicists, and appropriate definitions will be given when needed.

This thesis contains also many complex-valued images. Rendering separately the real and imaginary parts is in general not appropriate since most reconstructions are defined only up to an overall phase factor; depending on this phase factor, the real and imaginary parts are mixed in a non transparent way. Another possibility is to separate magnitude and phase; we have decided instead to express both quantities in a single image by mapping the magnitude to brightness and phase to hue (saturation is always kept constant and at its maximum value). Figure 1.1 shows the resulting colorwheel encoding of the complex plane.
Figure 1.1: The colorwheel used to encode complex-valued images. Images colored with this scheme have their magnitude rescaled such that the brightest pixel has the maximum brightness (on the edge of the wheel). Rotation of the wheel with respect to phase angle is generally free because of the overall phase factor indeterminacy in reconstruction; it is sometimes adjusted to improve perceived contrast in the image.
Chapter 2

Theory

This chapter gives an overview of the main theoretical tools needed to understand the physical aspects of various relevant experimental methods.

Overall, the required background knowledge is more broad than deep. We will first give a rapid overview of the nomenclature commonly used throughout this thesis (section 2.1). We will then give a relatively complete description of the scattering phenomena relevant to x-ray, optical and electron diffraction experiments (2.2). Because of the geometry of these experiments, the most convenient point of view is one that makes a natural separation between the propagation direction and the plane perpendicular to it. All quantities and concepts will thus be placed in this framework – the same as used in the formalism of Fourier optics. After presenting useful analytic approximations to the solution of the wave equation, we will describe a numerical method useful to model scattering where other analytic methods don’t apply. Once the various approximations to the formation of diffraction patterns are exposed, we state the phase problem, which is central to this thesis (2.3). We will finally give a description of diffraction microscopy and similar experimental techniques (2.4).

2.1 Preliminaries: definitions of useful tools

We first give definitions of tools and signal processing concepts repeatedly used in this work. The definitions are stated in one dimension as generalization to more than one dimension is straightforward.
The continuous-space Fourier transform is
\[
\tilde{f}(q) = \mathcal{F} f = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{iqx} \, dx.
\] (2.1)

A quantity expressed as a function of spatial coordinates will be said to be in direct space, whereas it is in Fourier space or reciprocal space when it is the Fourier transform of the direct-space function. Direct space coordinates will be expressed with the usual notation \( \mathbf{r} = (x, y, z) \) and \( \mathbf{q} = (q_x, q_y, q_z) \) will always be used for the Fourier space coordinates.

Various elementary properties of the Fourier transform will be used repeatedly in this thesis.

- A property – often coined as Parseval’s theorem – is the conservation of the \( L_2 \) norm:
  \[
  \int |\tilde{f}(q)|^2 \, dq = \int |f(x)|^2 \, dx.
  \] (2.2)

- Fourier transform of a translated function appears as a phase tilt or phase ramp factor
  \[
  \mathcal{F} f(x + a) = e^{iqa} \mathcal{F} f(x).
  \] (2.3)

- The Fourier transform of a real function \( f(x) = f^*(x) \) has the property
  \[
  \tilde{f}(q) = \tilde{f}^*(-q).
  \] (2.4)

In diffraction, this property is called the “Friedel law”, and \( \tilde{f}(q), \tilde{f}(-q) \) is often called a “Friedel pair”. 
Convolution is defined as
\[
(f * g)(x) = \int f(x')g(x - x')dx',
\] (2.5)
and is related to the Fourier transform operation through
\[
f * g = \sqrt{2\pi} \mathcal{F}^{-1}(\mathcal{F}f \mathcal{F}g).
\] (2.6)

Closely related to the convolution operation is the cross-correlation:
\[
(f \circ g)(x) = \int f(x')g^*(x' - x)dx'
\] (2.7)
\[
= \sqrt{2\pi} \mathcal{F}^{-1}[\mathcal{F}f \quad (\mathcal{F}g)^*],
\] (2.8)
of which a very important special case is the autocorrelation:
\[
A_f = f \circ f = \sqrt{2\pi} \mathcal{F}^{-1}(|\mathcal{F}f|^2).
\] (2.9)

From (2.5), it can be seen that convolution and cross-correlation are the same if one of the functions is real and even.

Practical applications involving the use of digital computers generally imply a discretization of data, either as a “binning” for real number approximations, or as a sampling of space on a grid. With single and double precision floating point numbers, the former case is rarely problematic. The latter case, ubiquitous in imaging-related data processing, has more important consequences. Interestingly, some physical systems having a periodic structure show a “natural discretization” which is very well suited to computer-based data manipulation. In other cases, continuous space concepts generally have an obvious discrete equivalent.

The one-dimensional discrete Fourier transform of an \(N\) long vector \(f_m\) is given by
\[
\tilde{f}_n = \mathcal{F}f = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} f_m e^{2\pi i mn/N},
\] (2.10)
and the inverse Fourier transform differs only by a complex conjugation of the exponential:

\[ f_n = \mathcal{F}^{-1}\tilde{f} = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} \tilde{f}_m e^{-2\pi imn/N}. \]  

The convolution definition is similar to the continuous case,

\[ (f * g)_n = \sum_{m=0}^{N-1} f_m g_{n-m} \]  
\[ = \sqrt{N}\mathcal{F}^{-1}(\mathcal{F}f \mathcal{F}g) \]

with the important observation that the sum is cyclic: the argument of \( g \) in this definition has to be taken modulo \( N \).

The discrete Fourier transform of a continuous function \( f(x) \) can be approximated by choosing a suitably fine grid spacing \( \Delta x \) to sample the function\(^1\): \( f_m = f(m\Delta x) \). The resulting Fourier transform is also sampled on a grid \( \tilde{f}(n\Delta q) = \tilde{f}_n \). Grid samplings in continuous-to-discrete conversion of a Fourier transform follows from the relation

\[ \Delta x\Delta q = \frac{2\pi}{N}, \]  

where \( N \) is the length of the vectors \( f_m \) and \( \tilde{f}_n \). In practice, all discrete Fourier transforms are executed using the fast Fourier transform (FFT), whose computing time scales as \( N \log(N) \) (Cooley and Tukey, 1965).

Given a sampling interval \( \Delta x \), the highest frequency present in a discrete Fourier transform is the Nyquist frequency,

\[ q_N = N\Delta q/2 = \pi/\Delta x. \]  

A too large sampling interval can lead to aliasing: when this occurs, frequencies higher than \( q_N \) are wrapped and appear as lower frequencies. If \( f \) is band-limited,

\(^1\)It is also necessary to assume that the function is periodic.
that is,
\[ \tilde{f}(q) = 0 \quad \text{for } |q| > B, \quad (2.16) \]
and
\[ B < q_N, \quad (2.17) \]
then aliasing does not occur. The sampling theorem states that, in this case, \( f(x) \) is completely determined by its sampled version \( f_n \):
\[
f(x) = \sum_n f_n \frac{\sin (x - n\Delta x)}{(x - n\Delta x)}. \quad (2.18)
\]
A signal sampled while condition (2.17) is satisfied is said to be “oversampled”.

### 2.2 Scattering of light and electrons by matter

Imaging techniques covered in this work involve mostly x-rays, but also electron diffraction. After a quick derivation of the wave equation for both types of radiation, various approximations and solution approaches will be described.

#### 2.2.1 The wave equation for electromagnetic fields

A derivation of the wave equation from Maxwell’s equations can be found in almost any introductory book on electromagnetism (Jackson, 1975) or optics (Born and Wolf, 1999). A version of this derivation is reproduced here for completeness and to emphasize the approximation involved (equation 2.20). We consider the propagation of an oscillating electromagnetic field (frequency \( \omega \)) in a medium.
Maxwell’s equations read:

\[ \nabla \cdot [\epsilon(r) \mathbf{E}(r)] = 0 \]  
(2.19a)

\[ \nabla \cdot \mathbf{H}(r) = 0 \]  
(2.19b)

\[ \nabla \times \mathbf{E}(r) = i\omega \mu \mathbf{H}(r) \]  
(2.19c)

\[ \nabla \times \mathbf{H}(r) = -i\omega \epsilon(r) \mathbf{E}(r) \]  
(2.19d)

\( \epsilon(r) \) is the space-dependent and frequency-dependent electric permittivity of the medium,\(^2\) and \( \mu \) is its magnetic permeability, assumed in this case to have negligible variations. In general, \( \epsilon \) is a tensor describing birefringence effects in the medium.

The wave equation is obtained from this set of equations provided that the length scale of the variations in the dielectric medium is much longer than the wavelength associated with the oscillating electromagnetic field, that is:

\[ |\nabla \epsilon| \ll k |\epsilon|, \]  
(2.20)

where \( k = 2\pi/\lambda \) is the (vacuum) wavenumber associated with the oscillating electromagnetic field (\( k \sim |\nabla \times \mathbf{E}|/|\mathbf{E}| \)). This assumption will be needed again in section 2.2.3, in the derivation of the Rytov approximation. Using (2.19c), we find:

\[ \nabla \times (\epsilon \mathbf{E}) = \nabla \epsilon \times \mathbf{E} + \epsilon \nabla \times \mathbf{E} \approx \epsilon \nabla \times \mathbf{E} \]
\[ = i\omega \epsilon \mu \mathbf{H}. \]  
(2.21)

Using the identity \( \nabla \times \nabla \times \mathbf{H} = \nabla (\nabla \cdot \mathbf{H}) - \nabla^2 \mathbf{H} \), and equations (2.19b) and (2.19d), we finally get

\[ \nabla^2 \mathbf{H} + \epsilon \mu \omega^2 \mathbf{H} = 0. \]  
(2.22)

\(^2\)The medium’s conductivity, also frequency-dependent, is absorbed in the complex-valued \( \epsilon(r) \).
Taking the curl on both sides of this equation and using approximation (2.20) once more, $E$ is found to satisfy the same equation as $H$. Because we are not concerned with polarization, the equation relevant to our problem is the scalar ("Helmholtz") wave equation:

$$\nabla^2 \Psi + k^2 n^2 \Psi = 0,$$

with $k = \omega/c$ and $n^2 = c^2 \epsilon \mu$.

### 2.2.2 The wave equation for electrons

Electron microscopy is based on the fact that high energy electrons behave much the same way as an electromagnetic wave. Neglecting spin and interactions between particles, the Schrödinger equation already has the form of (2.23):

$$\nabla^2 \Psi + \frac{2m}{\hbar^2} (E - V(r)) \Psi(r) = 0,$$

where $\Psi$ is the electron’s wavefunction, $m$ is its mass, $E$ is its energy, and $V$ is the potential. In general, electrons are accelerated to relativistic speeds, giving (Spence, 2003)

$$\frac{2mE}{\hbar^2} \rightarrow \frac{2mE}{\hbar^2} \left(1 + \frac{E}{2mc^2}\right) = k^2.$$

Hence, letting $n^2(r) = (1 - V(r)/E)$, we recover the exact form of equation (2.23).

### 2.2.3 Solutions of the wave equation

The Helmholtz equation can be solved exactly only in special situations. Most of the time, numerical methods or series expansions are required. Formally, this partial differential equation requires that boundary conditions for $\Psi$ be specified. In all practical purposes, $\delta n$ will be non-zero only in a finite region of space (the
region occupied by the specimen) and the boundary conditions are that $\Psi(\mathbf{r})$ equals the free-space solution as $|\mathbf{r}| \to \infty$.

In the x-ray community, the refractive index is commonly expressed in terms of its purely refractive (real) and absorptive (imaginary) components, $\delta$ and $\beta$: \(^\text{3}\)

$$n = 1 + \delta n = 1 - \delta - i\beta. \quad (2.26)$$

Since, for x-rays, $|\delta n| \ll 1$, the approximation $n^2 \approx 1 + 2\delta n$ will always be used. Note also, that $2\delta n$ is approximately equal to the electric susceptibility of the medium, $\chi_e$, because $\epsilon = (1 + \chi_e)\epsilon_0$.

**Free-space propagation**

When $\delta n = 0$, Eq. (2.23) has a constant coefficient function and therefore takes a particularly simple form in Fourier space:

$$\left(k^2 - q^2\right)\tilde{\Psi}(\mathbf{q}) = 0. \quad (2.27)$$

Clearly, $\tilde{\Psi}(\mathbf{q}) = 0$ unless $|\mathbf{q}| = k$, with the obvious interpretation that the free-space solution of $\Psi(\mathbf{r})$ is a superposition of all plane waves which have wavenumber $k$. From a quantum mechanical point of view, this is merely a consequence of energy conservation. In crystallography (as well as other fields), the sphere $|\mathbf{q}| = k$ is called the “Ewald sphere” (Ewald, 1913).

Scattering experiments generally involve the measurement of the wavefield amplitude in a plane transverse to its main propagation direction. This suggests a separation of the transverse $[\mathbf{r}_\perp = (x, y)]$ and parallel ($z$) components of the free-propagating wavefield. Taking the Fourier transform of Eq. (2.23) only in the

\(^{3}\text{In optics, } \beta \text{ is sometimes called the “extinction coefficient”}\).
transverse plane gives:

$$
(-\mathbf{q}_\perp^2 + \partial_z^2 + k^2) \bar{\Psi}(\mathbf{q}_\perp; z) = 0. \tag{2.28}
$$

This can be seen as a family of one-dimensional differential equations in $z$, decoupled in $\mathbf{q}_\perp$. The general solution thus reads

$$
\bar{\Psi}(\mathbf{q}_\perp; z) = \bar{\Psi}^-(\mathbf{q}_\perp) e^{-i\kappa z} + \bar{\Psi}^+(\mathbf{q}_\perp) e^{i\kappa z}, \tag{2.29}
$$

where $\bar{\Psi}^\pm(\mathbf{q}_\perp)$ are two independent functions representing forward (+) and backward (−) scattering, and

$$
\kappa = \sqrt{k^2 - \mathbf{q}_\perp^2}. \tag{2.30}
$$

In view of eq. 2.27, the two terms of eq. 2.29 should not come as a surprise; for a fixed $\mathbf{q}_\perp$, the Ewald sphere condition allows only for two different $z$-components of $\mathbf{q}$: $q_z = \pm \kappa$. In scattering experiments, there are generally no back-propagating terms, so that $\bar{\Psi}^-(\mathbf{q}_\perp)$ can be safely set to zero when considering propagation past the specimen. $\bar{\Psi}^+(\mathbf{q}_\perp)$ is then interpreted as the wavefield at some arbitrary $z$, and equation (2.29) gives how the wave evolves at any further $z$:

$$
\Psi(\mathbf{r}_\perp; z) = \mathcal{F}^{-1} \left[ \bar{\Psi}(\mathbf{q}_\perp) e^{i\kappa z} \right]. \tag{2.31}
$$

Figure 2.1 illustrates the geometry of wave propagation in Fourier space.

Far-field (or “Fraunhofer”) diffraction is given by this equation in the limit $z \to \infty$. It is common to derive the expression for the far-field diffraction pattern from the Kirchhoff-Fresnel integral (Born and Wolf, 1999). The same result can be obtained in the current Fourier optics framework:

$$
\Psi_{\text{far field}} \rightarrow \mathcal{F}^{-1} \left[ \bar{\Psi}(\mathbf{q}_\perp) e^{i\kappa z} \right]
\propto \int d^2 \mathbf{q} \, \Psi(\mathbf{q}) \exp \left[ ikz \left( \sqrt{1 - (\mathbf{q}/k)^2} + \mathbf{q}/k \cdot \mathbf{r}/z \right) \right]. \tag{2.32}
$$
As $kz \to \infty$, the integrand vanishes unless the phase is stationary. This happens when

$$\frac{r_\perp}{z} = \frac{q_\perp}{\sqrt{k^2 - q_\perp^2}} = \frac{q_\perp}{\kappa}. \quad (2.33)$$

Using standard asymptotic methods, we find that the far-field intensity measured in direction $u = r_\perp/z$ is

$$I(u) = |\Psi_{\text{far field}}(zu)|^2 \propto \frac{1}{1 + u^2} |\bar{\Psi}(q_\perp = \kappa u)|^2. \quad (2.34)$$

This fundamental result is well known: the far-field diffraction pattern is proportional to the absolute value of the 2D Fourier transform (in the transverse plane) of the wavefield past the specimen (the exit wave). Since $|u| = \tan \theta$, where $\theta$ is the angle between the direction of the measurement and the propagation direction, $(1 + u^2)^{-1}$ can also be written as $\cos^2 \theta$.

A simpler and sometimes more useful expression can be obtained in the small-
angle scattering case, where $\overline{\Psi}(q)_{\perp}$ is assumed not negligible only when $|q_{\perp}| \ll k$.

Then, one can expand $\kappa$ to the first non-zero order in $q_{\perp}$ (this is called the paraxial approximation) and (2.31) becomes

$$\Psi(r_{\perp}; z) = \mathcal{F}^{-1} \left\{ \overline{\Psi}(q)_{\perp} \exp \left[ ikz \left( 1 - \frac{q_{\perp}^2}{2k^2} \right) \right] \right\}.$$  (2.35)

This can be identified at once as a direct-space convolution,

$$\Psi(r_{\perp}; z) = \frac{1}{\sqrt{2\pi}} \Psi(r_{\perp}) * P(r_{\perp}; z),$$  (2.36)

where

$$P(r_{\perp}; z) = \mathcal{F}^{-1} \exp \left[ ikz \left( 1 - \frac{q_{\perp}^2}{2k^2} \right) \right] = \frac{-i k}{z} \exp \left[ ikz \left( 1 + \frac{r_{\perp}^2}{2z^2} \right) \right]$$  (2.37)

is sometimes called the “Fresnel propagator”. Computations become easier within this approximation. Eq. (2.36) is written

$$\Psi(r_{\perp}; z) = \frac{-i k}{2\pi z} e^{ikz} \int d^2 r_{\perp}' \Psi(r_{\perp}') \exp \left[ \frac{ik}{2z} (r_{\perp} - r_{\perp}')^2 \right].$$  (2.38)

Using, as above, the unitless vector $u = r_{\perp}/z$ to describe the propagation direction, and expanding the square term in the exponent, we get

$$\Psi(zu; z) = \frac{-i k}{2\pi z} \exp \left[ ikz \left( 1 + \frac{u^2}{2} \right) \right] \int d^2 r_{\perp}' \Psi(r_{\perp}') \exp \left[ \frac{ikr_{\perp}^2}{2z} - iku \cdot r_{\perp}' \right].$$  (2.39)

The integral in this equation has the form of a Fourier transform when one identifies $q$ with $ku$, thus yielding:

$$\Psi(zu; z) = \frac{-i k}{z} \exp \left[ ikz \left( 1 + \frac{u^2}{2} \right) \right] \mathcal{F} \left[ \Psi(r_{\perp}') \exp \left( \frac{ikr_{\perp}^2}{2z} \right) \right]|_{q_{\perp} = ku}.$$  (2.40)

Sometimes, $\Psi(r_{\perp})$ is 0 except within a finite region of space (that is, if $\Psi(r_{\perp})$ has a support). In these cases, the phase modulation caused by the rightmost exponential factor in (2.40) reaches a maximum along the edges of the non-zero
region. If $a$ is the typical width of this region, then the phase at the edges is $\pi f/4$, where

$$f = \frac{a^2}{\lambda z} \quad (2.41)$$

is known as the “Fresnel number”. This unitless number is often used to capture the regime in which an experiment is done: small Fresnel numbers correspond to the far-field regime, while large Fresnel numbers are related to near-field diffraction.

As $f \to 0$ (or increasing $z$), the phase factor inside the Fourier transform can be replaced by 1, and one recovers the small scattering angle version of (2.34):

$$I(u) \propto |\bar{\Psi}(q_\perp = ku)|^2. \quad (2.42)$$

Equation (2.40) has another interesting interpretation; an ideal thin lens has the effect of introducing a spatially varying phase lag in an incoming wavefield $\Psi \to e^{i\varphi} \Psi$. If the lens is converging, the phase is of the form

$$\varphi(r_\perp) = -\frac{k r_\perp^2}{2l}, \quad (2.43)$$

where $l$ is the focal length of the lens. Substituting $e^{i\varphi} \Psi$ in (2.40), one finds that the field propagated a distance $z = l$ is equal to the Fourier transform of $\Psi$ (with an additional quadratic phase factor). This is the well-known Fourier transforming property of converging lenses, used for spatial filtering and in many microscopy techniques (Born and Wolf, 1999).

**Small-angle approximation**

We now consider the case $\delta n \neq 0$. Solving (2.23) is especially simple when only small-angle scattering is considered. Factoring out the quickly oscillating component of the wavefield, $\psi(r) = \Psi(r)e^{-ikz}$, equation (2.23) becomes:

$$\nabla^2 \psi + 2ik \partial_z \psi + 2k^2 \delta n \psi = 0. \quad (2.44)$$
In this formulation (which is just a translation in Fourier space) spatial derivatives of $\psi$ do not scale like $k$. Hence, in the $k \to \infty$ limit, the first term can be neglected and one gets a first-order eikonal equation:

$$\partial_z \psi(r) = ik\delta n(r) \psi(r), \quad (2.45)$$

wherein the perpendicular ($r_{\perp}$) components of $\psi$ are now independent. The solution is

$$\psi(r) = A(r_{\perp}) \exp[ik\delta n_{\perp}], \quad (2.46)$$

where

$$\delta n_{\perp}(r_{\perp}) = \int \delta n(r_{\perp}, z) \, dz \quad (2.47)$$

is the projected refractive index along the propagation direction and $A(r_{\perp})$ is a function defined by the boundary conditions (the form of the incident wave).\(^4\)

This approach is satisfactory to model some experimental situations, for instance in some hard x-ray imaging techniques. We will define $k\delta n_{\perp}$ as the optical thickness of the scattering object. If

$$|k\delta n_{\perp}| \ll 1, \quad (2.48)$$

the object is said to be optically thin, and the “projection approximation” is valid:

$$\psi(r) \propto 1 + ik\delta n_{\perp}. \quad (2.49)$$

Otherwise, the specimen is optically thick.

**Born approximation**

The Born expansion is a common approach to solving eq. (2.23). The method is essentially perturbative: let $\epsilon$ be an expansion parameter for $\delta n$ (it will later be

\(^4\)Note that this solution is mathematically equivalent to the WKB approximation for classical motion (nearly free particle).
set to 1). We write (2.23) as

\[ \nabla^2 \Psi + k^2 \Psi = -2k^2 \epsilon \delta n \Psi \quad (2.50) \]

and let the solution have the form

\[ \Psi(r) = \Psi_0(r) + \epsilon \Psi_1(r) + \epsilon^2 \Psi_2 + \cdots \quad (2.51) \]

This is the Born series. The zeroth order solution is the homogeneous wave equation:

\[ \nabla^2 \Psi_0 + k^2 \Psi_0 = 0, \quad (2.52) \]

and higher order terms have the form

\[ \nabla^2 \Psi_n + k^2 \Psi_n = -2k^2 \delta n \Psi_{n-1}. \quad (2.53) \]

Since \( \Psi_{n-1} \) is known when solving for \( \Psi_n \), all that remains to be done is writing the general solution to the inhomogeneous Helmholtz equation:

\[ \nabla^2 \Psi(r) + k^2 \Psi(r) = -2k^2 f(r) \quad (2.54) \]

where \(-2k^2 f\) is a source term. From a kinematic point of view, equation (2.53) represents multiple scattering events: \( \Psi_1 \) is the wave produced in response to a source \( \delta n \Psi_0 \), that is, the direct scattering of the incoming wave produced by the disturbance \( \delta n \). \( \Psi_2 \) is the “second scattering” amplitude: it is the scattering amplitude coming from \( \Psi_1 \).

The standard way to solve (2.54) is with the use of the Green function. Taking the Fourier transform gives

\[ (k^2 - q^2) \tilde{\Psi}(q) = -2k^2 \tilde{f}(q). \quad (2.55) \]

Defining

\[ \tilde{G}(q) = -(2\pi)^{-3/2} \frac{2k^2}{k^2 - q^2}, \quad (2.56) \]
we have
\[ \tilde{\Psi}(q) = (2\pi)^{3/2} \tilde{G}(q) \tilde{f}(q), \] (2.57)

or, back in direct space,
\[ \Psi = G \ast f. \] (2.58)

\( G \) is the Green function for equation (2.54):
\[ G(r) = -\frac{2k^2}{(2\pi)^{3/2} 4\pi |r|} e^{i k |r|}. \] (2.59)

In terms of the Green function, the terms of the Born series are:
\[ \Psi_n = G \ast (\delta n \Psi_{n-1}) \] (2.60)

The first Born approximation (or just Born approximation) keeps only the linear term in \( \delta n \):
\[ \Psi_{\text{Born}} = \Psi_0 + G \ast (\delta n \Psi_0). \] (2.61)

The practical significance of this solution can be made more apparent. As has been seen above, the far-field diffraction pattern is proportional to the Fourier transform (in the transverse plane) of the exit-wave. In the “hybrid space” \( (q_\perp, z) \), eq. (2.61) reads
\[ \tilde{\Psi}_{\text{Born}} = \tilde{\Psi}_0 + \tilde{G} \ast_z [\delta n \ast_{q_\perp} \tilde{\Psi}_0], \] (2.62)

where the convolution symbols have been indexed with the variables over which they operate. The convolution in \( z \) can be calculated explicitly:
\[ G \ast_z \tilde{f} = \sqrt{2\pi} \mathcal{F}_z^{-1} \left[ \tilde{G}(q) \tilde{f}(q) \right] \]
\[ = \int_{-\infty}^{\infty} \frac{2k^2}{(2\pi)^{3/2} |q|^2 - k^2} e^{i q_z z} dq_z \]
\[ = \frac{2k^2}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} \frac{\tilde{f}(q) e^{i q_z z}}{(q_z - \kappa) (q_z + \kappa)} dq_z. \] (2.63)
with \( \kappa = \sqrt{k^2 - q_{\perp}^2} \), as above. This integral is done in the complex plane. The path of integration is chosen to ensure that only outgoing waves are present. \( +\kappa \) is therefore included in the contour and \( -\kappa \) excluded for \( z > 0 \). For \( z < 0 \), the situation is reversed and only \( -\kappa \) is included in the contour.

\[
G *_{z} \tilde{f} = \frac{1}{\sqrt{2\pi}} \frac{ik^2}{\sqrt{k^2 - q_{\perp}^2}} \begin{cases} 
\tilde{f}(q_{\perp}, q_z = \kappa)e^{ikz} & \text{if } z \geq 0, \\
\tilde{f}(q_{\perp}, q_z = -\kappa)e^{-ikz} & \text{if } z < 0.
\end{cases}
\quad (2.64)
\]

We are, however, only interested in the transmission term \((z > 0)\).

It is common to assume that the incoming wave is a plane wave \( \Psi_0 = Ae^{ikz} \).

In this case, one finds

\[
\tilde{\Psi}_{\text{Born}} = 2\pi A \delta(q_{\perp}) e^{ikz} + \frac{\sqrt{2\pi} A k^2}{\sqrt{k^2 - q_{\perp}^2}} \tilde{\delta n}(q_{\perp}, q_z = \kappa - k) e^{ikz}.
\quad (2.65)
\]

Using (2.34), the measured intensity in direction \( u \) (except for \( u = 0 \)) becomes

\[
I(u) \propto \left| \tilde{\delta n}(\kappa u, \kappa - k) \right|^2.
\quad (2.66)
\]

In the Born approximation, assuming a plane wave illumination, the measured diffraction pattern is the three-dimensional Fourier transform of the scattering strength measured on the Ewald sphere. This result is of utmost importance for three-dimensional reconstructions (in crystallography for instance) as it provides the recipe to find the magnitude of the three-dimensional Fourier transform of a specimen through the collection and assembly of many two-dimensional measurements taken at different incidence angles.

Note that when \( k \to \infty \), one recovers the “projection approximation” [eq. (2.49)]:

\[
\Psi_{\text{Born}} = \mathcal{F}^{-1}_{q_{\perp}} \tilde{\Psi}_{\text{Born}} \\
\approx Ae^{ikz} + ik A e^{ikz} \mathcal{F}^{-1}_{q_{\perp}} \left[ \sqrt{2\pi} \tilde{\delta n}(q_{\perp}, q_z = 0) \right] \\
= Ae^{ikz} (1 + ik \delta n_{\perp}).
\quad (2.67)
\]
This allows a more precise definition of the validity conditions of the Born approximation. The condition $|\Psi_1| \ll |\Psi_0|$ simply becomes $|k\delta n_\perp| \ll 1$, that is, the first Born approximation is valid only for optically thin specimens.

**Rytov approximation**

The Rytov approximation (Rytov, 1937; Born and Wolf, 1999) is based on first assuming that the perturbation of the incident wave is multiplicative (instead of additive, as in the Born approximation):

$$\Psi = \Psi_0 e^\phi$$

(2.68)

It is convenient to use the exponential of a function as the factor since $\phi = 0$ is the free-space solution $\delta n = 0$. It is also a natural choice as the $k \to \infty$ case has this form [Eq. (2.46)].

Substitution of (2.68) in the Helmholtz equation involves the calculation of the laplacian:

$$\nabla^2 (\Psi_0 e^\phi) = \left[ \nabla^2 \Psi_0 + \nabla \Psi_0 \cdot \nabla \phi + 2 \nabla \Psi_0 \cdot \nabla \phi + \nabla^2 \phi \right] e^\phi.$$  

(2.69)

Noticing that

$$2 \nabla \Psi_0 \cdot \nabla \phi + \Psi_0 \nabla^2 \phi = \nabla^2 (\Psi_0 \phi) - (\nabla^2 \Psi_0) \phi,$$  

(2.70)

The laplacian becomes

$$\nabla^2 (\Psi_0 e^\phi) = \left[ \nabla^2 (\Psi_0 \phi) + \nabla \Psi_0 \cdot \nabla \phi + \nabla^2 \Psi_0 (1 - \phi) \right] e^\phi.$$  

(2.71)

Then, using $\nabla^2 \Psi_0 = -k^2 \Psi_0$, we finally obtain

$$\nabla^2 (\Psi_0 \phi) + k^2 (\Psi_0 \phi) + \Psi_0 \nabla \phi \cdot \nabla \phi = -2k^2 \delta n \Psi_0.$$  

(2.72)
The Rytov series is obtained by writing

$$\phi = \epsilon \phi_1 + \epsilon^2 \phi_2 + \cdots , \quad (2.73)$$

where the perturbation parameter $\epsilon$ has the same significance as above in the Born series treatment. The $\epsilon^0$ term is the homogeneous equation for $\Psi_0$, the next order in $\epsilon$ yields

$$\nabla^2 (\Psi_0 \phi_1) + k^2 (\Psi_0 \phi_1) = -2k^2 \delta n \Psi_0, \quad (2.74)$$

and the next few higher order terms in $\epsilon$ give

$$\nabla^2 (\Psi_0 \phi_2) + k^2 (\Psi_0 \phi_2) = -\Psi_0 \nabla \phi_1 \cdot \nabla \phi_1 \quad (2.75)$$
$$\nabla^2 (\Psi_0 \phi_3) + k^2 (\Psi_0 \phi_3) = -2\Psi_0 \nabla \phi_1 \cdot \nabla \phi_2 \quad (2.76)$$
$$\nabla^2 (\Psi_0 \phi_4) + k^2 (\Psi_0 \phi_4) = -2\Psi_0 \nabla \phi_1 \cdot \nabla \phi_3 - \Psi_0 \nabla \phi_2 \cdot \nabla \phi_2 \quad (2.77)$$

The **Rytov approximation** is the truncation of the Rytov series to the first order in $\epsilon$. Equation (2.74) is the inhomogeneous Helmholtz equation for $\Psi_0 \phi_1$ already encountered above when solving the first Born approximation. Thus, $\phi_1 = \Psi_1 / \Psi_0$, and

$$\Psi_{\text{Rytov}} = \Psi_0 \exp \left[ \Psi_1 / \Psi_0 \right], \quad (2.78)$$

where $\Psi_1$ is the first term in the Born series. Assuming that the incoming wave is a plane wave, this becomes

$$\Psi_{\text{Rytov}} = A e^{ikz} \exp \left\{ \mathcal{F}_{q_\perp} \left[ \frac{\sqrt{2\pi i k^2}}{\sqrt{k^2 - q_\perp^2}} \delta n (q_\perp, q_z = \kappa - k) e^{i(\kappa - k)z} \right] \right\}. \quad (2.79)$$

A few points are worth noting. The linear term in the expansion of the exponent of equation (2.79) gives the Born approximation, (2.65). For this reason, the Rytov approximation is often seen as a higher order solution which captures a portion of the multiple scattering events of the Born series. Another point in favor of the use...
of this method is the fact that the small scattering angle limit \( k \gg |q_\perp| \) is equal to the solution (2.46), so that optically thick specimens seem to be included in its range of validity. Note, however, that (2.79) does not have the form of (2.29), which means that it is approximate even for empty regions of space.

The Rytov solution we have just derived was obtained using only two assumptions: that the fluctuations of \( \delta n \) are slow enough at the scale of one wavelength [condition (2.20)], and that \( |\delta n| \ll 1 \) everywhere. There is apparently no constraint on the optical thickness of the specimen. The exact conditions for the Rytov approximation to be considered valid are still subject to contention. The question is especially important for tomographic applications, where an analytic expression for the scattered field is very useful to infer the structure of the scatterer. Some calculations seem to demonstrate that the range of validity of the Rytov approximation is nearly the same as the Born approximation (Brown Jr, 1966; Lin and Fiddy, 1992) but numerical and experimental evidence show that Rytov gives often considerably better results than Born (Oristaglio, 1985; Devaney, 1986).

From a pragmatic point of view, a simple-minded argument for the limited validity of the Rytov approximation comes from the realization that the exponential in (2.78) and (2.79) is unstable to real-valued arguments. A large enough curvature of the Ewald sphere (that is, departure from the approximation \( \kappa \approx k \)), combined with a large amplitude of \( \delta n \), result in “defocus ripples” around the specimen in \( \Psi_1 \). As a result, the ratio \( \Psi_1/\Psi_0 \) can acquire a large real amplitude which, when exponentiated, make the Rytov approximation fail. The Rytov approximation is thus limited to “optically relatively thin” specimens.
2.2.4 Numerical simulations

In situations where analytic solutions to the Helmholtz equation fail to be valid, approximate solutions can be found numerically. The analytic tools developed in the previous section can be adapted and applied on simulated specimens sampled on a grid. Such methods yield both near-field and far-field simulated diffraction patterns, and can be valid even in cases where the Born and Rytov approximations fail.

The setting will be assumed to be as follows: the scattering strength $\delta n$ is sampled on a $N^3$-voxels 3D array (assumed to be cubic for simplicity). The linear dimension of one voxel is $\Delta x$. Scattering will be assumed to result from an incident plane wave of wavelength $\lambda$, assumed to be smaller than the voxel size, $\lambda < \Delta x$ (typically by one order of magnitude).

The main result of this section is a multislice formula which covers both Born and Rytov validity regimes. The demonstration appearing below is different from the one originally published (Thibault et al., 2006).

A multiscattering implementation

A possible approach is the direct implementation of the multiscattering recursion [eq. (2.60)]. The use of discrete Fourier transforms (FFT) allows for a quick computation of the convolutions in this recursion formula. Let $\psi^{(n)}$ be the $n$-scattering solution, with the fast modulation of the incoming wave removed:

$$
\psi^{(n)} = \psi_0 + \psi_1 + \cdots + \psi_n = (\Psi_0 + \Psi_1 + \cdots + \Psi_n) e^{-ikz}.
$$

Then, starting with $\psi_0 = 1$, Eq. (2.60) becomes

$$
\psi^{(n+1)} = (2\pi)^{3/2} \mathcal{F}^{-1} \left[ \tilde{G}' \cdot \mathcal{F} \left( \delta n \cdot \psi^{(n)} \right) \right],
$$

(2.81)
where
\[
\tilde{G}' = \tilde{G}(\mathbf{q}_\perp, q_z + k) = (2\pi)^{-3/2} \frac{2k^2}{\mathbf{q}^2 + 2kq_z}
\]
(2.82)
is the Green function (2.56) with a translated \(q_z\) component. This function cannot
be simply sampled on an array since the denominator vanishes for points lying on
the Ewald sphere \(|\mathbf{q} + k\hat{z}| = k\). An approximate \(\tilde{G}'\) can be obtained by considering
its form in the hybrid \((\mathbf{q}_\perp, z)\) space. This has already been indirectly computed
above [equation (2.64) with \(\tilde{f} = \mathcal{F}[\delta(z)] = 1/\sqrt{2\pi}\)]:
\[
\bar{G}' = \frac{1}{2\pi} \frac{ik^2}{\sqrt{k^2 - q^2_\perp}} e^{i|k||z| - k|z|}.
\]
(2.83)
For small \(\mathbf{q}_\perp\), this function has a very different behavior on the two sides of the
\(z = 0\) plane: when \(z < 0\), \(\bar{G}'\) has rapid oscillations of spatial frequency close to \(2k\),
while the phase oscillates slowly in the forward direction. Sampling \(\bar{G}'\) on an array
and applying a FFT in the \(z\) direction should give a suitable \(\tilde{G}'\) to use in (2.81).
The \(z < 0\) portion of the array can be set to 0. This corresponds to neglecting
backscattering, a natural assumption since it was determined from the start that
the modeled specimen is smooth at the lengthscale of one wavelength.

A multislice formula

The neglect of the backscattering terms in the method presented above suggests
that an alternative numerical method could be used to achieve the same result by
gradually propagating the wave through the sample in the forward direction. This
approach is called the “multislice method” in the electron microscopy community
(Cowley and Moodie, 1957), and the “beam propagation method” in fields related
to optics (Feit and J. A. Fleck, 1978; Van Roey et al., 1981). Both cases are
essentially identical and, in their original form, apply only to low scattering angle
cases.
Let us consider the Green function formulation of the wave equation solution (2.60), expressed in the hybrid \((q_\perp, z)\) space:

\[
\Psi(z) = G^* z [\delta n *_{q_\perp} \Psi]
\]

\[
= \int_{-\infty}^{\infty} G(z') [\delta n * \Psi](z - z') \, dz'. \tag{2.84}
\]

As has been done above, the next step in the derivation is the neglect of backscattering, which simply entails to setting the lower bound of the integral to 0. The field in the plane \(z + \Delta z\) is

\[
\Psi(z + \Delta z) = \int_{\Delta z}^{\infty} G(z') [\delta n * \Psi](z + \Delta z - z') \, dz'
\]

\[
+ \int_{0}^{\Delta z} G(z') [\delta n * \Psi](z + \Delta z - z') \, dz', \tag{2.85}
\]

where the integral has been divided into two terms. For \(z\) positive, the relation \(\tilde{G}(z + \Delta z) = e^{ikz}G(z)\) holds. Hence, after a simple shift in the integration variable, the first integral is equal to \(e^{ikz}\tilde{\Psi}(z)\). The second integral is \(O(\Delta z)\). For sufficiently small \(\Delta z\), this can be approximated by the value of the integrand at the upper bound:

\[
\int_{0}^{\Delta z} G(z') [\delta n * \Psi](z + \Delta z - z') \, dz' \approx \Delta z \tilde{G}(\Delta z) [\delta n * \Psi](z), \tag{2.86}
\]

whence we finally find that, to first order in \(\Delta z\),

\[
\Psi(z + \Delta z) \approx e^{ikz} e^{ik\Delta z} \left[ \tilde{\Psi}(z) + \frac{1}{2\pi} \frac{i k^2 \Delta z}{\sqrt{k^2 - q_\perp^2}} \left( \delta n(z) * \tilde{\Psi}(z) \right) \right]. \tag{2.87}
\]

In all numerical applications, it is necessary to factor out the \(e^{ikz}\) phase factor, yielding the modified multislice formula (Thibault et al., 2006):

\[
\tilde{\psi}(z + \Delta z) \approx e^{i(k - \kappa)\Delta z} \left[ \tilde{\psi}(z) + \frac{1}{2\pi} \frac{i k^2 \Delta z}{\sqrt{k^2 - q_\perp^2}} \left( \delta n(z) * \tilde{\psi}(z) \right) \right]. \tag{2.88}
\]
As the above derivation demonstrates, the formula (2.88) is an exact expansion, to first order in $\Delta z$, of the complete Born series without back-scattering. It differs from the usual multislice formula by the square root factor of the last term. Assuming small angle scattering, one finds

$$\bar{\psi}(z + \Delta z) = e^{i(\kappa - k)\Delta z} \left[ \bar{\psi}(z) + \frac{1}{2\pi} ik\Delta z (\delta n(z) * \bar{\psi}(z)) \right]$$

$$= e^{i(\kappa - k)\Delta z} \mathcal{F} \left[ \psi + ik\Delta z \delta n(z) \psi(z) \right]$$

$$\approx e^{i(\kappa - k)\Delta z} \mathcal{F} \left[ \psi(z) e^{ik\Delta z \delta n(z)} \right].$$  (2.89)

This is the traditional multislice formula. It differs from the simple equation (2.46) only by the exponential factor, responsible for a part of the propagation effects. The range of validity of this multislice formula depends on many factors and (2.88) should be used for large scattering angles.

### 2.3 The phase problem

The phase problem is the fundamental limitation originating from the well known quantum mechanical statement that any physical measurement is of probabilistic nature, with the probability proportional to the squared amplitude of some quantum state vector. The phase is an entity essential for all phenomena of wave mechanics, and is always hidden to the observer. In some cases it can be deduced via interference effects, but it appears that, as a rule, a portion of the real world is always out of sight of the observer.

In diffraction experiments, the phase represent half of the necessary information to characterize completely a wavefront (neglecting other degrees of freedom such as polarization or spin). Reconstructing the image or the density of a scattering object from a diffraction dataset (through the inversion of equations (2.34) or
(2.66), for instance) thus requires additional information. For our purposes, the phase problem can be defined as follows.

**Definition 1 (The phase problem).**

Given diffraction data (a diffraction pattern or a set of Bragg reflections), find the corresponding phases which satisfy additional constraints.

Unlike what the name suggests, the phase almost never is the quantity we are really interested in. It is simply a mean of extracting information from the wavefield scattered by some object. This seemingly obvious statement allows an important shift of focus which will be exploited in the next chapter, where phase retrieval algorithms will be described.

### 2.4 Diffraction-based imaging techniques

The phase problem is present in many imaging and scattering experiments. That there is a phase to look for implies that the scattered wave is sufficiently coherent to exhibit measurable interference effects. Diffraction-based imaging techniques thus have in common large transverse coherence lengths, compared to the scale of the scattering object. Depending on the type of radiation and the type of specimen, intensity measurements occur sometimes in the far-field, sometimes in the near field, sometimes with the help of optical elements, sometimes without.

This section is concerned with the definition and theoretic aspects of experimental techniques involving the phase problem or the use of coherent sources. Difficulties related to experimental realities (such as noise, partial coherence and radiation damage) will be discussed in chapter 4.
2.4.1 Diffraction microscopy

Diffraction microscopy is an imaging technique based on the reconstruction of a scattering specimen from its diffraction pattern. It can be described as “crystallography without the need of crystals” (Sayre, 2005), or “microscopy without lenses”.

In traditional crystallography, scattered x-rays form sharp peaks in directions determined by the orientation and the symmetry of the crystal. These Bragg peaks are the result of the constructive interference between the very many copies of a specimen in the crystal. This phenomenon makes possible the measurement of diffraction data even with faint and low-coherence x-ray sources, and was therefore the first x-ray diffraction application available to scientists. X-ray diffraction microscopy, like crystallography, amounts to measuring the far-field intensity of radiation scattered off of a specimen. However, the specimen is not required to be a crystal\(^5\) and thus the diffraction pattern does not exhibit systematic signal concentration into Bragg peaks. The absence of constructive interference means that much brighter sources (and more sensitive detectors) need to be used to obtain a useful signal. It also means that, unlike crystallography, it is possible to sample the diffraction pattern on a fine grid, hence providing the possibility of *oversampling* the diffraction pattern.

To a microscopist, diffraction microscopy is simply a method where the task of the optical elements, normally placed downstream from a specimen, is accomplished algorithmically. This provides a way of avoiding experimental limitations caused by imperfect lenses (aberration, limited resolution). On the other hand,

\(^5\)In a sense, it is actually better if it does not have strongly crystalline features, see section 4.2.4.
the computer is provided only half the information with which lenses can work, as only the intensity of the wavefront is measured.

X-ray diffraction microscopy is typically applied to specimens much larger than the unit cell of a macromolecule crystal,\(^6\) thus requiring a transverse coherence length much longer than what is generally needed for protein crystallography. The technique is often called “coherent diffractive imaging” to emphasize this high coherence requirement.

Scattering from a crystal forms Bragg peaks and therefore dictates how a diffraction pattern is sampled. Diffraction microscopy is different in that the sampling of diffraction patterns can be chosen. In most cases, sampling is determined by the experimental conditions; in particular, it depends on the geometry of the pixel detector array used to measure the data. Sampling is a central concept of diffraction microscopy because it is closely related to uniqueness. The question whether only one specimen image can yield — and thus be reconstructed from — a given diffraction pattern is obviously of the utmost importance; a guarantee — or at least very high confidence — of uniqueness is essential for diffraction microscopy to be a valuable tool to scientists.

\textbf{Oversampling}

More than fifty years ago, Sayre (1952a) used a recently published sampling theorem (Shannon, 1949) to argue that the phase problem in crystallography might be solved if a diffraction pattern is sufficiently oversampled. As mentioned in the first section of this chapter, oversampling is possible only if a “signal” is “band-limited”. In diffraction microscopy, this corresponds to having a sufficiently isolated spec-

\(^6\)Although going to ever smaller particle sizes is something many scientists are working on.
imen. The typical length scale of the modulations in the resulting diffraction pattern is determined by the size of the specimen.

In what follows, we will assume that the diffraction data is a two dimensional diffraction pattern, and thus that the specimen is a two dimensional image. Expressions or concepts where extension to three dimensions is not obvious will be mentioned explicitly.

We consider a diffraction pattern measured on a $N_x \times N_y$ grid. The sampling interval (grid spacing) in both directions is $\Delta q_x$ and $\Delta q_y$. This sampling interval alone determines the dimensions of the field of view in direct space:

$$L_x = \frac{2\pi}{\Delta q_x}, \quad L_y = \frac{2\pi}{\Delta q_y}. \quad (2.90)$$

$L_x$ and $L_y$ correspond to twice the “Nyquist frequencies” [see eq. (2.15)]. The direct space image is also sampled on a $N_x \times N_y$ grid, and the sampling interval $(\Delta x, \Delta y)$ is given by eq. (2.14).

In more than one dimension, the expression “band-limited” can be replaced by “supported”. We define the support of a function $f(\mathbf{r})$ as the region of space $\mathcal{S}$ (assumed to be finite) where the function is non-zero. It is often convenient to also define a characteristic function $\chi_{\mathcal{S}}(\mathbf{r})$, unity in the support and zero everywhere else. The area of the support is

$$A_S = \int \chi_S. \quad (2.91)$$

The supported function $f$ sampled on the direct space grid has $N_S = A_S/(\Delta x \Delta y)$ non-zero grid-points (pixels).

When placed in a cartesian system, $\mathcal{S}$ has a maximum extent along each axis $s_x$ and $s_y$. If the phases associated with the diffraction pattern were known, the sampling condition (in the sense of Shannon) would just be [from eq. (2.17) with
\[ B = s_x/2 \text{ or } s_y/2 \]:

\[ \Delta q_x \leq \frac{2\pi}{s_x}, \quad \Delta q_y \leq \frac{2\pi}{s_y}. \quad (2.92) \]

The measured diffraction pattern, however, is the intensity of the signal, and thus has to be measured on a finer grid. As noted at the beginning of this chapter, the Fourier transform of the intensity is called the autocorrelation. The sampling criterion needs therefore to be adjusted according to the autocorrelation support, which is the domain where the function

\[ A_\chi(r) = \int \chi_S(r')\chi_S(r' - r)dr' \quad (2.93) \]

is non-zero. In other words, the autocorrelation support is the set of points \( s = r' - r \) such that \( r' \in S \) and \( r \in S \), or in set theoretic notation, \( S - S \). Thus, the width of the smallest rectangle containing the support of \( A_\chi \) is \( a_x = 2s_x \) (similarly for \( y \)). Hence, a diffraction pattern is oversampled if

\[ \Delta q_x \leq \frac{\pi}{s_x}, \quad \Delta q_y \leq \frac{\pi}{s_x}. \quad (2.94) \]

Put in simple words, this condition is that the intensity be sampled finely enough to avoid aliasing in the autocorrelation.

A different approach to oversampling is based on counting degrees of freedom. Assuming that the image (that is, the function \( f \) sampled on the direct space grid) is complex-valued, \( 2N_S \) independent real variables need to be recovered. The diffraction pattern has \( N = N_xN_y \) measured amplitudes. A reconstruction cannot be uniquely determined from a diffraction pattern if the number of degrees of freedom is larger than the number of constraints, that is, it is undetermined unless

\[ 2N_S \leq N = N_xN_y. \]

Defining the oversampling ratio (Miao and Sayre, 2000) as

\[ \sigma = \frac{A_{\text{FOV}}}{A_S} = \frac{N}{N_S}. \quad (2.95) \]
where $A_{\text{FOV}}$ is the area of the field of view, a necessary condition for a unique reconstruction is $\sigma \geq 2$. This condition is valid for lower- and higher-dimensional systems, as it is true for real-valued images (because the Friedel symmetry decreases by a factor of 2 the number of independent measurements).

The oversampling ratio is a convenient number in the sense that it captures roughly the conditions in which a diffraction pattern was measured. It is also useful to compute the one-dimensional oversampling ratios, simply given by:

$$\sigma_x = \frac{L_x}{s_x}, \quad \sigma_y = \frac{L_y}{s_y}.$$  (2.96)

If both $\sigma_x$ and $\sigma_y$ are greater than two, than the diffraction pattern is oversampled, in the sense of Shannon — that is, the continuous diffraction pattern is completely characterized by the samples (neglecting experimental limitations). Thus, for a rectangular support, the equivalent oversampling ratio is $\sigma = \sigma_x\sigma_y \geq 4$. From now on, we will describe as “Shannon-oversampled” such diffraction patterns, to avoid confusion with the cases where $2 \leq \sigma < 4$.

Figure 2.2 shows an example of an oversampled diffraction pattern. The smallest smooth features on such diffraction patterns are called speckles. The speckle size is inversely proportional to the specimen size, $\sim \pi/s_x, \pi/s_y$.

It should be noted that diffraction patterns do not need to be measured in the far-field (Fraunhofer) limit for the notion of oversampling to be well-defined. The diffraction pattern of a supported specimen, even when $z < \infty$, is also supported, as can be seen in equation (2.40).

**Overdetermination**

To evaluate the information content and the difficulty of an eventual reconstruction, one needs to evaluate the number of degrees of freedom and the number of
Figure 2.2: A simulated diffraction pattern, oversampled, and showing speckles.

constraints, like we did above when defining the oversampling ratio. However, be-cause constraints provided by the intensity measurements are not independent, it is incorrect to assume that increasing oversampling ratio is synonymous with tighter constraints. In fact, increasing the oversampling of a Shannon-oversampled diffraction pattern does not provide any additional information. Considerations about aliasing suggest that a more sensible evaluation of the number of constraints can be done with the autocorrelation. The constraint ratio is defined as

$$\Omega = \frac{1}{2} \frac{A_{\text{auto}}}{A_{S}},$$  \hspace{1cm} (2.97)$$

where $A_{\text{auto}}$ is the area of the autocorrelation support (Elser and Millane, 2007). $\Omega$ is a ratio between the number of independent constraints and the number of degrees of freedom. Since the measured intensity is always real, the autocorrelation
Figure 2.3: Example of support shapes and their constraint ratios.

has Friedel symmetry, and only half of the autocorrelation area has independent values. This explains the factor $\frac{1}{2}$ in the definition. In principle, a reconstruction is possible without additional information if $\Omega > 1$. Shannon-oversampled diffraction patterns always have $\Omega \geq 2$, the equality occurring when the support is convex and centrosymmetric. Figure 2.3 shows examples of support shapes and their constraint ratio.

In crystallography, the “specimen” (the electron density in the unit cell) occupies the totality of the field of view. As a result, the highly aliased autocorrelation, called *Patterson map* in the field (Patterson, 1939), has the same area (or rather, volume) as the support, giving $\Omega = \frac{1}{2}$. This exemplifies the difference between
phase retrieval in crystallography and phase retrieval in diffraction microscopy.

Uniqueness

We have stated in a previous section that $\sigma \geq 2$ is a minimal condition for uniqueness to be possible. However it is not, in general, sufficient.\footnote{In crystallography, the sampling imposed by Bragg peaks is such that the unit cell occupies the totality of the field of view, which gives $\sigma = 1$. Hence, the phase problem can be solved only with additional constraints.} It turns out that, in one dimension, uniqueness is never guaranteed (Wolf, 1962; Walther, 1963; Millane, 1990). The proof of this statement is short enough to be stated here.

Let $f_n$ be a one-dimensional image sampled on an $N$ long grid and with a support of width $M$. Then, its discrete Fourier transform can be written as

$$\tilde{f}(z) = \frac{1}{\sqrt{N}} \sum_{n=0}^{M-1} f_n z^n$$

$$= A \prod_{i=0}^{M-1} (z - z_i), \quad (2.98)$$

where $z$ takes the values $z_m = \exp(2\pi im/N)$, $m = 0, 1, \ldots, N - 1$. The expression on the second line is a factorization of this $M$th order polynomial in $z$. Let us now consider another function $\tilde{f}'$ defined as:

$$\tilde{f}'(z) = |z_0| \frac{z - 1/\overline{z_0}^*}{z - z_0} \tilde{f}(z), \quad (2.99)$$

where $z_0$ is one of the roots of $\tilde{f}(z)$. The factor on the right hand side has a magnitude of 1 (because $|z| = 1$), so that the squared magnitudes of $\tilde{f}$ and $\tilde{f}'$ are identical. Yet, $\tilde{f}'(z)$ is a new $M - 1$ degree polynomial where $z_0$ has been replaced with $1/\overline{z_0}^*$:

$$\tilde{f}'(z) = A |z_0| (z - 1/\overline{z_0}^*) \prod_{i=1}^{M-1} (z - z_i)$$

$$= \sum_{n=0}^{M-1} g_n z^n, \quad (2.100)$$
and is the Fourier transform of a (different) direct space image \( g_n \) with the same support (since the sum is also up to \( M - 1 \)). This completes the proof, since we have found two different images with the same diffraction pattern. Of course, one can use any combination of factors in (2.99), and not only the one involving \( z_0 \). This results in a \( 2^M \) degeneracy in the possible solutions.

The key concept on which this proof is based is the possibility of factorizing any one-dimensional polynomial. The proof also used the assumption that the support is connected. Uniqueness can be proven for disconnected supports. Hence the condition \( \omega > 1 \) still applies, without modification. Fortunately, in higher dimensions, the set of factorizable polynomials is of measure zero, which allows to guarantee that most solutions are unique (Bruck and Sodin, 1979; Bates, 1982) with Shannon-oversampled diffraction patterns.

### 2.4.2 Other techniques

We complete this section with a quick overview of other imaging techniques which make use of diffraction data.

**Fourier holography**

Holography is a method which encodes phase information in a scattered wave through its interference with a reference wave (Gabor, 1948). In Fourier holography, the reference is typically a point-like source in the same plane as the specimen, and the interference pattern is measured in the far-field (Stroke and Falconer, 1964; McNulty et al., 1992; Howells et al., 2000; Eisebitt et al., 2004). The reference wave is sometimes formed in the specimen plane with a focussing element placed upstream, but can also simply originate from a strongly scattering element (a gold
ball, or a pinhole for instance) placed directly in the specimen plane.

The important benefit of this method is that it can reduce phase retrieval to a single Fourier transformation operation. Let $\psi(\mathbf{r})$ be the exit wave formed by the specimen, and $\eta(\mathbf{r} - \mathbf{r}_0)$, the reference wavefield, in the same plane. The offset $\mathbf{r}_0$ is the distance between the centers of the specimen and of the reference wavefield. These two waves add coherently to form a diffraction pattern, measured in the far-field. If the oversampling condition is satisfied (for the sum $\psi + \eta$), the Fourier transform of the measured intensities (the autocorrelation) is

$$A(\mathbf{r}) = (\psi + \eta) \circledast (\psi + \eta) = (\psi \circledast \psi)(\mathbf{r}) + (\eta \circledast \eta)(\mathbf{r}) + (\psi \circledast \eta)(\mathbf{r} + \mathbf{r}_0) + (\eta \circledast \psi)(\mathbf{r} - \mathbf{r}_0).$$

(2.101)

The first two terms in (2.101) are the autocorrelations of the specimen wave and the reference wave, superimposed at the origin. The two other terms are the complex-conjugate mirrors of each other, and located on opposite sides of the central autocorrelation. When $\eta$ is localized enough to be approximated with a point, these two last terms are proportional to $\psi(\mathbf{r} + \mathbf{r}_0)$ and $\psi^*(\mathbf{r} - \mathbf{r}_0)$, thus giving the “reconstruction” without much effort. The two twin images are not overlapping if the “holographic condition” is satisfied, that is if $|\mathbf{r}_0| > \frac{3}{2}a$, where $a$ is the extent of the support in the direction of $\mathbf{r}_0$.

The ease with which the phase problem is solved makes Fourier holography a very attractive experimental method. It has, however, some drawbacks.

First, it may involve a more complicated sample preparation than “classical” diffraction microscopy, as a strong scatterer has to be placed in the vicinity of the specimen. Also, because the reference scatterer has to be placed a good distance
Figure 2.4: Simulated Fourier holography. (a) The specimen is isolated and a point-like source is placed at a large enough distance from it (satisfying the “holographic condition”). (b) The resulting diffraction pattern. The inset shows high spatial frequency modulations caused by the interference of the two distant objects. (c) The Fourier transform of the diffraction pattern, giving the autocorrelation of the specimen. The two copies of the specimen result from the cross-correlation of the specimen image with the point source.
away from the specimen, the field of view needs to be much larger than the over-
sampling condition dictated by the specimen size, thereby limiting the resolution
that can be achieved. The resolution of a reconstruction, if it is obtained solely
with the Fourier transformation of the measured intensities, is also limited by the
finite support of the reference wave in the specimen plane.

Phase retrieval algorithms are likely to be useful even on datasets meant to be
used for Fourier holography, thus allowing reconstructions in intermediate cases,
where either the reference wave is extended or the holographic condition is not
satisfied. Even then, reconstructions should be relatively easy: the Fourier holo-
graphy configuration is such that the overdetermination ratio is always high (see
Figure 2.3(f) for instance).

**Multiple-measurements methods**

Many methods involve the use of multiple measurements to overconstrain the phase
problem. In all cases, a single specimen is exposed many times, with a component
of the experiment setup varying in a controlled way. This idea is similar to multiple
measurement methods used in crystallography (multiple wavelength anomalous
data, multiple isomorphous replacement, radiation-induced phasing).

Perhaps the oldest application of this type is the image reconstruction from
focal series, as is commonly used nowadays in high resolution transmission elec-
tron microscopy (Schiske, 1973; Misell, 1973; Saxton, 1978). This method simply
consists in measuring a series of images with different defocus values. Although
this method does not involve the measurement of far-field diffraction patterns,
it has the same flavor as the method for which the Gerchberg-Saxton algorithm
was devised. Commonly used algorithms with focus series are typically based on
maximum-likelihood concepts (Coene et al., 1996), but there also exists iterative algorithms (Allen and Oxley, 2001). A very similar method, proposed by Maleki and Devaney (1994), involves reconstructing the phase information from two near-field holograms (with a plane reference wave).

Faulkner and Rodenburg (2004) suggested the use of many diffraction patterns formed by a specimen and a “window” placed right behind it (downstream from it). Overlap of the windows provides additional information on the visible portions of the specimen.

A similar approach (called “PIE”, for “ptychographical iterative engine”) used successfully with optical light (Rodenburg et al., 2007a) and with hard x-rays (Rodenburg et al., 2007b) places instead a window upstream from the specimen, which is thus illuminated with a slightly defocussed image of the window. Yet again, the overlap of neighboring regions makes the data highly redundant.

In the same spirit of shaping the incident beam, Allen et al. (2001) and McBride et al. (2005) suggested using well controlled wavefront aberrations (for instance in a transmission electron microscope). Similarly, Nugent et al. (2003) also proposed using incident wavefronts with various curvatures (see also Nugent et al., 2005).

Finally, Mentes et al. (2002) introduced the possibility of imaging magnetic domains through the controlled tuning of the wavelength of the incident beam – an idea very close to the MAD method of crystallography.
Chapter 3

Iterative algorithms

The phase problem, as formulated in the previous chapter, identifies the sought solution as a set of $N$ phases. Owing to the periodicity of these variables, the search space, in this formulation, is therefore an $N$-dimensional torus. Original approaches to solving the phase problem in crystallography are formulated in this language. In this scheme, Fourier transforms become very complex operations, nevertheless required because additional constraints lie in direct space.

A different approach is very common when one is confronted with many non-compatible constraints; by allowing the search space to be expanded, the formulation of constraints can be simpler and lends itself to more diverse applications. The use of Lagrange multipliers (Lagrange, 1813) in traditional optimization is probably the most eloquent example of this philosophy. In the case of phase retrieval, one can simply expand the search space to include not only the phases but also the magnitudes of the object’s Fourier transform, yielding a simple $N$-dimensional complex-valued Euclidean search space. A completely equivalent embedding is the direct space counterpart, where the search space contains the $N$ pixels (or voxels) on which the reconstructed object is sampled. The problem is now expressed as the search for the unique element of the space that satisfies two constraints at the same time: one in Fourier space (the amplitudes are fixed by the measurements), one in direct space.

In section 3.1, we describe how the reformulation of the problem allows for the use of formalism and ideas more geometrical in nature: that of constraint sets and projections onto these sets. The next section introduces the difference map.
algorithm and rederives its most important properties. Section 3.3 surveys all the reconstruction algorithms currently in use.

3.1 Formalism

3.1.1 Constraint sets

As we have noted above, the phase problem can be seen as belonging to a much larger family of problems defined as follows:

Definition 2 (The two-constraints problem). Given two constraint sets $C_1$ and $C_2$, find $x$ such that $x \in C_1 \cap C_2$.

In all practical applications, we will assume that $C_1$ and $C_2$ are subsets of a finite-dimensional Hilbert space (or Euclidean space), $\mathcal{E}$. Obviously, one cannot tell from this formulation alone if the problem is hard to solve or not – it all depends on the structure of the constraint sets. For instance, if $C_1$ and $C_2$ are linear constraint spaces, then $C_1 \cap C_2$ is also linear and, as such, finding the solution (provided it exists) does not represent a computational challenge – apart from possibly prohibitively large memory requirements to do the computation. On the other hand, $C_1$ and $C_2$ could be large sets of random points with no structure, possibly leading to complete enumeration as the only solution strategy. Fortunately, most interesting problems lie somewhere in between.

The special case where all constraint sets (possibly more than two) are convex, known as the convex feasibility problem, has been studied in details for a long time (von Neumann, 1950; Youla and Webb, 1982). Convexity allows the use of

$^{1}$Recall that a set $\mathcal{C}$ is convex if $x, y \in \mathcal{C}$ implies that all points in the segment $xy$ also belong to $\mathcal{C}$. 
powerful mathematical theorems, and general conclusions – about the convergence rate of algorithms for instance – can be drawn without explicit knowledge of the individual constraint sets. This type of problem has numerous applications in medical imaging and various other optimization problems. For a good review, see the article by Bauschke and Borwein (1996) and references therein. Ibrahim Sezan (1992) also offers a review more focussed on image recovery applications.

By contrast, very few general results are known in the case of non-convex constraint sets, and the gathered knowledge is problem-specific and often of an empirical nature. For some years, approaches taken for phase retrieval have been developed independently from the convex set framework and formalism, and in a more pragmatic spirit. An overview of these algorithms, along with more recent ones, is given in section 3.3.

3.1.2 Projections

Although many general results from the convex-set problem break down with non-convex sets, it is still useful to use similar concepts. One of the most important entities is the projection onto a set.

**Definition 3** (Projection). Given a closed constraint set \( C \in \mathcal{E} \), a projection \( P_C \) onto this set is a map which, to every \( x \in \mathcal{E} \) returns a point \( P_C(x) = y \in C \) and such that \( \|x - y\| \) is minimized.\(^2\)

In other words, a projection gives the closest point of a set. It can be shown that if \( C \) is convex, then \( P_C(x) \) is unique for any \( x \). For non-convex sets, some points can have multiple equally distant projections, so that Def. 3 is ill-defined for these \( C \) needs to be closed, since otherwise the minimum may not exist. Generally, this is of little relevance in actual applications.
points. This is rarely a problem in practical applications since the probability of having to compute a projection for a point of this type is zero most of the time.

Projections are idempotent: $P \circ P = P$. This, however, can not be seen as a defining property, since idempotency does not guarantee that the distance minimizing property is satisfied.

We can now restrict the type of problem we are interested in solving by requiring that projections onto both constraint sets are easy to compute. The meaning of “easy” can be in terms of computational complexity (for instance that the computation be of polynomial time) or simply that the effective calculation time is short. It is natural to see the projections onto the constraint sets $P_{C_1}$ and $P_{C_2}$ (abbreviated from now on as $P_1$ and $P_2$) as the building blocks of an algorithm looking for a point in $C_1 \cap C_2$. Because $x \in C \iff x = P_C(x)$, all solutions to the two-constraints problem satisfy the condition

$$x_{\text{sol}} = P_1(x_{\text{sol}}) = P_2(x_{\text{sol}}). \quad (3.1)$$

Implicit in the formulation of projections is the definition of the search space itself. The way an image (or a three dimensional density sampled on a grid) is embedded mathematically is very straightforward and in general does not require much refinement. In other problems, however, the structure of the search space (and, therefore, of the constraint sets therein) can critically affect the ease of implementation and efficiency. That the sought solution can be expressed in diverse ways should always be kept in mind when formulating a new problem.

### 3.1.3 Iterative maps

The formulation of each projection is problem-specific; to be general, an algorithm solving the two-constraints problem has to be expressed in terms of these
projections only (plus possibly a small number of adjustable parameters). Most algorithms are iterative\(^3\); this is especially appropriate when searching in a very large space. Algorithms consist in repeatedly applying a map

\[
x_{n+1} = F(x_n)
\]  

(3.2)

until it reaches a fixed point

\[
x^* = F(x^*).
\]  

(3.3)

It is then said that the iteration converged. A useful algorithm needs to have its fixed points related to the solution of the problem (3.1), although it is not necessary that the fixed point itself be a solution. Iterative techniques are ubiquitous in the mathematical and physical sciences. In mathematics, iterative maps of the form (3.2) are often viewed as dynamical systems. When discussing the properties of algorithms, it will therefore be natural to borrow concepts and nomenclature from this field. Concepts (such as “trajectories”, “attractor” or “ergodicity”) will be introduced below when we describe in detail one of these algorithms, called the difference map.

### 3.2 The difference map

The difference map (Elser, 2003a) is defined as:

\[
x_{n+1} = x_n + \beta D(x_n),
\]  

(3.4)

where

\[
D(x) = y_2 - y_1,
\]  

(3.5)

\(^3\)This is to be expected since the structure of complicated constraint sets can be probed by a numerical tool only through multiple calls to the corresponding projection.
with
\[ y_1 = P_1[(1 + \gamma_2)P_2(x) - \gamma_2 x], \quad y_2 = P_2[(1 + \gamma_1)P_1(x) - \gamma_1 x]. \quad (3.6) \]
The projections \( P_i \) are defined as above, the two \( y_i \) are called the estimates (for reasons soon to become clear), and \( \beta, \gamma_1 \) and \( \gamma_2 \) are complex (but generally taken to be real) parameters.

Starting from some initial point \( x_0 \) in the search space, the map (3.4) is applied repeatedly until a fixed point is reached. Then, one has:
\[ x^* = x^* + \beta D(x^*), \quad (3.7) \]
which implies that both estimates are equal:
\[ y_2(x^*) = y_1(x^*), \quad (3.8) \]
and the solution has been found, since (3.8) states that a point on \( C_1 \) also belongs to \( C_2 \). Hence, at convergence we have
\[ x_{sol} = y_1 = y_2 \quad (3.9) \]
Note that one can easily obtain (3.1) from (3.9) using the idempotency of the projection operators. This property should not be overlooked: in general, a fixed point is not the solution. The difference map often has many fixed points, even though the solution itself is unique.

Convergence can be monitored with the difference map error, defined as the norm of the displacement:
\[ \varepsilon_n = \|x_{n+1} - x_n\| = |\beta| \|y_{2,n} - y_{1,n}\|. \quad (3.10) \]
At any point in the iteration process, \( \varepsilon_n/|\beta| \) is the distance between the two estimates. The smaller the error, the closer to each other are the estimates (hence their name).
An important property of the difference map is its ability to avoid *traps*, that is, regions of the search space where the distance between the two constraint spaces is a local minimum. Let $c_1 \in C_1$ and $c_2 \in C_2$ be two points such that $\|\Delta c\| = \|c_1 - c_2\|$ is a local minimum. Assuming the iterate is momentarily attracted to this region of the search space, it will slowly drift away as, locally, $\|x_{n+1} - x_n\| > |\beta| \|\Delta c\|$, whence $x_n \rightarrow x_n + \beta \Delta c \rightarrow x_n + 2\beta \Delta c \rightarrow \ldots$ until other more remote regions of either constraint sets divert the iterate trajectory. Because this argument is valid only locally, the iterate could still be trapped by being attracted again by the same local minimum. A possible instance of this problem can be found in the case of *phase vortices* sometimes observed in two-dimensional phase retrieval problems (Robinson, 2003, as cited in Peele et al., 2004).

### 3.2.1 Discussion

**Convergence**

How and why the difference map works (as well as other similar algorithms, for that matter) is only incompletely understood. Obviously, little can be said because of the incredible variety of problems it can tackle. One could expect more formal results if the class of constraint sets is reduced, as in the convex sets case. In the last four years, experiments on various problems have lead to some intuitive understanding of the search process. A rough but probably accurate picture is as follows (Carrier, 2002). All iterates, irrespective of the initial conditions, seem to follow a trajectory on a low-dimensional subspace of the search space, analogous to *strange attractors* encountered in the qualitative description of dynamical systems. The analogy is not perfect, since fixed points lying on this subspace have a bassin of attraction of their own which has the effect of terminating the progress of
the iterate. Because the dynamics are in general chaotic, the number of iterations needed to reach convergence is mostly explained by the ratio of the volume of the basin of attraction to the volume of the search subspace. In favor of this description is the observed exponential distribution of iteration numbers needed to reach a solution, indicative of a “memory-less” hoping process. In this picture, tuning the parameters $\beta$ and $\gamma_i$ corresponds to reducing the volume (or dimensionality) of the search subspace while ensuring that the fixed points have a large basin of attraction. In instances where this picture holds, the system appears to be ergodic, that is, the search subspace is connected and the solution can be reached from any starting point.

An optimal choice for the parameters $\beta$ and $\gamma_i$ are also, to a very large extent, problem dependent. In the picture given above, the overall convergence rate of the algorithm for a given problem depends both on the dimensionality of the effective search subspace and on the “attractiveness” of the fixed points. The former condition is a global property, whereas the latter is more local and, as such, much easier to tackle. The approach taken by Elser (2003a,c) is to attempt the optimization of the free parameters based on the properties of the projection operators in the vicinity of fixed points. Assuming that the constraint spaces are locally smooth, these projections can be approximated by their Taylor expansion up to linear order. The resulting projections are affine transformations, that is, linear transformation followed by a translation. Assuming that the two projections
are locally orthogonal, one finds that the difference map converges fastest if

\[ \gamma_1 = \beta^{-1}, \quad \gamma_2 = -\beta^{-1}. \] (3.11)

With these parameters, convergence takes exactly one iteration if \( C_1 \) and \( C_2 \) are orthogonal. With this choice of parameters, the difference map also becomes nicely symmetric, with the interchanging of \( \beta \) with \( -\beta \) being equivalent to switching \( P_1 \) with \( P_2 \).

**Incompatible constraints**

A crucial feature of the difference map, shared by other iterative methods, is its robustness in instances where constraint sets do not intersect. This situation, although formally ill-posed, is as a rule always met in real applications. As mentioned above, the difference map has this useful characteristic of getting attracted by near solutions without being trapped by them. This continuous progress, interrupted only if an intersection point is found, is an essential feature when looking for intersection of non-convex constraint sets, as these problems are generally plagued with many regions of space where the distance between the constraint sets comes to a local minimum. As the two constraint sets are pulled closer and closer, the steps of iterates close to the near-solution decrease in the same proportion, and iterates thus spend a longer and longer time in the vicinity of the near-solution.

Because the iterate is not trapped, it sometimes happens that it moves away from a near-solution enough to show a sudden increase in the error [eq. (3.10)]. This “wandering” behavior is sometimes seen as inconvenient (Russel Luke, 2005).

\(^4\)This assumption is especially good when considering ensembles of low-dimensional subspaces in a high-dimensional space. For entropic reasons, two low-dimensional planes intersecting in a high-dimensional space have high probability of being nearly orthogonal.
In the view presented here, sudden rises of the error simply indicate that the search is not over and that the iterate has headed to another region in the search space. Constraint sets that are too far apart create this type of wandering, persisting indefinitely. In these cases, it may be necessary to relax the constraints. This subject certainly needs further investigation as there is currently no simple way to characterize the degree of incompatibility between constraint sets.

In common usages, the progress of the difference map is fairly well-behaved: the error generally starts at a high value and reaches a plateau during which the main search occurs. Progress is shown by a sudden drop (or many successive drops) in the difference map error. Finally, the error reaches another lower plateau where the qualitative appearance of the iterate does not seem to improve anymore (see figure 3.1). With simulated data, this last plateau almost always corresponds to the solution, and further fluctuations are caused by the incompatibility of the constraints. Often, when the data used in the definition of the constraints is of good quality, the iterate is “trapped” in the vicinity of the solution, with only small wandering. This trapping is certainly different, more global, than it is with simple alternating schemes (see the “error-reduction” algorithm below). Since, however, the dynamics never halts, one needs to find a new operational definition of “solution”.

The approach that we have proposed is that of using the average of many estimates. If the search were not chaotic and ergodic, the right averaging ensemble would be the many estimates resulting from different (randomly selected) starting points. However, owing to the chaotic and ergodic nature of the search, this “ensemble average” can be replaced with a “time average”, where estimates are taken at regular intervals on a single run, the only requirement being that the time
Figure 3.1: A typical difference map error plot showing the characteristic plateaus before and after convergence.

interval be large enough to allow consecutive estimates to be uncorrelated.\(^5\)

The benefit of the definition of solution in terms of an average is that it is reproducible: two different iteration sequences with different starting iterates should give the same average (in the limit of very long iteration times). In the context of imaging, it tends to lower the resolution of a reconstruction by averaging out potentially misleading artifacts that vary quickly from one iterate to the other (see section 4.1.2).

Averaging is a useful definition, but it is probably not the best one can do; the dynamics near the solution (in the last plateau) are probably biased by the very structure of the difference map, and the average is likely to be influenced by this bias. As an operational definition, it is probably the simplest and most reliable

\(^5\)This condition can be hard to meet when some "slow modes" are present. See sec. 4.1.3 for a known situation where this occurs.
available for now.

**Pseudo-projections and non-optimal projections**

In the general presentation on the difference map above, it was assumed that the projections were exact and distance-minimizing. In real applications, it may be very difficult to define such projections. This leads to definitions of *pseudo-projections* and *non-optimal projections*.

A pseudo-projection is an operation (not a projection in general) designed to behave like a true distance-minimizing projection at the solution (and in its vicinity). In general, a pseudo-projection may not project on a well-defined constraint set, and may not be idempotent. However, convergence of the difference map, if it occurs, still means that the solution is obtained because of the true projection behavior at the solution. Apart from the fact that the search may be less efficient (a highly problem-dependent question), convergence of the difference map should always be possible if one requires that the pseudo-projection linearized at the solution be equal to the true projection. For instance, a relaxed projection (defined below in section 3.3.1)

$$ T(x) = x + \alpha (P(x) - x) $$

(3.12)

is such that $T(P(x)) = P(x)$, but is not a pseudo-projection since its linear expansion around a point $x_0 = P(x_0)$ is not equal to the expansion of $P(x)$ around the same point. This requirement is needed to ensure that the parameters $\gamma_1, \gamma_2$ make the map contracting in the vicinity of the solution.

A non-optimal projection is a projection that is not distance-minimizing. An example of such a projection is the thresholding operation, often used in dynamic support determination and discussed on page 75.
Various empirical evidence shows that the use of non-optimal projections in the difference map can still lead to convergence. However, behavior can also be unpredictable and it is preferable, when possible, to use distance-minizing projections.

**Extension to more than two projections**

The difference map does not lend itself to problems formulated in terms of more than two constraint spaces. How to generalize the algorithm’s structure for these situations is still unclear. At the core of the algorithm is the difference (3.5), which in itself ensures that a solution is attained when the iteration reaches a fixed point. It seems that the same principle cannot be expressed as simply for three or more constraint spaces.

Many-constraint problems can still be tackled with the current formulation of the difference map, through an expansion of the search space. If $n$ is the number of constraint sets originally embedded in the space $\mathcal{E}$, the new iterate $X$ is a vector in the search space $\mathcal{E}^n$, that is, it is the outer product of $n$ vectors of $\mathcal{E}$:

$$X = x^{(1)} \otimes x^{(2)} \otimes \cdots \otimes x^{(n)}$$  \hspace{1cm} (3.13)

The iterate is now a higher dimensional vector formed by the concatenation of duplicates, or replicas (Gravel, 2007). Then, one defines the “average projection”, $P_A$ and the “direct product projection”, $P_D$ as follows:

$$P_A(X) = \bar{x} \otimes \bar{x} \otimes \cdots \otimes \bar{x},$$  \hspace{1cm} (3.14)

$$P_D(X) = P_1(x^{(1)}) \otimes P_2(x^{(2)}) \otimes \cdots \otimes P_n(x^{(n)})$$  \hspace{1cm} (3.15)

where $\bar{x} = \frac{1}{n} (x^{(1)} + x^{(2)} + \cdots + x^{(n)})$. The meaning of these two projections is simple: $P_D$ projects, at once, each of the $x^{(i)}$ onto its own constraint set, while
$P_A$ enforces the additional constraint that all replica be equal. If a fixed point is reached, all replicas are equal, and each of them also satisfies its individual constraint. This method was suggested by Bauschke et al. (2004) for convex sets algorithms. We will describe below how this framework can be used for phase retrieval, in applications where the information from many datasets need to be combined.

When this projection reformulation is used with only two constraints, there results a “disguised” version of the difference map. For instance, the mapping

\[
x_{n+1} = \frac{1}{2}(x_n - y_n) + P_1(y_n)
y_{n+1} = \frac{1}{2}(y_n - x_n) + P_2(x_n)
\]

(3.17)
is a special case of the difference map in this formalism. Note however, that the conventional choice of the parameters $\gamma_1$ and $\gamma_2$ is no longer optimal as $P_A$ and $P_D$, as defined above, are not orthogonal at the solution.

3.3 Other algorithms

3.3.1 Algorithms for convex sets

As briefly noted above, there exists already a considerable body of work on the convex feasibility problem. Most methods successful with convex sets fail with non-convex sets; nevertheless, many important ideas and concepts remain important even in the latter situation. Diverse iterative algorithms found in the literature can be expressed in a unified form (Flåm and Zowe, 1990; Bauschke and Borwein, 1996). Although it is more common to see problems formulated for an arbitrary number of intersecting sets, we will mostly be interested in the two constraints.
The general iteration then becomes

\[ x_{n+1} = \lambda_1^{(n)} T_1(x_n) + \lambda_2^{(n)} T_2(x_n), \]  

(3.18)

where \( T_i \) are the “relaxed projections”, defined as

\[ T_i(x) = x + \alpha_i (P_i(x) - x). \]  

(3.19)

The case \( \alpha_i < 1 \) is often called “under-relaxed”, while it is “over-relaxed” if \( \alpha_i > 1 \). In the case \( \alpha_i = 2 \), \( T_i \) is sometimes called the “reflector”, as it gives the mirror image of \( x \) with respect to its projection on \( C_i \). The occurrence of relaxed projections should be seen as very natural: it is the only origin-independent linear combination of a given point and its projection.

With some assumptions on the weights \( \lambda_i^{(n)} \), the relaxation parameters \( \alpha_i \), and the projection operators \( P_i \), Flåm and Zowe (1990) have shown that the sequence formed by the iteration of (3.18) always converges to a point of \( C_1 \cap C_2 \). Equation (3.18) covers many of the early algorithms used to solve either linear systems of equations, system of linear inequalities, or other convex sets problems. For instance, when \( \alpha_1 = \alpha_2 = 2 \), \( \lambda_1^{(n)} = \lambda_1 > 0 \), \( \lambda_2^{(n)} = \lambda_2 > 0 \) with \( \lambda_1 + \lambda_2 = 1 \), the algorithm becomes of the form first described by Cimmino (1938) for solving linear systems of equations:

\[ x_{n+1} = \lambda_1 (2P_1(x_n) - x_n) + \lambda_2 (2P_2(x_n) - x_n). \]  

(3.20)

An interesting aspect of this scheme, which becomes apparent only when searching for the intersection of many more than two constraint sets, is that it can easily be implemented for parallel computing.

When \( \alpha_1 = \alpha_2 = 1 \), and the variables \( \lambda_i^{(n)} \) take in turn the values 0 and 1, one
gets the very simple alternating scheme

\[ x_{n+1} = P_1(x_n), \]
\[ x_{n+2} = P_2(x_{n+1}), \]  

attributed to Kaczmarz (1937) and von Neumann (1950) in the case of linear constraints and to Brègman (1965) for convex sets. This simple idea was often rediscovered later on, and is still often used even in the case of non-convex sets (Gerchberg and Saxton, 1972; Fienup, 1978), especially when there are more than two projections (Allen et al., 2001; Nugent et al., 2003), a very little studied situation up to now (see section 3.2.1 above). It appears that having highly redundant data makes the alternating scheme less likely to stagnate. In crystallography, the method known as “density modification” typically uses alternating projections (Shiono and Woolfson, 1992; Read and Kleywegt, 2001). The method is almost always used in a late stage of the reconstruction and is not used to search for a solution but rather for refinement of the electron density map.

Some algorithms for the convex optimization problem do not belong to the family described by eq. (3.18). Among them, Dykstra and Douglas-Rachford algorithms have been related to algorithms used in phase retrieval (Bauschke et al., 2002). The difference map can be seen as a generalization of the latter.

### 3.3.2 Common phase retrieval algorithms

The two main lineages of phase retrieval algorithms are those coming from the crystallography and imaging communities. Although both cases are devised to solve the phase problem, their scope and their flavor is very different.

Nowadays, crystallographers have a very large and varied toolbox for tackling the phase problem. Phasing algorithms in this field are commonly divided into two
categories: direct and indirect methods. Direct methods are the algorithms that solve directly the positions of atoms in the unit cell given only a single dataset. These methods were pioneered in the early 50’s by Hauptman and Karle (1953) and Sayre (1952b). Direct methods are used on high resolution datasets (roughly above 1.5 Å) and rely on the atomistic character of the reconstructed electronic density. They can also be applied to lower resolution “heavy-atom” datasets to first identify the positions of the heavy atoms in the unit cell. Indirect methods encompass a large group of methods, which often use more than one dataset to provide a stronger constraint on the reconstruction, especially when resolution is too low to enforce atomicity. Millane (1990) gives a good overview of phase retrieval algorithms in crystallography, in a wider context including imaging techniques.

**Gerchberg-Saxton algorithm**

The first phase retrieval algorithm widely used in imaging was the Gerchberg-Saxton algorithm (Gerchberg and Saxton, 1972). This algorithm is an alternating algorithm of type (3.21) between two non-convex set. Created for electron microscopy measurements, it attempts to recover the phase of a direct space complex image using two intensity measurements (one in direct space, the other in Fourier space). Although very simple, this algorithm played an important role as a precursor. In particular, it introduced the “modulus projection” (see projection 1, page 66) in a non-crystallographic context.

**Error-reduction and hybrid input-output algorithms**

The most significant contribution to the field to date was brought by Fienup, who introduced new algorithms in the late 70’s, among which one (HIO) is still
widely used today (Fienup, 1978, 1982). These algorithms were developed to solve the phase problem given a diffraction pattern, with the additional constraint that the reconstructed object has a sufficiently small support (See section 2.4.1). Like the difference map, all three algorithms can be expressed in terms of projections onto constraint sets. One projection enforces the consistency with the measured diffraction pattern (see projection 1 below), while the other projects onto the set of all (possibly positive) images with a known support (see projection 8 below).

The “Error-reduction algorithm” (or ER) is a direct application of Gerchberg and Saxton’s algorithm with the direct space constraint replaced with a non-negativity (and possibly support) constraint. It is therefore of the alternating type [Eq. (3.21)]. It’s name comes from the fact that the map is non-expanding: the distance between an iterate and the next $\|x_{n+1} - x_n\|$ is non-increasing, a property previously observed by Gerchberg and Saxton (1972). Like the Gerchberg-Saxton algorithm, this algorithm is not used often anymore as it is known to be prone to stagnation.

The hybrid input-output algorithm (HIO) is one of the most popular algorithms in the imaging field. It is defined as

$$x_{n+1} = x_n + \beta \left[ P_S \left( (1 + \beta^{-1}) P_M(x_n) - \beta^{-1} x_n \right) - P_M(x_n) \right], \quad (3.22)$$

where $P_M$ is the projection onto the constraint set of known Fourier magnitudes, and $P_S$ is the support projection.

Manifestly, HIO is a special case of the difference map with $\gamma_1 = -1$ and $\gamma_2 = \beta^{-1}$. Consequently, it shares with the difference map most of its properties, such as its propensity to avoid traps, and the fact that fixed points are in general not the same as the solution. Although Fienup mentions that “the input (...) no longer is an estimate of the object” (Fienup, 1982), this last point might have been
overlooked by some authors, making necessary the use of additional error-reduction steps to force the iterate itself to converge to the solution.

In his original article, Fienup (1982) suggested that an alternating scheme between ER and HIO seemed to yield the fastest convergence rates. To our knowledge, there exists no serious study of this claim, although many practitioners have adopted it. This type of periodic “algorithm switching” has never been necessary for reconstructions presented in this thesis. In the difference map (and thus in HIO), the iterate not being identified with the estimate of the solution suggests a picture where the iterate is “hovering” over local minima in the distance between the constraint spaces. The effect of the alternating map corresponds to plunging into the nearest local minimum to check if the intersecting point might lie there. In this sense, the HIO/ER periodic combination appears to be a “watched kettle” technique, which in the end does not seem to accelerate the process.

The relaxed averaged alternating reflections algorithm

In 2005, Russel Luke (2005) introduced the “relaxed averaged alternating reflections” (RAAR) algorithm, a descendent of the “hybrid projection-reflection” algorithm\(^6\) (Bauschke et al., 2003). This algorithm is given by

\[
x_{n+1} = x_n + \beta [P_S (2P_M (x_n) - x_n) - P_M (x_n)] + (1 - \beta) (P_M (x_n) - x_n).
\]  

(3.23)

When \(\beta = 1\), this algorithm is equivalent to HIO (with \(\beta = 1\)). Russel Luke studied the convergence behavior of this algorithm when applied to convex sets, in order to gain insights for the non-convex case.

\(^6\)The HPR method turned out to be a special case of the difference map.
The charge flipping algorithm

Introduced by Oszlániy and Sütő (2004), the charge-flipping algorithm is so simple that, in the words of the authors, “it is surprising that it works at all”. It consists in the iteration of the map

\[ x_{n+1} = P_M (2P_A(x_n) - x_n), \]  

(3.24)

which is a special case of the form described above [eq. (3.18)]. In its original formulation, the projection \( P_A \) enforces the “atomicity” constraint through a simple thresholding operation (akin to projection 9).

In actuality, the charge-flipping algorithm is found to be seldom successful and using it in place of more established algorithms is somewhat controversial.

Discussion

In view of all algorithms now available for phase retrieval, the difference map is currently the one which has been applied for the broadest range of applications, making it a “meta-algorithm”. The rigorous and firmly established formalism of constraint sets and projections is certainly a main advantage of this algorithm.

There exists few comparisons between these algorithms (Marchesini, 2004; Williams et al., 2007). In simple, well-controlled, phase retrieval experiments, the conclusion of these works is that the hybrid input output (and hence the difference map as well) is the most efficient algorithm. This seems to be corroborated by the popularity of hybrid input output in the recent literature. More thorough comparisons are a hard task for many reasons. First, the original range of application of these algorithms overlaps only weakly, so that not very many benchmarks can be considered legitimate applications for all algorithms. Also, aspects to be
considered in the evaluation of an algorithm’s “performance” are varied and not always easily defined. Of course one wants an algorithm which converges quickly; other important issues are (1) the behavior of the algorithm when the constraint sets are incompatible and, (2) the algorithm’s robustness when used in various contexts. These issues have great relevance in real-life applications and will need to be explored in much greater depth in the future.

3.4 Projections

In the preceding section, we have presented a general framework to solve various problems, as long as they can be formulated in terms of finding the intersection of two (or even more) constraint sets. The power of this formalism is that the problem-dependent part of the algorithm is contained solely in the definition of these constraint sets — or more precisely, for computational purposes, in the definition of the projections onto these sets. In this section, we provide a list of projections that can be used in various problems, with a special emphasis on phase retrieval applications. Some projections are important in that they are very commonly used, while others are worth noting to illustrate how the technique can be applied in less obvious cases.

In all cases below, the search space is a simple (real or complex) Euclidean space $E = \mathbb{R}^N$, or $E = \mathbb{C}^N$. Projections are thus defined in terms of minimizing the usual Euclidean distance. In most instances, an element of $E$ is a $N$-pixel image or a $N$-voxel three-dimensional density map. In all cases below, elements of $E$ will be noted by the symbol $\rho$ and indexed by a position vector $r$, to recall its interpretation as a density of some sort. It should also be understood that $r$ always represents discrete positions.
In many cases, the distance minimizing character of a projection is trivial. In other cases, however, some calculations are needed and the use of minimization with Lagrange multipliers can be a very convenient tool. In some cases (for instance when the constraint set has a discrete nature), the projections cannot be obtained through the simple minimization of a Lagrange function — but can still be exact and well-defined. However, some constraint sets also yield no analytic solution for the distance minimizing projection onto them. Numerical methods or approximations (“pseudo-projections”) have then to be used.

Because not all projections can be written as a single equation, we decided instead to use a procedural description. Although not as compact, it has the benefit of making transparent the implementation on a computer. The first set of projections presented are those related to diffraction measurements. We then present projections to various direct-space constraints, such as the support constraint. Additional examples of more complex projections will be defined in section 5.4.2.

3.4.1 Diffraction projections

In this section we present projections involving measured data from diffraction experiments.

Simple modulus projection

This projection is the simplest and most commonly used operation in phase retrieval. It is present in virtually all phase retrieval algorithms, from the “density modification” methods in crystallography, to the Gerchberg-Saxton and Hybrid-Input-Output algorithms. It was not originally called a “projection” since this
language is only of recent usage in the imaging community.

Let us assume that a set of diffraction pattern intensities has been measured. For now, we assume that this measurement is complete and exact. In crystallographic applications, these measurements are sampled on Bragg peaks, and commonly obtained by the integration of the signal over rocking curves, then assembled in a three-dimensional dataset. In diffraction microscopy, these are most commonly the intensity values gathered by a pixel-array detector. In both cases, these measurements represent the Fraunhofer diffraction patterns of a specimen and as such can be identified with the squared magnitude of the object’s Fourier transform — the “object” being either a 2D exit wave or a 3D refraction index map.

Thus, the constraint set for this problem is defined as:

$$\mathcal{C} = \{ \rho(r) \in \mathcal{E} \mid |\mathcal{F}\rho|^2 = I \}.$$  \hspace{1cm} (3.25)

The projection of a point \( \rho \in \mathcal{E} \) onto this set is given by the point \( \rho' = P(\rho) \) which minimizes \(|\rho' - \rho|^2\) and such that \( \rho' \in \mathcal{C} \). Formally, \( \rho' \) can be found as the minimization of a Lagrange function,

$$\mathcal{L} = \sum_r |\rho'(r) - \rho(r)|^2 + \sum_q \lambda_q \left( |\tilde{\rho}'(q)|^2 - I(q) \right),$$  \hspace{1cm} (3.26)

where the second term contains the constraints expressed in Fourier space, and \( \lambda_q \) are the corresponding (real-valued) Lagrange multipliers. Since the Fourier transform is a unitary transformation, the norm of any element of \( \mathcal{E} \) is conserved (“Parseval’s theorem”), so that \(|\rho' - \rho|^2 = |\tilde{\rho}' - \tilde{\rho}|^2\). Eq. (3.26) then becomes

$$\mathcal{L} = \sum_q |\tilde{\rho}'(q) - \tilde{\rho}(q)|^2 + \sum_q \lambda_q \left( |\tilde{\rho}'(q)|^2 - I(q) \right).$$  \hspace{1cm} (3.27)

This function is now easily minimized with respect to the Fourier space represen-
tation of $\rho$. One finally gets:

$$\tilde{\rho}'_q = \sqrt{I(q)} \frac{\tilde{\rho}_q}{|\tilde{\rho}_q|},$$

(3.28)

that is, the simple modulus projection simply consists in replacing the magnitudes (or moduli) of $\tilde{\rho}$ with the measured magnitudes, while conserving the phases. Geometrically, this is just projecting a point onto a circle in the complex plane. This constraint is clearly not convex (the interior of the circle is not included in the constraint set).

**Projection 1** (Simple modulus projection).

Let $I(q)$ be the measured intensities.

1. Compute the FFT of the current iterate $\rho$:

$$\tilde{\rho} = \mathcal{F}\rho$$

(3.29)

2. Replace the magnitudes of $\tilde{\rho}$ with the measured magnitudes:

$$\tilde{\rho}'(q) = \sqrt{I(q)} \frac{\tilde{\rho}(q)}{|\tilde{\rho}(q)|}$$

(3.30)

3. Inverse Fourier transform

$$\rho' = \mathcal{F}^{-1}\tilde{\rho}'$$

(3.31)

**Close variants**

The simple modulus projection has many variations. We list here a few which are commonly used.
It is very common that some of the intensities are not measured. In two-dimensional datasets, this may be caused by saturation of the detector or the use of a beam-stop. Because three-dimensional datasets are formed by the assembly of many 2D datasets, many points (voxels) are often missing, in the “gaps” between measured data. The effect of this missing data is discussed in chapter 4. The appropriate projection for this case is as simple as the previous one.

**Projection 2** (Modulus projection with missing data).

Let $M$ be the set of Fourier space coordinates for which the intensities are measured.

1. Compute the FFT of the current iterate, $\tilde{\rho} = \mathcal{F}\rho$
2. For all $\mathbf{q} \in M$, replace the magnitudes of $\tilde{\rho}$ with the measured magnitudes and leave unchanged the unmeasured ones.
3. Inverse Fourier transform: $\rho' = \mathcal{F}^{-1}\tilde{\rho}'$.

It is often possible to model the overall decay of the diffraction pattern intensity, so that an upper bound can be set on the missing intensities. This knowledge is also easily included in the modulus projection.

**Projection 3** (Modulus projection with bounds).

Let $M$ be the set of Fourier space coordinates for which the intensities $I(\mathbf{q})$ are measured. Let $B(\mathbf{q})$ be the upper bound on the unknown intensities.
1. Compute the FFT of the current iterate, \( \tilde{\rho} = \mathcal{F}\rho \)

2. For all \( q \in M \), rescale the magnitudes of \( \tilde{\rho}(q) \) to the measured magnitudes. For all \( q \not\in M \), replace the magnitudes of \( \tilde{\rho}(q) \) with \( \sqrt{B(q)} \) only if larger, otherwise leave unchanged.

3. Inverse Fourier transform: \( \rho' = \mathcal{F}^{-1}\tilde{\rho}' \).

**Modulus Projections with noisy data**

Because all physical measurements are noisy, the simple modulus projection formally never projects on a constraint set which contains the actual (physical) solution. As noise increases in a dataset, this projection becomes a worse approximation, eventually threatening convergence and reproducibility of reconstructions. Limited information content makes noisy measurements fundamentally harder to reconstruct. However, when the second constraint is strong enough, it is possible to relax the Fourier space constraint so that the constraint set lies closer to physical solution.

A straightforward approach is to “leave room” to the Fourier magnitudes according to the uncertainty of the corresponding measurement. The “annulus constraint” enforces the condition that every single Fourier magnitude in a reconstruction lies within a range determined by the uncertainty of the measurement: \( I_{\text{inf}}(q) < I(q) < I_{\text{sup}}(q) \). The inferior and superior bounds are set according to the uncertainty on the measurements. For low count numbers, a Poisson model could be preferred to a gaussian one.

**Projection 4** (Modulus annulus projection).
Let $I_{\text{inf}}(\mathbf{q})$ and $I_{\text{sup}}(\mathbf{q})$ be the inferior and superior bounds on the measured intensities.

1. Compute the FFT of the current iterate, $\tilde{\rho} = \mathcal{F}\rho$

2. Project onto the annulus:

   $$\tilde{\rho}'(\mathbf{q}) = \begin{cases} 
   \sqrt{I_{\text{sup}}(\mathbf{q})} \frac{\tilde{\rho}(\mathbf{q})}{|\tilde{\rho}(\mathbf{q})|} & \text{if } |\tilde{\rho}(\mathbf{q})|^2 > I_{\text{sup}}(\mathbf{q}), \\
   \sqrt{I_{\text{inf}}(\mathbf{q})} \frac{\tilde{\rho}(\mathbf{q})}{|\tilde{\rho}(\mathbf{q})|} & \text{if } |\tilde{\rho}(\mathbf{q})|^2 < I_{\text{inf}}(\mathbf{q}), \\
   \tilde{\rho}(\mathbf{q}) & \text{otherwise.}
   \end{cases}$$  \hfill (3.32)

3. Inverse Fourier transform: $\rho' = \mathcal{F}^{-1}\tilde{\rho}'$

The weakness of the annulus constraint comes from the fact that the realization of the constraint space will be mostly made of very improbable elements. For instance, it is very unlikely that the physical solution corresponds to a case where all intensities are larger than the actual measurement (meaning that all noise fluctuations would somehow be negative). Yet, this situation is included in the annulus constraint set.

At the cost of additional computation, it is possible to do better by taking into account the well characterized statistics of the noise. Although the photon counting process follows fundamentally the Poisson distribution, the error on the measurement can often be assumed to have a gaussian distribution because of additional sources of error and high counting rates. Given an element $\rho$ of $\mathcal{E}$, one can define unit-less variables

$$x(\mathbf{q}) = \frac{|\tilde{\rho}(\mathbf{q})|^2 - I(\mathbf{q})}{\sigma(\mathbf{q})}. \hfill (3.33)$$

If $\rho$ is the solution, following the gaussian noise assumption, the “law of large numbers” dictates that the set of $x(\mathbf{q})$ has a normal distribution with mean 0 and
standard deviation 1. This suggests the formulation of a stronger constraint set:

\[ C = \{ \rho \mid \text{the set } \{x(q)\} \text{ has a standard normal distribution.} \}. \quad (3.34) \]

How to project onto this type of constraint set is known as a “histogram projection”, discussed in more detail below (projection 13).

**Projection 5** (Modulus histogram projection).

Let \( I(q) \) and \( \sigma(q) \) be the measured intensities and their estimated error. Let also \( F^{-1} \) be the inverse of the integrated distribution function, that is \( x = F^{-1}(p) \) is such that

\[ p = F(x) = \int_{-\infty}^{x} f(x') \, dx', \quad (3.35) \]

where \( f \) is the distribution function of a normalized gaussian variable.

1. Compute the FFT of the current iterate, \( \tilde{\rho} = \mathcal{F}\rho \)
2. Calculate \( x(q) \), the normalized squared magnitudes of \( \tilde{\rho}(q) \) [Eq. (3.33)]
3. Apply the histogram projection on \( x(q) \); Find the ordered sequence of indices \( \{q_i\} \) such that \( x(q_i) \) is sorted in an increasing order; then, replace the \( x(q_i) \)'s with the histogram values:

\[ x(q_i) \rightarrow x'(q_i) = F^{-1}\left( \frac{i}{N+1} \right) \quad (3.36) \]

4. \[ \tilde{\rho}'(q) = \sqrt{I(q) + \sigma(q)x'(q)}\frac{\tilde{\rho}(q)}{|\tilde{\rho}(q)|} \quad (3.37) \]
5. Inverse Fourier transform: \( \rho' = \mathcal{F}^{-1}\tilde{\rho}' \).
Near-field and Far-field modulus projections

For two-dimensional data, the main idea of the modulus projection for Fraunhofer diffraction patterns can be easily generalized to other type of diffraction data, at large or small Fresnel numbers (near-field or far-field diffraction). For this case, simple analytic expressions allow to “back-propagate” numerically the wavefield to the specimen plane.

In the near field case, the measurements are 2D real-space intensities $I(r)$, and the constraint set is made of all the $\rho(r)$ such that $I(r) = |\psi(r; z)|^2$, where $\psi(r; z)$ is given by equation (2.31):

$$\psi(r; z) = \mathcal{F}^{-1} \left[ e^{i(k-\kappa)z} \tilde{\rho}(q) \right].$$

(3.38)

Using again the unitarity of the Fourier transformation, the distance to be minimized between a current element of $\mathcal{E}$ and an element of $\mathcal{C}$ is

$$\sum_r |\rho'(r) - \rho(r)|^2 = \sum_q |\tilde{\rho}'(q) - \tilde{\rho}(q)|^2$$

$$= \sum_q |e^{i(k-\kappa)z} (\tilde{\rho}'(q) - \tilde{\rho}(q))|^2$$

$$= \sum_r |\psi'(r; z) - \psi(r; z)|^2.$$

(3.39)

The distance-minimizing projection is therefore analogous to the simple modulus projection. This projection has been used in many applications already – although it is seldom called “projection” – (see for instance Allen et al., 2001).

Projection 6 (Near-field modulus projection).

Let $I(r)$ be the intensities of the near-field diffraction image measured a distance $z$ downstream from a specimen.
1. Propagate $\rho$ forward a distance $z$: $\psi = \mathcal{F}^{-1} \left[ e^{i(k-\kappa)z} \mathcal{F} \rho \right]$

2. Replace the magnitude of $\psi$ with $\sqrt{I}$:

$$\psi'(r) = \sqrt{I(r)} \frac{\psi'(r)}{|\psi(r)|} \quad (3.40)$$

3. Propagate $\psi'$ back a distance $z$: $\rho' = \mathcal{F}^{-1} \left[ e^{-i(k-\kappa)z} \mathcal{F} \psi' \right]$.

In the special case where $z = 0$, the first and last steps are of course not needed. This corresponds to the direct-space projection originally used in the Grechberg-Saxton algorithm (see section 3.3.2).

The far-field projection, based on eq. (2.40), is similarly defined. Its use in phase retrieval was suggested by Xiao and Shen (2005).

**Projection 7** (Far-field modulus projection).

Let $I(q)$ be the intensities of the far-field diffraction image measured a distance $z$ downstream from a specimen. Define

$$G(r) = \exp \left( \frac{ikr^2}{2z} \right) \quad (3.41)$$

1. Propagate $\rho$ into the far-field: $\tilde{\psi}(q) = \mathcal{F} [\rho(r)G(r)]$.

2. Replace the magnitude of $\tilde{\psi}$ with $\sqrt{I}$:

$$\tilde{\psi}'(q) = \sqrt{I(q)} \frac{\tilde{\psi}'(q)}{|\tilde{\psi}(q)|} \quad (3.42)$$

3. Propagate $\tilde{\psi}'$ back a distance $z$: $\rho' = G^*(r) \mathcal{F}^{-1} \tilde{\psi}'$.

Similar to the near-field case, the limit $z \to \infty$ corresponds to the simple modulus projection (projection 1). Although this projection is called “far-field”, it
can be used even if the Fresnel number is much larger than unity (see, for instance, the simulation in Section 5.4.2).

Both near-field and far-field projections can be adapted to accommodate missing data and noise in the same way this is done for the simple modulus projection.

3.4.2 Direct space projections

Phase retrieval require the use of additional information in direct space. We present here a list of the most commonly used projections. When possible, we give the original reference of their introduction.

Support projections

The notion of support in phase retrieval goes back to the very idea of oversampling a diffraction pattern. Its use as a constraint in phase retrieval seems to be due to Fienup (1978).

The constraint set is simply stated: a density (image) satisfies the support constraint if it is 0 everywhere outside the support. It is equally trivial to define the distance-minimizing projection onto this constraint set: set to 0 all values outside the support.

**Projection 8** (Support projection).

Let $S$ be the support. Then, execute the following operation:

1. 

$$
\rho'(r) = \begin{cases} 
\rho(r) & \text{if } r \in S, \\
0 & \text{otherwise},
\end{cases}
$$

(3.43)
It should be noted that, in some instances (see for instance the reconstructions in sections 5.1 and 5.2) the assumption that the non-support values are 0 is formally wrong. The exit wave formed by an isolated specimen bathed in an incoming plane wave is certainly not zero outside the support: it is simply equal to the amplitude of the incoming wave. Setting the non-support to 0 in a reconstruction therefore implies that a constant value has been subtracted of the complete density. Fortunately, this leaves a measured diffraction pattern perfectly unchanged except for its exact center (sometimes called the “DC mode”). This statement is known as the “Babinet principle” in optics. Since the center of a diffraction pattern is almost never measured,\(^7\), the above definition of support projection is general enough for almost all purposes.

The support constraint requires that the support be known in the first place. In two dimensions, this is not a big problem, and with enough patience, it is generally possible to obtain a well-adjusted support simply with the use of a drawing program. It is, however, much more convenient to have an automated support refinement method.

Dynamic support constraint sets are looser constraints which allow the shape of the support to vary from iteration to iteration. Dynamic support constraint sets can be defined in terms of a threshold on the density values, or in terms of a total allowed area. The threshold-based constraint set is given by the set of all \(\rho(\mathbf{r})\) which have no value such that \(0 < |\rho(\mathbf{r})| \leq T\), that is, the amplitude of \(\rho(\mathbf{r})\) is either 0 or is above a given threshold \(T\). The area-based constraint set is the set of all \(\rho(\mathbf{r})\) which have exactly \(N_S\) non-zero pixels. Projections onto these two constraint sets follow.

\(^7\) Instances where the center of a diffraction pattern is measured is the case where the surrounding 0 values are valid, as when a mask is used.
**Projection 9** (Dynamic support projection – threshold).

Let $T$ be the threshold on the amplitude of $\rho(r)$. Execute the following step:

1. 

\[
\rho'(r) = \begin{cases} 
\rho(r) & \text{if } |\rho(r)| \geq T, \\
T \rho(r)/|\rho(r)| & \text{if } T/2 < |\rho(r)| < T, \\
0 & \text{if } |\rho(r)| \leq T/2
\end{cases}
\]  

(3.44)

**Projection 10** (Dynamic support projection – area).

Let $N_S$ be the total number of pixels in the support.

1. Sort the values of $|\rho(r)|$ and set to 0 all but the $N_S$ largest.

The area-based method was used in crystallography, in a density modification technique called “solvent flattening” (Wang, 1985), although in a slightly modified form (a low-pass filter is applied on $\rho$ before applying the projection). We could not find instances of this method in the diffraction microscopy or image processing literature.

Thresholding might be a more popular approach in part because it is faster and also because a threshold may be easier to determine and adjust than an area.\textsuperscript{8} It is, for instance, a central ingredient to the charge-flipping (Oszlányi and Sütő, 2004) and low-density elimination (Shiono and Woolfson, 1992) algorithms. Interestingly, in all known applications, the thresholding operation is not the same as

\textsuperscript{8}Although, in crystallography the *solvent content* is information often readily available for the calculation of the support area.
Figure 3.2: Illustration of the threshold projection (a) for a real variable and (b) for a complex variable.

the projection shown above. It always consists in simply setting to 0 all values of $|\rho(r)|$ lying below $T$. This operation is not a distance-minimizing projection onto the constraint set described above.

Popular variations on this approach are the use of a low-pass filter on $\rho$ prior to the application of the threshold (Wang, 1985; Marchesini et al., 2003). Again, this operation is not distance-minimizing, and is most of the time not even a projection. The fact that good results can still be obtained with it suggests that it is still well behaved. One can also see the use of dynamic support as a procedure external to
the reconstruction process. The actual projection used is then the simple support projection 8, with an additional intervention that redefines the support as the iteration evolves. This is especially consistent with what Marchesini et al. (2003) suggest in their “Shrink-wrap” algorithm, as the definition of the support is not updated at each projection application, but rather is changed at longer intervals.

We conclude our tour of dynamic support methods with the atomicity projection (Elser, 2003b), very similar to the area-based support, except for the fact that a small region around the $N_{\text{atom}}$ highest pixel values is also left untouched, where $N_{\text{atom}}$ is the expected number of atoms to be reconstructed.

**Projection 11** (Atomicity projection).

Let $N_{\text{atom}}$ be the number of atoms in $\rho$ and let $a$ be the radius of one atom.

1. Find the list of the $N_{\text{atom}}$ highest pixel values, excluding nearest-neighbor pixels in the list.$^9$

2. Set to 0 all pixel values that are not within a distance $a$ of the $N_{\text{atom}}$ points.

**Value projections**

Value projections form an important second category of direct space constraints. They consist of all constraints applied directly on the pixel values. Most often they

---

$^9$A possible refinement is the definition of a fractional pixel position of the maximum based on a fit to neighboring pixels.
are used in conjunction with fixed support constraints, since formally support and value projections are commuting operations.

The discussion of section 4.2.3 explores the question of the possible values taken by pixels in a reconstructed exit wave. In some special conditions, these values can be assumed, to a very good approximation, to be real and positive. When it can be used, positivity is known to be a very powerful constraint. For completeness, we state here the formulation of the positivity projection even though it is straightforward.

**Projection 12** (Positivity).

Let \( \rho(\mathbf{r}) \) be real.

1. \[
\rho'(\mathbf{r}) = \begin{cases} 
\rho(\mathbf{r}) & \text{if } \rho(\mathbf{r}) \geq 0, \\
0 & \text{otherwise},
\end{cases}
\] (3.45)

Many other value constraints have a similar formulation. For instance, one can enforce that

1. \( \rho(\mathbf{r}) \) is real: set the imaginary part to 0.

2. The imaginary part of \( \rho(\mathbf{r}) \) is positive: apply positivity to the imaginary part.

3. The amplitude of \( \rho(\mathbf{r}) \) is bounded: leaving the phase unchanged, project all pixel values outside a circle onto it.

---

\(^{10}\)This constraint can be useful as a result of the fact that physical materials cannot have a negative absorption (Miao and Sayre, 2000)

\(^{11}\)This constraint also corresponds to the impossibility of negative absorption, but in the case of an optically thick specimen (see for instance, Figure 5.17).
4. The amplitude of $\rho(\mathbf{r})$ is fixed: project all pixels values onto a fixed radius circle, leaving the phase unchanged.

Another type of important value constraint is the histogram constraint, (Zhang and Main, 1990; Elser, 2003a) already mentioned above for the modulus histogram projection. This projection enforces that the solution have a well-defined histogram of values. While the histogram is not a known quantity in many cases, some types of specimen are well characterized, such as the electron density in a protein crystal, or the magnetic scattering by a thin ferromagnetic specimen (Loh et al., 2007).

While the histogram constraint set formulation is simple ($\rho(\mathbf{r})$ has a given distribution), projection onto this set does not seem simple. Zhang and Main (1990) suggest a method requiring the pixel values to be binned and rescaled. It is not clear that this is a projection (in the description of the method by the author, this question does not seem to be a concern). Elser (2003a) suggested instead a method based on the sampling of the values according to the distribution, and assignment of the optimally permuted elements of this list to the density $\rho$. This simple assignment method involves a sorting of the values of $\rho$ and replacing them with the sorted samples from the known histogram.

**Projection 13** (Histogram projection).

Let $f(x)$ be a known probability distribution ("histogram"). Let $F^{-1}(p)$ be the inverse of the integrated distribution function, that is $y = F^{-1}(p)$ is such that

$$p = F(y) = \int_{-\infty}^{y} f(x)dx. \quad (3.46)$$

---

$^{12}$This occurs for an optically thick pure phase specimen.
Define the sorted samples as

\[ a_i = F^{-1}\left(\frac{i}{N + 1}\right), \quad (3.47) \]

where \( N \) is the total number of pixels in \( \rho \).

1. Sort the values of \( \rho \) in ascending order, that is find the sequence of indices \( \{r_i\} \) such that \( \rho(r_i) \leq \rho(r_j) \) if and only if \( i < j \) for all \( i, j \).

2. Replace the values of \( \rho \) with the samples \( \rho(r_i) = a_i \).

The histogram projection is a distance-minimizing projection onto the set of all \( \rho \) with list of values given by a permutation of the list \( \{a_i\} \). That this constraint set is almost the same as the histogram requirement is a natural consequence of the “law of large numbers”.
Chapter 4

Working with experimental data

One of the ultimate goals of the reconstruction techniques described in the previous chapters is to produce successfully the image or the density of the scattering object. In this context, “successful” means that the image is unique and accurate.\(^1\)

In this chapter, we will explore how diffraction microscopy is affected by experimental realities. The impact of “real life” is twofold. First, some experimental limitations can compromise the quality of the data itself. Coherence, noise, limited dynamic range are examples of this. Second, even if the data is good in principle, some physical situations can still make reconstructions harder. We will discuss, for instance, how complex-valued reconstructions are in general more challenging. This second category will be loosely defined as “algorithm limitations”

4.1 Experiment limitations

The picture presented up to now is that of ideal experiments: in the previous chapter, we have assumed at times perfect knowledge of the diffraction pattern, at times perfect isolation of the specimen. It is now time to examine how experimental conditions can limit the quality of the data, and ultimately its reconstructibility.

In what follows, we will describe the effect on diffraction data and reconstructions of partial coherence, noise, and missing data (caused by saturation or obstruction with a beam stop).

\(^1\)One may also add that a condition for success is that, over all, the resolution and quality are high enough to make the technique competitive.
Preliminaries

Most of the time, a far-field diffraction pattern is measured on a square or rectangular detector, whose size determines the resolution of the reconstruction. If $L$ is the distance between the specimen and the detector and $w$ is the width of the detector, only light scattered at an angle up to $\theta_{\text{max}} \approx \tan^{-1}(w/L)$ is recorded. In Fourier space, this corresponds to an upper bound on the spatial frequencies in the perpendicular plane (or momentum transfer): $q_{\text{max}} = k \sin(\theta_{\text{max}}) = w/\sqrt{L^2 + w^2}$. The limit on the recorded spatial frequency sets the achievable resolution of the reconstruction. The resolution is defined as the shortest spatial period in the reconstruction: $d_{\text{min}} = 2\pi/q_{\text{max}}$.

4.1.1 Coherence

Coherence is a measure of the degree of correlation between two points in space and time of a field. Essentially of a statistical nature, coherence originates from the fact that individual sources are, in general, uncorrelated. A wave is said to be incoherent if, on average, all interference effects are suppressed. We have already stated that diffraction microscopy needs coherence to yield valuable data. Since no source is perfectly coherent, it is important to know how partial coherence is defined and how coherent a wave needs to be to give useful data.

Theory

In the following, we will restrict ourselves to a few important concepts defined with classical fields (For a quantum mechanical treatment of coherence, see for

2When the diffraction pattern is formed in the back focal plane of a lens, the same relation holds if $L$ is the lens’s focal length.
instance Glauber, 1963a,b). Let $\Psi(r, t)$ be a scalar field having fluctuations. For a thermal source, these fluctuations result from the superposition of modes with different frequencies, uncorrelated between different point sources. For sources such as lasers and synchrotron radiation, it is wrong to assume that sources (atoms) are uncorrelated, but some fluctuations occur nevertheless. In any case, we now consider the wavefield as a stochastic quantity for which ensemble averages give the measured values.

Let $\langle I(q, t) \rangle$ be the averaged intensity of a far-field diffraction pattern. Its Fourier transform is an autocorrelation $\langle A(r, t) \rangle$,

$$
\langle A(r, t) \rangle = \int d^2r' \langle \Psi(r', t) \Psi^*(r' - r, t) \rangle.
$$

(4.1)

The time dependence of the above quantities expresses the fact that we have departed from the assumption of chapter 2 that the wavefield is monochromatic. All wavefield expressions considered in chapter 2 should now be labeled with a $\omega$ (of course, $\omega = kc$), and the full time-dependent wavefield is given by

$$
\Psi(r, t) = \int s(\omega) \Psi_\omega(r) e^{i\omega t} d\omega,
$$

(4.2)

where $s(\omega)$ is the complex amplitude of the frequency $\omega$, and $\Psi_\omega(r)$ are normalized (or, in case they are not integrable, scaled according to a well-defined rule). Typically, $s(\omega)$ is peaked around a central frequency $\omega_0$ and has a width $\Delta \omega$. The ratio $\Delta \omega/\omega_0$, sometimes called “spectral density”, is a commonly used quantity to characterize the monochromaticity of a source.

We now proceed to integrate over time, in accordance with the fact that measurements have a span in time. Formally, we compute

$$
\langle A(r) \rangle = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \langle A(r, t) \rangle \, dt,
$$

(4.3)
but we will assume for simplicity that \( \langle A(r,t) \rangle \) is square-integrable, so that
\[
\langle A(r) \rangle \propto \int dt \langle A(r,t) \rangle.
\] (4.4)

Prefactors are unimportant in the present analysis, so proportionality symbols will be replaced with equalities. Substituting (4.2), we find that
\[
\langle A(r) \rangle = \int d\omega \int d\omega' \langle |s(\omega)|^2 \Psi_\omega(r') \Psi_\omega^*(r' - r) \rangle
\] (4.5)

To progress further, we assume that \( \Psi(r) \) can be factored as \( \Psi_0(r)O(r) \), where \( O(r) \) is the modulation of the incident wave \( \Psi_0 \) produced by the object (the specimen). This form is valid when \( \Psi_0 \) is approximately a plane wave and the specimen is optically thin [Eq. (2.65)], or in the small-wavelength limit [Eq. (2.46)]. Equation (4.5) becomes
\[
\langle A(r) \rangle = \int d\omega \int d\omega' \langle O_\omega(r') \Psi_\omega^*(r' - r) \rangle \Psi_0(r') \Psi_0^*(r' - r).
\] (4.6)

**Spatial and temporal coherence**

Two distinct paths from equation (4.6) are now considered. Most of the following results can be found in a recent paper from Vartanyants and Robinson (2001).

First, we assume that the wave is quasi-monochromatic; this corresponds to assuming that \( |s(\omega)| \) is narrowly peaked around a central frequency \( \omega_0 \). Furthermore, the integrand is assumed to vary slowly enough with \( \omega \), so that equation (4.6) becomes
\[
\langle A(r) \rangle = \int d\omega d\omega' \langle O_{\omega_0}(r') \Psi_{\omega_0}^*(r' - r) \rangle \langle \Psi_{\omega_0}(r') \Psi_{\omega_0}(r' - r) \rangle.
\] (4.7)

where \( W(r_1, r_2, \omega) \), sometimes called the “cross-spectral density”, or “mutual spectral density”, is the Fourier transform of the *mutual intensity* (Zernike, 1938), or
mutual coherence function (Wolf, 1955):

\[
\Gamma(r_1, r_2, t) = \mathcal{F}_\omega^{-1} W(r_1, r_2, \omega) = \langle \Psi_0(r_1, t) \Psi_0(r_2, 0) \rangle.
\] (4.8)

In terms of \( \Gamma \), the average intensity at a point \( \mathbf{r} \) can be written

\[
\langle I(\mathbf{r}) \rangle = \Gamma(\mathbf{r}, \mathbf{r}, 0).
\] (4.9)

It is sometimes convenient to normalize \( \Gamma \) to obtain the “complex coherence factor”, or “complex degree of coherence”,

\[
\gamma(r_1, r_2, t) = \frac{\Gamma(r_1, r_2, t)}{\sqrt{\langle I(r_1) \rangle \langle I(r_2) \rangle}},
\] (4.10)

so that \(|\gamma|\) is always between 0 and 1. If \(|\gamma| = 0\), the field is incoherent at these points, whereas it is completely coherent if \(|\gamma| = 1\). In between, the wave is said to be partially coherent. The normalized cross-spectral density is the Fourier transform (in time) of the complex degree of coherence:

\[
w(r_1, r_2, \omega) = \mathcal{F}_t \gamma(r_1, r_2, t)
\] (4.11)

A good approximation to \( W \) (or \( \Gamma \)) is that of a quasi-monochromatic source far from the measurement plane. Let \( S(R_\perp) \) be the average intensity of the source as a function of the transverse position \( R_\perp \), \( \tilde{S}(Q) \) its Fourier transform, and \( L \) the distance between the measurement plane and the source. The result is (Mandel and Wolf, 1965):

\[
W(r_1, r_2, \omega) \propto \left( \frac{k}{L} \right)^2 \exp \left[ \frac{ik(r_1^2 - r_2^2)}{2L} \right] \tilde{S} \left( Q = k\frac{r_2 - r_1}{L} \right).
\] (4.12)

If \( a \) is the spatial extent of the object, the exponential factor is negligible when \( ka^2/L \ll 1 \), and \( W \) becomes a function of \( r_2 - r_1 \) only.

\[
W(r_1, r_2, \omega) \approx W(r_2 - r_1) \propto \tilde{S} \left( k\frac{r_2 - r_1}{L} \right).
\] (4.13)
The averaged autocorrelation then takes the very simple form

\[ \langle A(r) \rangle = S(kr/L) A_O(r) \]
\[ \propto \tilde{w}(kr/L) A_O(r), \]  

(4.14)

where \( A_O(r) \) is the autocorrelation of the object modulation \( O(r) \), and \( \tilde{w} \) is the Fourier transform of \( w(r_2 - r_1) \). The resulting intensity is

\[ \langle I(q) \rangle = I_O(q) * S(qL/k), \]

(4.15)

where \( I_O \) is the intensity that would be obtained if the specimen had been illuminated by a perfect plane wave. The final result is easy to interpret: the intensity \( I_O \) is convolved with the intensity profile of the source. The function \( S \) thus acts as a point spread function, whose effect is to blur the measured diffraction pattern.

The normalized form of eq. (4.13) \[ w, \text{ from eq. (4.11)] \] is close to 1 in a region around the origin. The area of coherence, or coherence patch, is an evaluation of the extent of this region. A circular incoherent source of radius \( R \) has an area of coherence \( \Delta A \) roughly equal to

\[ \Delta A = \alpha \frac{\lambda^2 L^2}{\pi R^2}, \]

(4.16)

where \( \alpha \) is a number of order 1 depending on the specific criterion used to define “coherence”. For \( |w| > .88 \), \( \alpha \) is \( 1/4 \) (Mandel and Wolf, 1965). Hence, with a given geometry, the area of the coherence patch scales like the square of the wavelength.

Figure 4.1 shows the modification of a diffraction pattern caused by a mild degree of incoherence. In this specific example, the Fourier transform of the source was approximated by a gaussian function\(^3\).

\(^3\)This approximation is reasonable when modeling synchrotron x-rays from an electron beam.
Figure 4.1: Illustration of the effect of partial transverse coherence. (a) A diffraction pattern obtained with a fully coherent incident wave. (b) The same diffraction pattern with a partially coherent incident wave. In this example, the radius of the coherence patch [eq. (4.16), with $\alpha = 1/4$] is equal to the radius of the specimen (half the radius of the autocorrelation).

The second situation we consider is that of limited temporal coherence. Assuming that the incident wave is nearly a plane wave for all frequencies and that the transverse coherence is high, the averaging brackets in the right-hand side of eq. (4.6) can be removed, leading to

$$\langle A(r) \rangle = \int d\omega |s(\omega)|^2 A_\omega(r),$$

(4.17)

or:

$$\langle I(q) \rangle = \int d\omega |s(\omega)|^2 I_\omega(q).$$

(4.18)

In this last equation, the intensity on the right is that from a monochromatic wave with frequency $\omega$. The averaged intensity is simply an average over all frequency terms. Usually, the main dependence on $\omega$ is through the scaling of the argument $q$. Using again $u$ as the direction of observation $u = q/k$, the intensity can be
written as

$$\langle I(u) \rangle = \int dk |s(k/c)^2 I(ku)\rangle.$$ \hspace{1cm} (4.19)

This case is well known; it illustrates for instance the dispersion of light by a grating, as is used in spectroscopy. The effect of the non-zero bandwidth of $s$ is stronger and stronger as $|u|$ increases (radially). Temporal coherence therefore has the effect of limiting the resolution of the diffraction pattern. If the scattering object has a size $a$, the speckles’ size in Fourier space is roughly $\pi/a$. The width of the spread caused by the spectrum bandwidth is of this order when

$$|q|_{\text{max}} \sim \frac{\pi}{a} \left( \frac{\Delta \omega}{\omega_0} \right)^{-1}. \hspace{1cm} (4.20)$$

The maximum resolution of a diffraction pattern from a non-monochromatic source is inversely proportional to the spectral purity. Equation (4.20) provides a good rule of thumb: the extent of the usable part of the diffraction pattern is equal to the inverse of the spectral purity, in units of speckle size. In terms of the *longitudinal coherence length*,

$$\Delta \ell = c \Delta t = \frac{2\pi c}{\Delta \omega}, \hspace{1cm} (4.21)$$

equation (4.20) also reads

$$|q|_{\text{max}} \sim \frac{\pi}{a} \left( \frac{\Delta \ell}{\lambda} \right). \hspace{1cm} (4.22)$$

The effect of partial longitudinal coherence is illustrated on Fig. 4.2. The effect is mostly visible at the edge of the diffraction pattern, where radial blurring is more significant.

**Discussion**

From the description above, it is clear that partial coherence can affect the quality of the data to the point of compromising the success of reconstructions. While
Figure 4.2: Illustration of the effect of partial longitudinal (temporal) coherence. (a) A diffraction pattern obtained with a fully coherent incident wave. (b) The same diffraction pattern with a partially coherent incident wave. This example corresponds to a spectral purity of \( \Delta \omega / \omega \approx 0.6\% \) or, equivalently, to a longitudinal coherence length of about \( 160\lambda \).

Spectral purity (temporal coherence) limits the usable extent of a diffraction pattern,\(^4\) an insufficient spatial coherence may render the entire diffraction pattern unusable. In practice, with x-rays spectral purity seems easier to achieve: x-ray monochromators use diffraction to disperse waves having different wavelength in different directions [by eq. (4.19)], after which it is possible to select a narrow band of wavelengths.

Transverse coherence is generally achieved with effectively small source sizes, with the important exception of lasers. In synchrotrons, the source is the electron beam, generally too large for diffraction microscopy, so that a pinhole needs to be used in addition. Electron microscopes can have high transverse coherence with field-emission guns, which are effectively very small sources. According to (4.14),

\(^4\)The use of an **autocorrelation spectrogram** (Chapman et al., 2006b, figure 3) can show the loss of signal caused by a limited temporal coherence.
information recorded in the autocorrelation is available only if the coherence patch is larger than the autocorrelation support (that is, roughly, if the coherent length satisfies $\Delta l > 2a$, where $a$ is the diameter of the object). In well controlled situations, it is possible to measure the complex degree of coherence (Lin et al., 2003; Pfeiffer et al., 2005) as a function of the position in the transverse plane of the beam. If $|\gamma|$ departs only slightly from 1 out to the edge of the autocorrelation support, deconvolution of equation (4.15) could be attempted prior to a reconstruction.

A small transverse coherence length has the effect of making the Fourier transform of the measured intensity compact, as if supported. This can lead to the wrong conclusion that the specimen is smaller than it actually is, and that the oversampling condition is satisfied. The incorrect view that a small coherence length is equivalent to selecting a small field of view has been brought forward by Zuo et al. (2004) when attempting to reconstruct the image of a carbon nanotube from its diffraction pattern. The shortcomings of electron diffraction microscopy will be discussed in more detail in section 4.2.4.

4.1.2 Noise

Diffraction-based imaging techniques always involve the measurement of a particle flux during a finite amount of time. These measurements are fundamentally noisy because of the discrete character of the detection process. Most often, measurements are made with very efficient charge-coupled device (CCD) detectors. The counting statistics determine an upper bound on a measurement’s signal to noise ratio, but noise may originate from many other sources, including the background and CCD read-out.
We consider in this section any source of error in the data (even those which are not strictly due to noise). Like in any type of measurement, sources of error can be placed in two categories: random and systematic. After having described common error sources in both categories, we will discuss how the reproducibility and the resolution of the reconstruction is affected by the resulting incompatible constraint sets.

**Sources of random errors**

In all cases of interest, intensity measurements originate from particle counting, which has a Poisson distribution,

\[ f(n) = \frac{e^{-I/I_0}(I/I_0)^n}{n!}, \]

where \( n \) is the particle count, \( I \) is the time-integrated intensity and \( I_0 \) is the energy of one particle (of course, this simple model only applies to monochromatic waves).

Given a particle count \( n_m \) (for “measurement”), the most probable integrated intensity is \( n_m I_0 \), with measurement error \( \Delta n = n_m^{1/2} I_0 \). Increasing the measurement time increases the particle count and thus reduces the relative error, \( n_m^{-1/2} \).

Uncontrolled sources of radiation and heat increase the particle count in a reproducible way. To eliminate this bias (systematic error), it suffices to subtract from a measurement the particle count obtained in the same condition but with the source removed. If \( n_b \) particles are counted from the background, the error on \( n_m - n_b \) is increased to \( \Delta n = (n_m + n_b)^{1/2} \).

Most CCDs introduce an additional source of noise when reading out the data (caused by the signal amplification). This “false count” \( n_r \) also contributes to the error - although, not being related to a flux, it does not scale with time. The signal-to-noise ratio (the inverse of the relative error, noted \( S/N \)) is therefore, in
an ideal situation,

\[ S/N = \sqrt{t} \frac{\sqrt{F_m}}{\sqrt{1 + F_b/F_m + n_r/(F_m t)}}, \]  

(4.24)

where \( t \) is the measurement time and the \( F \)'s are the particle detection rates corresponding to each of the sources mentioned above. Obviously, \( S/N \) is higher for longer exposure times and fewer read-outs.

In practice, many factors prevent arbitrarily good data from being taken (apart from the obvious limited time available to the experimenters). The finite dynamic range of detectors, combined with the typically large range of intensities in a diffraction pattern (up to 6 orders of magnitude) limits the time of single exposures, thus necessitating higher numbers of CCD read-outs.

**Radiation damage**

High-energy radiation causes damage to specimens, with important consequences especially on organic materials. In many situations, the limited dose a specimen can sustain is the most important limitation to the quality of datasets.

The importance of radiation damage in x-ray microscopy has been known for a long time (see for instance Sayre and Chapman, 1995) and has been studied in great detail lately (see Howells et al., 2005; Shen et al., 2004), especially as the resolution of images approaches the limit it imposes.

The dose is defined as the absorbed energy per unit mass. It is given by:

\[ D = \frac{\mu E_0}{\rho_m}, \]

(4.25)

where \( \mu = -k \text{Im}\{\delta n\} \) is the absorption coefficient, \( E_0 \) is the total energy per unit of area, and \( \rho_m \) is the mass density.

The maximum dose a specimen can sustain is hard to determine as it depends on many factors. Important ones are the nature of the specimen – and the way it is
prepared – and image (or reconstruction) resolution. Owing to their lower density and to the weaker chemical bonds, biological tissues are much more fragile than metallic structures. Normally, higher doses can be tolerated for lower resolution reconstructions since the radiation damage effects appear first at the atomic level and then propagate as more and more bonds are broken and more and more ions are free to diffuse. Structural changes of dehydrated biological specimens appear as shrinkage and mass loss, effects that have been observed to be much less prone to occur on frozen hydrated specimens (see for instance Beetz and Jacobsen, 2003).

Lately, Howells et al. (2005) have gathered data from many studies using x-rays and electron microscopy, giving an idea of the maximum dose as a function of resolution. This data and the evaluation of Shen et al. (2004) give a maximum dose growing roughly as the resolution to the power 3/2. Regardless of what explains this behavior, the maximum dose a specimen can sustain to be imaged at a given resolution needs to be compared to the minimum dose needed to get enough signal to image at this resolution. What constitutes “enough signal” can be different in different situations. Howells et al. (2005) define “enough” in terms of the Rose criterion (Rose, 1948): each voxel (volume element) is required to have a signal-to-noise ratio of at least 5, which corresponds to the measurement of at least 25 photons coming out of this voxel. This criterion is based on the interpretability of the reconstruction – the $S/N = 5$ threshold sets the limit between details that are considered discernable and those that are lost to noise. Shen et al. (2004) addressed the more specific problem of the dose needed to allow reconstruction. This noise criterion is more system specific, as it requires a model of the decay of the diffraction pattern envelope as a function of spatial frequency, and needs to take into account oversampling requirements. In both cases, the minimum dose
is found to scale like the resolution to the power $-4$, and the highest achievable resolution is thought to lie around 5 nm. Close to this limit, it is to be expected that the signal-to-noise ratio in the outer part of a diffraction pattern could fall close to, or even below 1. How reconstruction algorithms can cope with this type of data has yet to be investigated.

**Sources of systematic errors**

While random errors come from unavoidable physical limitations, systematic errors are biases that could in principle be eliminated. The systematic errors most commonly encountered in diffraction microscopy result from the presence of additional ("alien") sources of scattering. These sources can be, for instance, irregularities or dust particles on the substrate holding the specimen (see 5.1 and 5.2 for more discussion and examples of such sources). In electron diffraction microscopy, an amorphous substrate generates a "background" signal which adds coherently with the specimen, and hence cannot be subtracted the same way other (incoherent) background sources can. This problem poses an important challenge to electron diffraction microscopy. When they are not identified, alien scatterers contribute to make the constraint spaces (used in the reconstruction algorithm) farther apart; this in turn results in lower quality reconstructions. When possible, it is preferable to identify and incorporate these sources into the constraint.

**Averaging**

It was mentioned in the general discussion on the difference map that averaging over many estimates is a route for getting reproducible results even when the constraint sets are incompatible. In fact, averaging was especially devised for
reconstructions with noisy data. More can be said when the constraint sets are well defined, as is the case for the difference map applied to diffraction microscopy.

Averaging was adopted for all applications presented in the next chapter. Since the physically significant constraint comes from the experimental measurement, the solution is always defined as the average of the Fourier estimate. The solution is therefore a sum of estimates having all the same magnitudes (defined by experiment) and differing only by their phases. Assuming that the estimate fluctuations follow a gaussian distribution, the squared magnitude of the average is

$$\left| \tilde{\Psi}_{av}(q) \right|^2 = I(q)|\langle \exp(i\phi_q) \rangle|^2 \approx I(q) \exp \left( -\frac{1}{2} \Delta \phi_{q}^2 \right), \quad (4.26)$$

where $I(q)$ are the measured intensities, and $\Delta \phi_q$ is the width of the phase fluctuations upon averaging. The averaged magnitudes are thus always smaller than the measured diffraction data, that is, the ratio

$$r(q) = \frac{\left| \tilde{\Psi}_{av}(q) \right|^2}{I(q)} \quad (4.27)$$

is bounded above by 1. This ratio is a measurement of the fidelity with which the algorithm transfers the measured intensity to the reconstruction. A source of phase fluctuations, always present in practice, is the noise in the intensity measurement. A high relative error on $I(q)$ generally means a greater incompatibility between the direct space constraint and this measurement, which in turn means larger phase fluctuations. The result of averaging is to reduce the intensity of all incompatible regions. Since diffraction patterns exhibit a strong radial decay, this generally means that averaging is equivalent to a low-pass filtering operation, analogous to the modulation transfer function (MTF) in incoherent optical systems. In fact, one can define the algorithmic analogue of the MTF as the radial average of the
fidelity function $r(q)$. This quantity has been called PRTF (Chapman et al., 2006a), standing for phase retrieval transfer function. This function will be used in the next chapter to give an estimate of the resolution of a reconstruction.

In addition to the information in its radial average, $r(q)$ can provide a “map” in Fourier space of the incompatibility of constraints. This map can help identify, in some instances, the systematic sources of incompatibility. An example of this situation is the effect of the spider silk in the insect wing reconstruction presented in section 5.2.

We conclude with a word of caution: averaging may guarantee the reproducibility of a reconstruction, but it does not ensure that the reconstruction is correct. Poor data quality, or wrong constraints can give a reproducible, yet incorrect reconstruction.

### 4.1.3 Missing data

Measured diffraction patterns are sometimes incomplete, most often because of detector limitations. CCDs have a finite bin depth, causing saturation in regions where intensity is too high. Also, very high photon flux can damage certain detectors, making necessary the use of a beam-stop. Three-dimensional datasets often have regions of unmeasured data because of geometry limitations: the assembly of two-dimensional diffraction patterns taken around a single rotation axis generally leads to unmeasured cones because of the Ewald sphere curvature.

Missing data does not always constitute a problem. The corresponding projection in the reconstruction algorithm can always be adjusted to give a different treatment to those unknown regions, most commonly by leaving them untouched. If a missing region is isolated and smaller than a typical speckle size, it is likely to
be determined by redundant information contained in the rest of the oversampled diffraction pattern. A missing data region at the periphery of the diffraction pattern is also typically less problematic as most of the power is generally concentrated away from the edges.

Important difficulties are known to occur in a specific situation. The missing data region has to be large (compared to the speckle size) and situated close to the center of the diffraction pattern. In addition, the support must have a low perimeter to area ratio (which is somewhat similar to having a low constraint ratio). In this special case, there can be degrees of freedom that are very weakly constrained both by the support and the diffraction data. Figure 4.3 gives an illustration of this situation in one-dimension. If the missing data region is larger than the size of a speckle, there exist functions well localized both in Fourier and direct space. These functions are constrained by the Fourier and direct space constraints only in the regions outside the missing data region and the support respectively. Although perfectly localized functions in both Fourier and direct space are impossible, some functions have a power in the constrained region of the same order or less than the noise in the diffraction data. Their amplitude is therefore nearly free to vary.

Conditions leading to the occurrence of these weakly constrained degrees of freedom are specific but not uncommon. The use of unconstrained modes, defined below, was necessary when reconstructing the image of a yeast cell from its diffraction pattern (see section 5.1). Attempts at reducing problems caused by missing data regions are underway. The use of a partially opaque (Jacobsen, 2005) or even rotating beam-stop has been proposed to reduce the strong low-angle signal to a non-damaging and detectable level.
Figure 4.3: A one-dimensional illustration of unconstrained degrees of freedom. (a) A diffraction pattern with a missing data region in the center. (b) The support, in direct-space, shown as a white region in between gray bands. The bell-shaped curve in (a) is localized enough to have almost negligible power in the region where the data is known. The Fourier transform of this curve is also well localized and has also negligible power in the non-support region. Thus the amplitude of this mode is only weakly constrained by the combined information of the data and the support.
Unconstrained modes

Although data is insufficient to determine exactly the amplitude of the weakly constrained degrees of freedom, one can at least attempt to characterize and control them. For this purpose, we let the constrained power of a given function $\Psi$ be given by:

$$ (\Psi, W\Psi) = \int_{r \notin S} dr |\Psi(r)|^2 + \int_{q \notin C} dq |\tilde{\Psi}(q)|^2, $$

(4.28)

where $S$ is the support, $C$ is the region of missing data in Fourier space, $\tilde{\Psi}$ is, as usual, the Fourier transform of $\Psi$, and $(\cdot, \cdot)$ is the scalar product,

$$ (\Psi_1, \Psi_2) = \int \Psi_1(r)\Psi_2^*(r)dr. $$

(4.29)

Minimizing the functional (4.28), with the constraint that $\Psi$ is normalized, gives the function that is the least constrained. This would be the first of possibly many independent modes that are also weakly constrained. Equation (4.28) defines $W$ as a self-adjoint operator, with a spectrum bounded by 0 and 2. The eigenfunctions form a set of orthogonal modes, of which those corresponding to the lowest eigenvalues are negligibly constrained in both direct and Fourier space. When the regions $S$ and $C$ are sufficiently large, the relevant features of the spectrum yield to a semi-classical, wavepacket analysis. Consider wavepacket modes inside the region $S$ in direct space. The density of these modes in Fourier (momentum) space is given by the well known expression (Rayleigh, 1900; Ashcroft and Mermin, 1976)

$$ dM = \frac{A_S}{(2\pi)^D} dq, $$

(4.30)

where $A_S$ is the area (or volume) of the region $S$ and $D$ is the number of dimensions of the problem (typically 2 or 3). To count the modes that also have negligible
power in the region $C$ of Fourier space, we integrate (4.30) over $C$:

$$M = \int_{q \in C} dM = \frac{A_S \widetilde{A}_C}{(2\pi)^D},$$

(4.31)

where $\widetilde{A}_C$ is the area of $C$ in Fourier space. When $M \gg 1$, the condition for semi-classical analysis to be valid, there will be approximately $M$ negligibly constrained modes with $W \approx 0$. We adopt the simpler terminology “unconstrained” for these modes from now on.

The structure of the spectrum of $W$, for modes with higher constrained power, is also simple. There is an infinite near-degeneracy of modes with $W \approx 1$, and relatively few modes with $0 < W < 1$. This too can be understood in semi-classical terms. Consider wavepackets of very small width in the region $S$. For suitably small widths, the number of independent wavepacket modes within $S$ can be made arbitrarily large. On the other hand, all of these modes will have Fourier transforms with widths so broad that close to the maximum penalty, $W \approx 1$, is incurred by the second term in (4.28).

In the discrete (sampled) problem, the number of samples $N_S$ and $N_C$ are given by

$$\widetilde{A}_C = N_C (\Delta q)^D$$

(4.32)

and

$$A_S = N_S (\Delta x)^D.$$  

(4.33)

Using equation (2.14), the number of unconstrained modes can be expressed as

$$M = \frac{N_C N_S}{N},$$

(4.34)

where $N$ is the total number of samples, or simply as

$$M = \frac{N_C}{\sigma},$$

(4.35)
with $\sigma$, the oversampling ratio (defined in section 2.4.1). As in Weyl’s formula, for the density of modes of the Laplacian on a bounded region (Weyl, 1911), there are corrections associated with the boundary that diminish the density of modes. We therefore expect (4.34) to be an overestimate when the regions $S$ and $C$ have large surface to volume ratios.

**Remark**

The eigensystem associated with (4.28) can be written as an integral equation:

$$\lambda_n \Psi_n(r) = W \Psi_n(r) = \chi_S(r) \Psi_n(r) + \int dr' X_C(r - r') \Psi_n(r'),$$

where

$$X_C(r) = \mathcal{F}^{-1} \chi_C(q)$$

and $\chi_S(r)$ and $\chi_C(q)$ are characteristic functions over their respective sets. This formulation emphasizes the similarity with the important question of how much information can be transmitted in a band-limited optical system (see for instance Gabor, 1975). In this case, one seeks modes that satisfy (Slepian and Pollack, 1961)

$$\lambda_n \Psi_n(r) = \int_{r \in S} dr' X_C(r - r') \Psi_n(r'),$$

whose solutions are given by the so-called generalized prolate spheroidal wavefunctions. The eigenvalues of $W$ have a very similar behavior as those of the system (4.36). The current problem is however mathematically different, since a support is known in both direct and Fourier space, so that the modes cannot be strictly band-limited. The number of modes $M$ is analogous to what is called the “Shannon number” or the number of degrees of freedom (Gori and Guattari, 1973; Gabor, 1975) of a coherent optical system in the classical optics eigenproblem (4.38).
Mode calculation

The number of samples in $S$ and $C$ typically satisfy the relationship $M \ll N_C \ll N_S$, and neither the pixels in the support nor the missing samples in the diffraction pattern are well matched in number to serve as efficient bases for the unconstrained modes. Moreover, since the matrix elements of the operator $W$ are not sparse in either of these bases, their computation would be costly. A better basis, and one that exploits the symmetry between direct and Fourier space, is provided by the quantum harmonic oscillator modes. These modes have the property of having some number of oscillations within a classically limited area, yet they rapidly taper off outside the classical limit and thus can make a nice transition between missing and measured data regions.

We will shortly describe in detail how the unconstrained modes are calculated. The general idea is to expand the modes in the basis of the two-dimensional harmonic oscillator eigenfunctions:

$$\Psi(r) = \sum_i c_i \psi_i(r),$$

(4.39)

where $\psi_i(r)$ is a single-index labeling of the (two- or three-dimensional) eigenfunctions. Since we are only interested in the $M$ least constrained modes, the sum will be truncated to some multiple of $M$. The constrained power operator is evaluated in this basis,

$$\langle \Psi, W \Psi \rangle = \sum_{i,j} W_{ij} c_i^* c_j,$$

(4.40)

and diagonalized, giving the eigenvalues (constrained power) and the corresponding eigenvectors.

We now proceed with the detailed description of the mode calculation method. From now on, we will assume that the data and support are in two dimensions;
generalization to 3D is straightforward.

The one-dimensional harmonic oscillator eigenfunctions are

\[ \psi_n(x) = \frac{1}{\sqrt{\pi a}} H_n(x/a) e^{-\frac{1}{2}(x/a)^2}, \]  

(4.41)

where \( H_n \) is the \( n \)th Hermite polynomial, and \( a \) is a parameter setting the width of the function and known as the \textit{classical turning point} in the original quantum context. Because the harmonic potential is separable in cartesian coordinates, two-dimensional eigenfunctions are simply formed by a multiplication,

\[ \psi_{m,n}(r) = \psi_m(x)\psi_n(y) \]

\[ = \frac{1}{\sqrt{\pi a_x a_y}} H_m(x/a_x)H_n(y/a_y) e^{-\frac{1}{2}(x^2/a_x^2+y^2/a_y^2)}. \]  

(4.42)

The Fourier transform of \( \psi_{m,n} \) has a very similar form:

\[ \tilde{\psi}_{m,n} = i^{m+n} \sqrt{\frac{a_x a_y}{\pi}} H_m(q_x a_x)H_n(q_y a_y) e^{-\frac{1}{2}(q_x^2/a_x^2+q_y^2/a_y^2)}. \]  

(4.43)

In practical applications, it is useful to ensure that the regions \( S \) and \( C \) are centered, that is

\[ \sum_{r \in S} r = \sum_{q \in C} q = 0. \]  

(4.44)

With this assumption, the widths \( a_x \) and \( a_y \) can be defined in terms of the radius of gyration of the regions \( S \) and \( C \):

\[ a_x^4 = \frac{\sum_{r \in S} x^2}{\sum_{q \in C} q_x^2} \]  

(4.45)

\[ a_y^4 = \frac{\sum_{r \in S} y^2}{\sum_{q \in C} q_y^2} \]  

(4.46)

It can be shown that this choice is optimal in the case where \( S \) and \( C \) are two disks.

For computation purposes, it is necessary to expand the unconstrained modes over a finite basis; how to choose the set of \((m, n)\) to include in the sum depends on
$M$, the approximate number of unconstrained modes, and on the relative scaling of $a_x$ and $a_y$. If $S$ and $C$ are both much wider in one direction, higher order Hermite polynomials can be accommodated along this direction. Since the ordering of the energy levels of the harmonic oscillator has a similar dependence on the ratio $a_x/a_y$, we suggest that the maximum values of $m$ and $n$ be based on a cutoff on the harmonic oscillator energy. The energy of a given eigenstate $(m, n)$ is given by

$$E_{m,n} = C \left( m/a_x^2 + n/a_y^2 \right),$$

(4.47)

giving

$$m_{\text{max}} = \frac{E_{\text{max}}}{C} a_x^2, \quad n_{\text{max}} = \frac{E_{\text{max}}}{C} a_y^2.$$ 

(4.48)

The number of levels with energy lower than $E_{\text{max}}$ is roughly

$$N = m_{\text{max}} n_{\text{max}} / 2.$$  

(4.49)

Setting $N = cM$, where $c > 1$ is a factor controlling the total number of modes computed in excess of $M$, we find

$$m_{\text{max}} = \sqrt{2cM} \frac{a_x}{a_y}, \quad n_{\text{max}} = \sqrt{2cM} \frac{a_y}{a_x}.$$  

(4.50)

Once each harmonic oscillator eigenfunction is labeled with a single index $\psi_i(r) = \psi_{m_i,n_i}(r)$, the matrix $W_{ij} = (\psi_i, W \psi_j)$ can be computed, and then diagonalized. The $n$th eigenvector of this matrix, $c_i^{(n)}$, gives the $n$th mode:

$$\Psi^{(n)}(r) = \sum_i c_i^{(n)} \psi_i(r).$$  

(4.51)

Mode replacement

Once the degrees of freedom are well identified in the form of unconstrained modes, there remains the question of what should be done with them. Unavailability of
Figure 4.4: Examples of unconstrained modes, as computed with the method described above. Each mode is presented as a Fourier pair (direct space on the left, Fourier space on the right). The support and the missing data region are shown in (a) In this case, $M = 16.0$, but there are 8 modes such that $W < 5\%$. For each of the 5 modes shown, the constrained power is given in parentheses.
Figure 4.5: Examples of unconstrained modes, with a support less trivial than in figure 4.5. Each mode is presented as a Fourier pair (direct space on the left, Fourier space on the right). The support and the missing data region are shown in (a). As in figure 4.4, $M = 16.0$, but only 6 modes have a constrained power less than 5%. For each of the 5 modes shown, the constrained power is given in parentheses.
additional data makes these modes nearly free. Because they correspond to low spatial frequency contrast, their presence might not be essential for the interpretation of the reconstruction. If this is the case, the amplitude of the unconstrained modes can be simply set to zero, or modified to make the image clearer. This last approach was taken in the yeast cell reconstruction, discussed below (see Figure 5.5, in section 5.1).

4.2 Algorithm limitations

4.2.1 Data centering

The position of the center of a measured diffraction pattern is never exactly known. In some instances, it is not measured and needs to be deduced from the diffraction pattern itself. Not knowing how to center a diffraction pattern for a reconstruction can add difficulties, independent of the quality of data.

From simple properties of the Fourier transform, it can be seen that a translation of a diffraction pattern by $q_0$ introduces a phase ramp $e^{iq_0 \cdot r}$ multiplying the real-space image. As this phase ramp does not affect the support constraint, diffraction data miscentering has no effect on the convergence of general complex-valued reconstructions. With some assumptions on the reconstructed image, the phase ramp can be factored out \textit{a posteriori} (see for instance the results in section 5.2). Difficulties arise when one wishes to use \textit{value constraints} (see section 3.4) within the support. For instance, an optically thick specimen is expected to give an exit wave with pixel values confined to half of the complex plane, owing to the fact that the imaginary part of the refractive index is always negative. A diffraction pattern offset by only a fraction of one detector pixel can make this
value constraint unusable. Miscentering can also make harder the assembly of a tilt-series into three-dimensional datasets.

In an ideal experiment, the center of a diffraction pattern can be easily identified as the brightest pixel resulting from the central incident beam. In this ideal case, one only needs to determine where the center should lie within one pixel — a less essential information when the Fourier space grid spacing is small compared to the speckle size (high oversampling ratio). Most of the time however, a portion of the diffraction pattern center is lost, due to saturation of the detector or the use of a beam-stop (see also section 4.1.3 above, about missing central data). The center of the diffraction pattern has, then, to be determined using peripheral information.

When the image to reconstruct is known to be real-valued, the diffraction pattern is centrosymmetric; this information constrains the position of its center and it simply suffices to compute a cross-correlation of the diffraction pattern with its inversion through the presumed origin. Without centrosymmetry, and with the central portion of the dataset missing, only approximate methods can be used. To our knowledge, no well established and well documented method exists yet for this problem.

4.2.2 Support shape determination

Diffraction microscopy requires that the support of the specimen be known. In some rare cases, additional information (such as a low-resolution image) provides a good estimate of the support. Otherwise, the support needs to be determined directly from the data.

When introducing diffraction microscopy in section 2.4.1, we have hinted how the autocorrelation, computed directly from the measured diffraction pattern, can
be a convenient tool. For the support determination, one uses the autocorrelation support. In practical applications, the autocorrelation support is represented as a boolean array, equal to the characteristic function $\chi_A$ (see section 2.4.1) sampled on a grid.

Support determination and refinement techniques follow two main schools of thought, one more theoretic, the other more pragmatic.

There is already an important body of work on theoretic methods for the determination of tight supports from the autocorrelation (Fienup et al., 1982; Brames, 1987; Crimmins et al., 1990; Voicu et al., 1997). In general, the support shape is not uniquely determined by the autocorrelation contour (see figure 4.6 for an example). One is therefore interested in finding minimal support estimates that can still accommodate all possible supports. Such objects have been called “locator sets” by Fienup et al. (1982). Among ways of generating locator sets, the simplest is to use the parallelogram whose size is half the size of the parallelogram that contains the autocorrelation support. Another method worth noting is the triple-intersection method (Crimmins et al., 1990), especially useful for non-convex autocorrelation supports. This method is illustrated on figure 4.7.

Theoretic methods generally give rather loose supports when the autocorrelation support is convex. It is then necessary to resort to a more pragmatic approach.
Figure 4.7: Example of triple-intersection method for the determination of a tighter support. The method is mostly useful with non-convex autocorrelation supports. Figure 4.7(c) illustrates that a support estimate ("locator set") is found by the intersection of three copies of the autocorrelation support, translated according to a well-defined rule (see Crimmins et al., 1990, for more information).
to find a good support. In any case, when no value constraint is used, quick convergence of the iterative algorithms seem to require relatively tight supports (see also the next section for more details). Partially converged reconstructions can indeed be used for refinements of the support. This refinement can be done interactively, by repeatedly letting the reconstruction algorithm evolve for a certain (typically small) number of iterations and tightening the support based on the partial reconstructions. The fact that this method works well with “eye-ball” refinement suggests that it can be automated and done at the same time as the reconstruction progresses, with appropriate filters on current estimates. This idea is the essence of the “shrink-wrap” algorithm (Marchesini et al., 2003), which reduces the size of the support based on a thresholding operation (generally preceded by a low-pass filtering). Monitoring of the error metric [given by eq. (3.10) in the case of the difference map] controls the rate of shrinking. This method is known to work well on high-contrast specimens (it was used, for instance in the three-dimensional reconstruction by Chapman et al., 2006a), but could be more unstable with weaker objects that have soft edges.

4.2.3 Complex-valued reconstructions

It has been mentioned already a few times that diffraction microscopy can succeed with both real-valued and complex-valued reconstructions. The latter is easily identified by a non centrosymmetric diffraction pattern.\(^5\) It has been observed by many that complex-valued reconstructions are noticeably harder than real-valued, positive ones (Lane, 1987; Fienup, 1987; Millane, 1990; Spence et al., 2002; Chap-

\(^5\)It should be noted however that the converse is not true: not all centrosymmetric diffraction patterns correspond to real-valued reconstructions.
man et al., 2006a). In terms of degrees of freedom, this observation may seem surprising, as lifting the redundancy of symmetric diffraction patterns doubles the constraints on the reconstruction while also doubling the degrees of freedom. A part of the increase in difficulty can certainly be attributed to the enlargement of the search space. We will explain shortly that the nature of the direct-space constraint (the support), in relation to the analog of focus in traditional microscopy, is responsible for the difficulty as well. To give a complete picture, we first describe the range of validity of the scattering approximations introduced in chapter 2; this will provide a simple way of explaining the causes of complex-valued exit-waves.

**Origins of complex-valued images**

Figure 4.8 illustrates the regions of validity of various approximations for the propagation of a wave through a specimen. The two axes of this figure are the specimen thickness \(t\), in units of wavelengths, and the magnitude of \(\delta n\). As the index of refraction varies inside a specimen, a formal definition of the latter can be given by the average over the specimen volume:

\[
\langle \delta n \rangle = \frac{1}{t} \int \delta n(r) \, dz.
\]

The main approximations covered in chapter 2 are the eikonal, Born and Rytov approximations, given by equations (2.46), (2.65) and (2.79) respectively.

The eikonal approximation is valid only for specimens that are thin enough to be fully in focus or, in other words, for cases where the resolution is low enough for the Ewald sphere to be assumed flat. Since a speckle has longitudinal width of \(\pi/t\), the Ewald sphere can be approximated as flat as long as \(|k - \kappa| = k(1 - \cos \theta) < \pi/t\), that is

\[
t/\lambda < \frac{1}{2(1 - \cos \theta)} \approx \frac{d_{\text{min}}^2}{\lambda^2}.
\]

(4.53)
Figure 4.8: Regions of validity of various wave scattering approximations, as a function of the average index of refraction $\langle \delta n \rangle$ and the specimen thickness in units of wavelengths, $t/\lambda$. The maximum scattering angle is assumed to be constant for this graph (the resolution is not constant). The domain boundaries are explained in the text. The projection approximation [Eq. (2.49)] applies only in A. The Eikonal approximation [Eq. (2.46)] applies in A and B. Born approximation [Eq. (2.65)] applies in A and C. Rytov approximation [Eq. (2.79)] applies in A, B, C and D. No analytic expression is known for region E. The boundary between D and E is smooth to express the ambiguity on the range of validity of the Rytov approximation.
This expression defines the depth of focus, in terms of the resolution, as\(^6\) \(z = t_{\text{max}} = \frac{d_{\text{min}}^2}{\lambda}\). The region where \(t < t_{\text{max}}\) (with an arbitrary 5° maximum scattering angle) is identified by the labels A and B on figure 4.8.

The diagonal line in fig. 4.8 separates optically thin from optically thick specimens (defined in section 2.2.3 on page 18), a specimen being optically thin if

\[
\langle \delta n \rangle t < 0.1\lambda.
\]  

(4.54)

The factor of 0.1 corresponds to a 5% error in the approximation of \(e^{i\phi}\) by \(1 + i\phi\). The line defines the region where the Born approximation is valid (labeled A and C). The Rytov approximation is known to be valid in all cases A, B and C, and probably beyond, represented as a region with a fuzzy boundary (D).

It is now possible to address the question of complex-valued exit-waves: the perturbation of the exit-wave can be real-valued only if it is both physically thin and optically thin (region A in figure 4.8). It then satisfies the projection approximation, given by equation (2.49) (page 18). Even in this case, the exit-wave perturbation will be real-valued only if the ratio of imaginary to real part of the index of refraction is constant: then, one can write \(\delta n(r) = e^{i\phi} |\delta n(r)|\) and the global phase factor is lost in the measurement and does not affect the reconstruction. All these conditions are typically satisfied in hard x-ray scattering experiments with inorganic specimens made of only one atomic species; at short enough wavelengths, all materials become transparent, and condition (4.54) becomes easier to satisfy. Shorter wavelengths also mean that the Ewald sphere radius increases, making small-angle scattering sufficient to attain a satisfying resolution. In this special case, the reconstructed is not only real but also positive, making possible the use

\(^6\)This result can be seen as the distance \(z\) such that a feature of size \(d_{\text{min}}\) is propagated to the Fresnel number \(f = 1\) [see eq. (2.41)].
Figure 4.9: Illustration of different cases of complex-valued images. (a) The exit-wave is complex-valued because the object is thicker than the optical depth (region B in figure 4.8). (b) The exit-wave is complex-valued because the object is optically thick (region C in figure 4.8). (c) The object is neither thick nor optically thick (region A in 4.8), but is made of two components having a refractive index with different ratios of real part to imaginary part.

Reconstruction difficulties

In the previous section, we have shown that complex-valued reconstructions are the norm more than the exception. Being able to solve reconstruction difficulties in this situation is therefore of very high relevance.

Two decades ago, Fienup (1987) noted that complex-valued reconstructions are harder than real-valued positive ones, especially when using a loose support, or when the edges of the object are not sharp. In this case “harder” means that some
may simply fail to converge in a given amount of time. In view of the discussion in chapter 3, on the search mechanism of the difference map, the main explanation may reside in the fact that looser constraints lead to slower dynamics, which in turn lead to longer searches before convergence. Positivity, after all, is sometimes a strong constraint, playing an important role in Fienup’s initial work (Fienup, 1978, 1982) and in crystallographic “direct” methods (Karle and Hauptman, 1950; Sayre, 1952b). In the absence of noise, removal of the positivity constraint is probably the most important (if not only) reason for reconstructions to become harder (McBride et al., 2004). In particular, we have done small numerical experiments showing that there is no obvious difference in difficulty (measured in terms of the average number of iterations needed to reach convergence) between complex-valued and real-valued non-positive reconstructions.

With noisy data, an additional source of ambiguity in the reconstruction is related to focus. Let us recall that the free-space propagation of a two-dimensional wavefront $\psi(r)$ is given by a multiplication, in Fourier space, with the Fresnel propagator $\exp(-i \frac{1}{2} z q^2 / k)$. The argument of this exponential is quadratic in $q$; a reality constraint can “pin” the focus plane because even a small defocus distance $z$ eliminates the Friedel symmetry in the wave. This can also be seen in real space, expanding to first order in $z$. The propagation of a wave $\psi(r)$ can be written

$$\psi(r; z) \approx \psi(r; z = 0) + \frac{iz}{2k} \nabla^2 \psi(r; z = 0), \quad (4.55)$$

and the perturbation of $\psi$ is, to first order, purely imaginary. When $\psi(r)$ is complex, this “pinning” is lifted and the focus plane is instead controlled by the tightness of the support, as larger and larger defocus distances make the wave leak further outside the support. This analysis makes apparent the need for very tight supports. Still, if noise is present, the incompatibility of the support and Fourier-
space constraints effectively allow some wave amplitude in the non-support region, allowing a range of defocus values as if the support was looser. The combined effect of this defocus range and the averaging method is described in appendix B. Marchesini et al. (2005) point out that higher-order phase aberrations (astigmatism, coma, ...) and even phase vortices can also be considered in this type of analysis.

4.2.4 Crystalline domains

The effect of the small shape of crystals on the broadening of Bragg peaks has been described very early in the history of x-ray crystallography (Patterson, 1939; Ewald, 1940). The great interest for atomic-resolution imaging has incited many scientists to attempt reconstructions on datasets obtained with wavelengths short enough to reach such length scales. Lately, both x-rays (Williams et al., 2003; Pfeifer et al., 2006) and electrons (Zuo et al., 2003) have been used to reconstruct small crystalline specimens (although with greatly varying approaches). In this section, we will discuss the peculiarities of diffraction patterns containing highly periodic domains, and will explain how this can lead to complications – to the point of throwing doubt on results that have appeared in the recent literature.

Diffraction from a crystalline specimen

Specimens having nearly perfect periodicity in one, two or three dimensions are common at the atomic scale. It is well known that periodicity results in interference effects leading to dramatic intensity redistribution. For a long time, this effect has been used in crystallography, where the concentration of the signal in sharp Bragg peaks has made their detection possible almost as soon as x-rays were discovered.
With the availability of brilliant and highly coherent sources (from synchrotrons and high-energy electron microscopes), it is now possible to detect the diffraction signal from a crystalline specimen small enough to meet the oversampling conditions (see section 2.4.1). These diffraction patterns have speckles, as those from non-crystalline specimens, and yet exhibit sharp concentrations of the intensity in specific regions (points for 2D and 3D periodicity, or lines for 1D periodicity). As we will explain in more detail, this intensity concentration, a blessing in crystallography, can lead to serious problems in diffraction microscopy.

Figure 4.10 gives a two-dimensional illustration of a small crystalline domain and its Fourier transform. To better understand this figure, we consider the case of a two-dimensional exit-wave produced by a small crystalline domain, such as what can be obtained in a transmission electron microscope in “diffraction mode”. The exit wave has the form

\[
\psi(r) = e(r)P(r),
\]

where \( P(r) \) is a function periodic in two dimensions and \( e(r) \) is an envelope having a finite support. The far-field diffraction pattern of this wavefield is defined by its Fourier transform:

\[
\tilde{\psi}(q) = \frac{1}{2\pi} (\tilde{e} * \tilde{P})_q.
\]

Since \( P \) is periodic, its Fourier transform is a discrete sum of Dirac deltas,

\[
\tilde{P}(q) = f(q) \sum Q 2\pi \delta(q - Q),
\]

where the sum runs over \( Q \), the points in the reciprocal lattice. The function \( f(q) \) is the Fourier transform of the density variations within a unit cell (the structure factor). Because of the structure of \( \tilde{P} \), the convolution (4.57) is trivial:

\[
\tilde{\psi}(q) = \sum Q f(Q) \tilde{e}(q - Q).
\]
The diffraction pattern is made of multiple copies of $\tilde{e}$, each centered on a point determined by the lattice of the crystalline domain. In the simplest case, where there is a single atom per unit cell, $f$ is a smoothly varying function encoding the electron density surrounding the atom (the “form factor”) and the blurring caused by thermal vibrations. The exact form of $f$ is not relevant for the current discussion.

Possible problems in the reconstruction of diffraction patterns of the form (4.59) originate from the large difference between the scale of the lattice and the scale of the envelope. As $\tilde{e}(q)$ becomes narrower, the signal between the peaks becomes relatively weaker. The only known reconstruction attempts (enumerated below) on this kind of specimen have the problem that the signal in this region is lost to noise.

**Reconstruction of the crystal shape**

The safest approach in applying diffraction microscopy to the kind of dataset described above is the low-resolution reconstruction of the crystal shape from one of the diffuse Bragg peaks. Assuming that the crystal has no distortion, the intensity around any Bragg peak is a perfect copy of the low-resolution intensity one would measure in the center of the diffraction pattern, with the additional advantage of not having to work around the bright central beam. This method was used successfully by Williams et al. (2003) to reconstruct the three-dimensional shape of a gold nanocluster using hard x-rays.

In cases where strain is present in the small crystal, large-scale variations of the lattice produce a measureable modulation of the diffraction pattern. It can be shown (Robinson and Vartanyants, 2001) that the density reconstructed from a
Figure 4.10: Formation of a diffraction pattern by a small crystal. This figure illustrates the terms in equations (4.56) and (4.59).
given Bragg peak $Q$ acquires an imaginary part proportional to the displacement field $u(r)$ projected along $Q$:

$$\rho'(r) \approx \rho(r) \left(1 + i Q \cdot u(r)\right).$$  \hspace{1cm} (4.60)

A reconstruction of the strain field inside a nanocluster was first reported by Pfeifer et al. (2006).

**Atomic-scale reconstruction**

Reconstructions of full diffraction patterns have recently been attempted by a few research groups. In all cases, diffraction patterns come from electron diffraction. The first is a two-dimensional reconstruction of a double-walled carbon nanotube, a structure periodic in one dimension (Zuo et al., 2003). Although the specimen was very long, the authors claimed to have satisfied the oversampling criterion by reducing the size of the probe (the illumination profile) incident on the nanotube. Figure 4.11 shows a simulation of a double-walled nanotube in a finite illumination profile. 4.11(c) shows the diffraction pattern with a mild measurement noise. Another similar reconstruction was also reported by the same group (Zuo et al., 2004). The other publication, by Wu et al. (2005), shows the two-dimensional reconstruction of a gold nanocluster from its diffraction pattern. This work was, however, flawed in many aspects and was subsequently retracted (Wu et al., 2006).

All three electron diffraction reconstructions are characterized by (1) a very low signal-to-noise level in the region surrounding the diffuse Bragg peaks and (2) a large missing data region around the center. As we will show, the lack of diffraction data between the peaks has important consequences on the uniqueness of reconstructions.
Figure 4.11: Double-wall nanotube simulation. (a) Simulation of the projected density. (b) Noise-free diffraction pattern. (c) Diffraction pattern with Poisson noise (average of 5000 counts per pixel). Interference between the Bragg spots is lost to noise.
We first consider the diffraction pattern of the form (4.59). Because of noise, the measured intensity,
\[ I(q) = \left| \sum_Q f(Q) \tilde{e}(q - Q) \right|^2, \] (4.61)
is indistinguishable from an incoherent sum of the individual Bragg spots:
\[ I(q) \approx \sum_Q \left| f(Q) \tilde{e}(q - Q) \right|^2. \] (4.62)
As a result, the relative translation of the envelope \( e(r) \) and the crystal \( P(r) \), encoded in the interference between the peaks, is lost. An eventual reconstruction would then be ambiguous at least in this respect. In the case of the nanotube reconstructions, this leads to an ambiguity in the translation of the nanotube with respect to the illumination profile.

We now turn to the case of specimens having more than one crystalline domain. The exit wave is made of the coherent sum of each individual domain, with its own envelope and its own lattice:
\[ \psi(r) = e_1(r)P_1(r) + e_2P_2(r) + \ldots \] (4.63)
From now on, we consider only a pair of domains, as the analysis is the same for more than two. Assuming that the structure factor is the same for both crystals, the far-field wavefield is
\[ \tilde{\psi}(q) = \tilde{\psi}_1(q) + \tilde{\psi}_2(q) \]
\[ = \sum_{Q_1} f(Q_1) \tilde{e}_1(q - Q_1) + \sum_{Q_2} f(Q_2) \tilde{e}_2(q - Q_2), \] (4.64)
where \( Q_1 \) and \( Q_2 \) are the lattice points of the two reciprocal lattices. The relative position of the two domains is encoded through interference between the two terms. The measured intensity is
\[ I(q) = \left| \tilde{\psi}_1 \right|^2 + \left| \tilde{\psi}_2 \right|^2 + 2\text{Re} \left\{ \tilde{\psi}_1^* \tilde{\psi}_2 \right\}. \] (4.65)
Figure 4.12: Demonstration of the interference problem with multiple crystalline domains. Figures (a), (b) and (c) show the simulated electron densities from a simple two-dimensional model of a nanocluster containing about 5000 atoms divided into three domains.\textsuperscript{8} Figures (a), (b) and (c) have the exact same atomic domains, only the relative position of the domains is changed. Figures (d), (e) and (f) show the diffraction pattern of the corresponding density on the left, assuming no noise in the measurement. The inset images show two magnified Bragg peaks. Figures (g), (h) and (i) show the corresponding diffraction patterns with a noise level reproducing typical measurement conditions. The interference between the peaks is not visible anymore.
The cross-term is lost if the noise discards most of the data between the peaks. This can be avoided if two peaks happen to lie close to each other. The interference is also always present near the center of the diffraction pattern. Unfortunately, it is very common that such data is not usable because of the strong incident beam. If no peaks are close to each other, the intensity concentration into Bragg peaks makes the measured intensity appear as an incoherent sum of multiple diffraction patterns,

\[ I(q) \approx |\tilde{\psi}_1|^2 + |\tilde{\psi}_2|^2. \] (4.66)

From this type of data, it is thus impossible to extract the relative position of the domains, compromising the reproducibility of any atomic-resolution reconstruction attempt. Figure 4.12 is an illustration of this situation. Although the specimens shown on Fig. 4.12(a)–(c) are very different, their noisy diffraction patterns [Fig. 4.12(g)–(i)] are, for reconstruction purposes, identical. This claim is directly verified on Figure 4.13, where the 3 “reconstructions” shown differ only by their phases.

On the grounds of this simple analysis, it can be concluded that the nanotube reconstructions, mentioned above, may be plausible, but are very likely not unique; information on both the position of the nanotube in the illumination function, and the relative positions of the two concentric nanotubes, is simply not present in the data. The nanotube reconstructions therefore contain no more information than what can be deduced directly from the diffraction pattern, without the phases.

Discussion

Whether diffraction microscopy (with x-rays or with electrons) will ever reach atomic resolution is a question still open to debate. For crystalline specimens, the main challenge is measuring a signal strong enough to have clear interference be-
Figure 4.13: Demonstration of the possible outcomes of a reconstruction with a noisy diffraction pattern. All three figures have the same diffraction pattern, taken to be a high-pass filtered version of that shown in 4.12(g). The high-pass filtering operation mimics the missing central data problem. The only difference between the three images are the phases. Figures (a), (b) and (c) result from a combination of the Fourier amplitudes from 4.12(g) and the phases from 4.12(a) 4.12(b) and 4.12(c) respectively.

tween the peaks. But even for non-crystalline specimens, atomic resolution requires that each atom scatters enough photons or electrons to contribute noticeably to a diffraction pattern – a condition made very hard by the damage incurred by the incident high-energy particles. Arguably, the interest of atomic-resolution imaging is the possibility of observing those atoms that deviate from the crystal order, such as atoms at the surface or along a grain boundary. In fact, one could even ask whether crystalline order is needed at all – for the purpose of atomic resolution imaging, diffraction microscopy should be able to succeed even with completely amorphous specimens.

With electrons, high signal-to-noise ratios are even harder to obtain because of the systematic error provided by the scattering of the substrate needed to maintain the specimen in place. Among ways to reduce the effect of the substrate, a possibility is the use of a crystalline substrate which, combined with a large il-
lumination profile, would give bright but very well localized Bragg spots in the
diffraction pattern, leaving free most of the space for the detection of the diffuse
scattering of the specimen. This approach was used for direct-space imaging by
Iijima (1977a,b).

We conclude this section by mentioning that the interference between Bragg
spots has been a sought source of phase determination for many years. Long ago,
Hoppe (1969) suggested a method where a mask (or a thin slot) placed in front
of a crystal would lead to a spreading sufficient to make Bragg peaks interfere.
This method, called “ptychography” has variants still well alive nowadays (Nellist
et al., 1995; Plamann and Rodenburg, 1998).
Chapter 5

Reconstructions and simulations

This section contains examples of reconstruction from experimental data. Much of the arsenal of tools and analysis described above had to be developed to reach satisfactory results. Difficulties encountered while attempting the reconstructions were in fact the primary motivation for the methods described above.

We will first present the reconstruction of a yeast cell image from its x-ray diffraction pattern. Then, we will present reconstructions similar in essence, but from data gathered with a simple optics setup. These two cases are “classical” diffraction microscopy, as they use oversampling to encode the phase information. Departing slightly from this framework, we will then present the reconstruction of the electron density of a quasicrystal from its x-ray diffraction data. In this case, the very low occupation density of a six-dimensional unit cell is the key additional piece of information needed to retrieve unique phases.

The second part of this chapter contains simulations of soft x-ray propagation through a model yeast cell, and of a new algorithmic approach to some recent experiments involving multiple datasets.

5.1 The yeast cell reconstruction\textsuperscript{1}

X-rays and high-energy electrons, with their short wavelengths, are the best suited candidates for the pursuit of ever higher resolutions in microscopy and tomographic reconstructions. While electron microscopy can already offer nearly atomic resolu-

\textsuperscript{1}Results of most of this section have appeared in publications (Shapiro et al., 2005; Thibault et al., 2006).
tion for macromolecules, strong interactions limit its use to specimens thinner than about 800 nm (Grimm et al., 1998). X-rays have much higher penetration depth and can thus accommodate thicker specimens. However, the resolution of x-ray microscopy and tomographic reconstructions is limited by the quality of the optical components. For these reasons, diffraction microscopy could become the method of choice for the high resolution three-dimensional imaging of large organic specimens, such as whole biological cells. The reconstructions reported in this section constitute a step in this direction.

Figure 5.1 shows the diffraction pattern used for reconstruction. It was measured with an apparatus (Beetz et al., 2005) located at the Advanced Light Source in Lawrence Berkeley National Laboratory. The specimen, a freeze-dried dwarf yeast cell about 2.5 μm wide, was illuminated with 750 eV x-rays. The spatial and temporal coherence of the beam was enhanced with the combined effect of an off-axis zone plate and a pinhole, the latter being placed 25 mm away from the specimen. The pinhole has the double effect of defining the transverse coherence and selecting a small region of the beam dispersed by the zone plate. The resulting incident wavefront was nearly monochromatic: $\Delta E/E = 0.2\%$ and also had a sufficient transverse coherence. Because of the very strong central beam, a beam stop had to be placed in front of the detector array, a $1300 \times 1340$ CCD with 20 μm x 20 μm pixels. The specimen to CCD distance was 15 cm. The diffraction pattern shown in figure 5.1 is a combination of 26 different datasets, taken at different exposure times, in order to cover most of the dynamic range of the pattern (over 6 orders of magnitude).

Although, as fabrication techniques for x-ray zone plates are progressing, the resolution limit in traditional microscopy may soon reach 10 nm (Le Gros et al., 2005; Chao et al., 2005).
Figure 5.1: Soft x-ray (\(\lambda = 1.65\) nm) diffraction pattern (log scale) of a freeze dried yeast cell (Shapiro et al., 2005). The array extends to \(q_{\text{max}} = \pi(20.7\) nm\)^{-1} on the sides, giving the corresponding real-space array 10.3 nm wide pixels. Inset, left: Magnified portion of the diffraction pattern showing the speckles. Inset, right: Magnified central region showing the diamond-shaped missing data region.
The diffraction pattern shown in figure 5.1 is clearly asymmetric under rotation of 180°; as mentioned above (section 4.2.3), this is an indication that the reconstruction can only be assumed to be complex-valued. In section 4.2.3, we have enumerated three main mechanisms through which a wave is complex-valued: intrinsic refractive index variations, thickness (defocus) and optical thickness. Simulations of a model cell (section 5.4) indicate that the most important reason for departure from reality is the optical thickness of the cell. For this experiment, the phase winding through the center of the cell is expected to be of the order of $\pi/2$, a clear indication that the diffraction pattern comes from a strongly phase-shifted exit wave, and that the Born approximation is not valid.

The small dark region in the center of the diffraction pattern is the central missing data region. As will be seen below, this region is large enough to allow a few weakly constrained modes (see section 4.1.3).

The autocorrelation of the cell image is shown in figure 5.2. The image shown has been high pass filtered to reduce the effect of the sharp discontinuity caused by the missing data region. Values above a threshold have been clipped to avoid stretching the grayscale over too wide a range, and to emphasize the sharply defined autocorrelation support. The diameter of this support is twice that of the actual yeast cell. We also have direct evidence of the cell from the faint “ghosts” surrounding the central oval. These are the result of phase interference between the yeast cell and isolated point-like scatterers located a few cell diameters distant; this is an unintentional, and very faint, Fourier hologram (see section 2.4.2).

Estimating the area in pixels of the support, and considering other quantities from the experimental setup, we can quickly compile the following facts prior to the reconstruction:
Figure 5.2: Real part of the autocorrelation. This image is the high-pass filtered inverse Fourier transform of the diffraction pattern shown in figure 5.1 (the high-pass filter reduces the effect of the sharp discontinuity due to missing central data). The central oval-shaped structure is the autocorrelation of the cell. Its contour is used to determine the size of the support and to calculate the overdetermination ratio. The surrounding shapes are faint images of the cell caused by the interference with point-like scatterers around the cell.
1. The oversampling ratio [eq. (2.95)] is about 25, or about 5 in each dimension.

2. The overdetermination ratio [eq. (2.97)] is almost exactly 2 because of the nearly centrosymmetric support.

3. The Fresnel number [eq. (2.41)] for the 2.5 μm cell is \( f = 0.025 \), which means that the reconstruction is expected to have a quadratic phase factor going to \( \pi f/4 = 0.006\pi \) at the edges of the cell, a negligible effect compared to, for example, the weakly constrained (missing) modes.

4. Assuming that the wavefront incident on the pinhole is flat, the divergence of the wavefront incident on the cell introduces an additional quadratic phase factor, with a phase going from 0 to about \( 0.075\pi \) on the cell’s edges, an effect more important but still negligible, considering the missing modes.

5. The number of unconstrained modes [eq. (4.35)] is \( M = 14.5 \) (since the number of pixels in the missing region is \( N_C = 362 \)). This is an overestimation and only 4 were finally taken to be underconstrained.

5.1.1 Reconstruction procedure

The shape of the support was gradually refined, starting from a rectangle half the dimensions of the autocorrelation. The difference map was used for only a few hundred iterations, with simple support and modulus projections, then the Fourier estimate was used to guess the shape of a new support. Supports were simply refined by hand, with the use of a drawing program (“the gimp”). As the support improved, the boundary of the cell became more sharply defined, which in turn helped to further refine the support.
Early reconstructions were attempted with various value constraints, but with only partial success because of lacking information. As suggested by propagation simulations through a model cell, the pixel values of the exit wave are expected to lie within a slowly broadening spiral, as shown on figure 5.22(b), and as is probably observed in the optical experiment below (5.17). Problems with a value constraint forcing the pixels to lie within this spiral are two-fold. First, the allowed region within the support has free parameters that are hard to determine from the scattering data; the critical one is the amplitude of the incident (undisturbed) wave, which sets the radius of the circle enclosing the spiral. The second difficulty is that this distribution can be sensitive to a non-planar wavefront — in particular, the phase tilt resulting from a miscentering of the diffraction pattern and the quadratic term coming from the divergence of the incident beam and the finite specimen to CCD distance. For these reasons, value constraints were finally abandoned.

Figure 5.3: Early reconstruction using a tentative value constraint

The lack of a value constraint quickly made the unconstrained modes problem apparent. The method described above (section 4.1.3) was applied, using the best
support estimate and the known set of unmeasured data pixels. Figure 5.4 shows the 4 modes that were considered unconstrained in the reconstruction.

The absence of value constraints also creates an additional burden for the computation of the average: the global phase factor, if let free to vary, could bias the average and slowly reduce the overall amplitude of the reconstruction. Fixing the overall phase factor is as simple as computing the phase of the pixel average of the estimate, and then rotating it to a chosen value. This procedure is complicated by the unconstrained modes, whose amplitudes are also varying nearly freely, thus affecting the aforementioned pixel average. The procedure adopted to address these problems consists in first zeroing the amplitudes of the modes,

$$
\Psi_0 = \Psi_F - \sum_{n=0}^{M-1} (\Psi_F, \Psi^{(n)}) \Psi^{(n)}, \quad (5.1)
$$

where $\Psi_F$ is a current Fourier estimate, and then rotating the phase,

$$
\Psi_0 \rightarrow \Psi_0 e^{-i\alpha}. \quad (5.2)
$$

The phase $\alpha$ is defined by

$$
\alpha = \arg \left( \langle \Psi_0 \rangle_S \right), \quad (5.3)
$$

where the average $\langle \rangle_S$ is taken only on the support pixels, to avoid a bias from the very large number of pixels outside the support.

The resulting average $\Psi_0$, with the modes removed, is shown in figure 5.5, and figure 5.6 is the difference map error for the reconstruction.

The unnatural appearance of image 5.5(b) caused mostly by the removal of the first mode [Fig. 5.4(b)], prompted the use of an ad hoc rule for the mode amplitude replacement. The chosen rule dictates that the reconstruction be as flat (in values) as possible within the support. Thus, the final reconstruction $\bar{\Psi}$ is

$$
\bar{\Psi} = \bar{\Psi}_0 + \sum_{n=0}^{M-1} a_n \Psi^{(n)}, \quad (5.4)
$$
Figure 5.4: The five least constrained modes for the yeast cell reconstruction. Each mode is presented as a Fourier pair (direct space on the left, Fourier space on the right). The support and the missing data region are shown in (a). For each of the 5 modes shown, the constrained power is given in parentheses. In the reconstruction, only the first 4 modes were used.
Figure 5.5: Yeast cell reconstruction prior to mode replacement. (a) Reconstruction including the unconstrained modes. (b) After the unconstrained mode amplitudes have been set to 0.

where the modes’ amplitude $a_n$ minimize the variance of the pixel values within the support:

$$\langle (\Delta \Psi)^2 \rangle = \left\langle |\Psi - \langle \Psi \rangle_S|^2 \right\rangle_S. \quad (5.5)$$

The resulting reconstruction is shown in figure 5.7.

The variance-minimizing rule was chosen mainly for aesthetic reasons. The arbitrariness in this definition is a well-controlled one: only the 4 least constrained modes — with no information available to constrain them — are readjusted. Besides, the rule itself is arbitrary but its definition is precise; results are therefore easily reproducible.

### 5.1.2 Reconstruction resolution

Defining the resolution of the reconstruction is at first not obvious. While resolution is limited by the apparatus in conventional microscopy and by the quality of
Figure 5.6: Difference map error ($\epsilon_n$) for the first 10,000 iterates (thin line). The bold line is a running average (over a 1,000 iterate window) to emphasize the decay of the transient. The dotted line is the overall average, $\bar{\epsilon} = 0.384$. The sharp initial decrease of the error, usually an indicator of near-convergence, occurs within the first 50 iterations. Inset: the difference map error for the totality of the run. As explained in the text, the overall scale of the error can be considered as arbitrary because of the missing central data.

the specimen in x-ray crystallography, both effects play a role in x-ray diffraction microscopy. The dimension of the smallest features in Fig. 5.7(a) is about 3 pixels, leading to a rough evaluation of a resolution of 30 nm.

Because of the incompatibility of the two constraints – caused by the noise and possibly other sources of systematic error – we judged that it is not sensible to define the effective resolution of a reconstruction as being equal to $d_{\text{min}} = 2\Delta x$, where $\Delta x$ is the pixel size, which is solely based on the largest recorded spatial frequency. Certainly, the imperfections of the data and of the algorithm play a role in the final resolution. As mentioned in the previous chapter, the averaging
method has the natural effect of reducing the resolution of the reconstruction. Only features that survive averaging are reconstructed reproducibly.

Figure 5.8 shows a more rigorous way of determining the resolution. For any $q$ in Fourier space, the ratio of the reconstructed intensity to the observed intensity [equation (4.27)] is reduced from unity as a result of averaging over residual phase fluctuations in the steady state of the algorithm. The solid curve on the graph is the angular average of this ratio, the so-called algorithm transfer function (ATF). In the diffraction reconstruction, this figure indicates the fraction of power in the object against power in the data as a function of spatial frequency, which is quite analogous to the information presented in an MTF curve. For comparison, two
Figure 5.8: Resolution decrease caused by averaging. The black line shows the relative decrease of the reconstructed intensity, as a result of averaging over residual phase fluctuations. The two dashed lines show the classical MTF for an incoherent imaging system with 75% efficiency and a Rayleigh resolution of 15 and 30 nm respectively. The vertical dotted line indicates the approximate boundary of the beam stop. [from Shapiro et al. (2005)]
classical optics MTF curves (dashed lines) have been added to the graph, showing that the 30 nm resolution estimate is sensible.

5.1.3 Dust

As suggested by the autocorrelation (Fig. 5.2), weak scatterers were present around the yeast cell during the measurement of the diffraction pattern. Figure 5.9 shows the result of our attempt to reconstruct the “dust” surrounding the cell. This reconstruction was realized by a simple relaxation of the support constraint: the modified support projector sets to zero only those pixels outside the support with amplitude below a predetermined threshold $c$:

$$P_{S_c}(\Psi_r) = \begin{cases} 
\Psi_r & \text{if } r \in S \text{ or } |\Psi_r| > c \\
0 & \text{otherwise.}
\end{cases}$$

(5.6)

In the yeast cell work, we found that setting this threshold to about 6 times the amplitude of the error was appropriate. Of course, too low a value for the threshold weakens the constraint and slows down the dynamics of the algorithm. This modified projector can be unstable if most of the power is not already inside the support. The reconstruction shown in Figure 5.9 was obtained by starting with the final iterate of a previous run with the regular support projection. A total of 20,000 iterations was needed to generate this average of 400 estimates. As explained in section 2.4.2, the increase in the overdetermination ratio caused by the presence of external scatterers can noticeably help the convergence of the reconstruction. Small scatterers could also help fix the focal plane in which the reconstruction is averaged, yielding possibly higher resolutions (see appendix B). In the present work, these scatterers seemed too weak to be useful (less than 0.02% of the total power comes from the reconstructed dust).
Figure 5.9: Reconstruction magnitudes of the small scatterers surrounding the cell, saturated to allow weaker points to be seen. This image is the average of 400 Fourier estimates. This result is independent of the starting iterate.

5.1.4 Discussion

Eight additional diffraction patterns, similar to the one shown on Figure 5.1, were collected with the same yeast cell being rotated by a few degrees. The same reconstruction procedure was applied to each of the datasets. While the number of datasets is insufficient to attempt a three dimensional reconstruction, it was still possible to assemble the individual images in a movie, showing the cell rocking back and forth\textsuperscript{3}. High resemblance of the various views (from completely independent reconstructions) provides a strong verification of the validity of the data and of

\textsuperscript{3}The movie is available online (Shapiro et al., 2005))
the reconstruction process.

Current efforts are directed towards realizing three-dimensional reconstructions of whole biological cells through the collection of many two-dimensional diffraction patterns. Radiation damage and optical thickness may be the most important difficulties currently standing on the way. Both issues could be solved by using hydrated specimens instead of dried ones; It has indeed been observed that frozen hydrated specimens can withstand a stronger radiation dose before showing signs of damage. In addition, a cell embedded in a slab of ice has a lower contrast than when it is in vacuum, thus possibly making the specimen optically thin enough for the Born approximation to be valid. Unfortunately, hydrated specimens preparation is currently difficult: one has to avoid the formation of ice crystals within the cell and on the exterior of the ice slab, as these can also scatter strongly and thus make reconstruction attempts more difficult, if not impossible.

5.2 Optical diffraction microscopy

In this section we present reconstructions from measurements taken at Cornell during the summer of 2006 and spring of 2007. These are optical light measurements which require much simpler experimental setups than those used with x-rays or electrons. Surprisingly little work has been published on diffraction microscopy with laser light.\textsuperscript{4} The experiment presented here is similar in many aspects to the one briefly described by Spence et al. (2002). Although visible wavelengths do not offer the possibility of very high resolution, the availability of highly coherent sources and relatively cheap equipment are good arguments in favor of laser

\textsuperscript{4}Lately, Rodenburg et al. (2007a) used an optical setup to demonstrate a new imaging method meant to be used with hard x-rays and electrons.
experiments for testing purposes.

The first datasets were measured in an optics teaching lab. The purpose of this setup was to start developing a new teaching experiment to introduce the concepts of diffraction microscopy at the undergraduate or graduate level. We will say more about teaching aspects of this experiment below. Portions of the description below was submitted for publication to the American Journal of Physics (Thibault and Rankenburg, 2007).

The second datasets were collected in the laboratory of Prof. Cohen. The purpose of this measurement was to show that the method can be used to measure accurately the thickness of a fly wing. At the moment of writing this thesis, the reconstruction was successful, but further tests and calibration need to be done to assess the precision of the method.

We will first give a quick description of the experiment setup used, and will then proceed to show the reconstruction results. We will end with a discussion of possible variations of this type of experimental setup.

5.2.1 Experiment description

The principle of the experiment is utterly simple: shine a laser on a sample and measure the resulting far-field diffraction pattern. Unfortunately, the actual realization of the experiment is more complicated. In what follows, we describe the collimator and detector parts of the apparatus, and discuss the choice and the preparation of samples.

Figure 5.10 shows the schematic of the collimator assembly. Our setup used a standard He-Ne red ($\lambda = 632.8\ nm$) laser with nominal power of 5 mW. We found that this intensity was more than sufficient for a diffraction pattern measurement.
Figure 5.10: The collimator part of the experimental setup. Next to the laser (L) is placed a sliding wedge-shaped slit for the adjustment of intensity (W), shown on top in head-on view. The beam then goes through a pinhole (P), a collimating lens (C1), and an iris (I).

The intensity of the beam was controlled with a wedge-slit placed on an adjustable sliding platform. Our first attempt to control beam intensity involved the use of two linear polarizers, with the intensity being adjusted by the relative angle between the two. However, we have found that this approach led to slow but significant intensity fluctuations (of more than 15%) due to the dynamics of the two degenerate polarization states in the laser cavity.

Scattering resulting from the use of the wedge-slit was not problematic since a pinhole was placed immediately beyond it. The very small area of the hole made its illumination effectively uniform. We found that a pinhole was necessary to ensure a good quality wavefront. Because of the high intensity of the laser, we did not need to focus the laser on the pinhole, as is often done for spatial filtering in a condenser assembly. We did find, however, that the quality of the pinhole transferred to the quality of the beam.

A collimated beam was then formed by placing a converging lens one focal distance away from the pinhole. Different beam sizes can be achieved, depending on the focal length of the lens. The radius of the Airy disk from a circular pinhole...
at a distance $l$ from it is given by (Born and Wolf, 1999):

$$R = 0.61 \frac{l \lambda}{r_0},$$  \quad (5.7)

where $r_0$ is the radius of the pinhole. In one experiment, we used a 200 mm lens and a 50 $\mu$m pinhole, which resulted in a collimated beam of about 6 mm in diameter. In another experiment, the pinhole had a diameter of 10 $\mu$m, giving a much larger Airy disk. In the first case, higher order lobes of the Airy disk were removed with an adjustable iris-shutter placed at the exit of the condenser lens. There was no need for this procedure for the larger Airy disk. The size and position of this iris was carefully adjusted to admit only the central disk of the Airy pattern. We have observed that a very clean condenser lens was needed to avoid streaks caused by random scattering from the surface of the lens.

![Figure 5.11: The detector part of the experimental setup. Downstream from the sample (S) are placed the collector lens (C2) and the CCD detector (D). The CCD detector is at the back focal plane of the lens: the incident beam is focused on the CCD and the far-field diffraction pattern is formed at the CCD plane.](image)

The samples were placed in the nearly parallel beam emerging from the collimator assembly described above. The distance between the collimating lens and the sample is theoretically irrelevant. However, the overall quality of the diffraction pattern was noticeably improved by making this distance longer. In fact, increasing the distance had a “cleansing” effect on the beam, as it made non-parallel
components of the wavefront (as, for instance, those coming from the scattering off of small dust particles on the condenser lens) propagate gradually away from the central beam.

In this experiment it was not feasible to measure the actual far-field diffraction pattern of the sample, mostly because of the small area covered by CCD detectors. An objective lens was therefore placed as close as possible downstream from the sample. The task of this last lens was to form the Fourier transform of the sample’s exit wave in the lens’ back focal plane (see the discussion at the end of section 2.2.3, page 13). The position of the CCD had to be accurately adjusted to capture the diffraction pattern in this plane. This was accomplished by removing the sample and making the focused image of the beam as small as possible on the CCD.

The choice of the objective’s focal length is limited by the oversampling condition. Using equation (2.92) and the relation $\frac{\Delta q}{k} = \frac{\Delta X}{l}$, where $l$ is the focal length and $\Delta X$, the size of one detector pixel, one finds that the size $s$ of a specimen has to satisfy the inequality

$$s \leq \frac{l\lambda}{2\Delta X}. \quad (5.8)$$

The detector we used was a low-end amateur astronomy CCD. It was not ideally suited, in many aspects, to our experiment. This detector had an antiblooming system which compromised the linear response at high intensities. This effectively reduced the useful dynamic range to about 14 bits, though this limitation can be circumvented by making averages of many snapshots. The detector chip was covered with a protective glass which had to be removed to eliminate multiple reflections. The fact that the pixels are not square added manageable but unnecessary complications. A larger number of pixels would have also allowed us to reach higher resolutions in the reconstruction.
One difficult aspect of the red-laser-light experiment, in contrast to its x-ray counterpart, is that very few objects are tenuous enough to be weak scatterers in the visible light region. As a consequence, most reconstructed exit-waves are complex-valued. Reconstructions can still be forced to be real-valued if the specimen is a flat mask, such as a photographic slide. Photographic slides can be thin and uniform enough to affect the incident wave only through absorption, which leads to an effectively real-valued exit-wave modulation.

5.2.2 Reconstructions

Both reconstructions presented in this section are those of insect wings. This type of specimen is very well suited for a demonstration as it is partially transparent, has a sharp boundary, and has features easy to recognize. As will be shown below, they also have the property that the phase retardation within the thickest parts is about $2\pi$.

The first wing (from an unidentified insect) was suspended in the vertical position using spider silk. The diffraction pattern, shown in figure 5.12(a), was collected using 3 different exposure times (4 ms, 88 ms and 1 s) to accommodate the large dynamic range. For each exposure time, 25 shots were taken and averaged. The same number of shots was taken for background subtraction. For each exposure time, pixels having a value higher than a predefined threshold (20,000 on a scale going up to 65,536) were rejected because of the antiblooming system mentioned above. All rejected pixels were treated as if not measured: they were set to 0 and their exposure time was also 0. After background subtraction, the different exposures were added together and renormalized according to the pixel-dependent total exposure time. The few pixels that were rejected for all exposures were easily
Figure 5.12: (a) Diffraction pattern of an insect wing (log scale). This diffraction pattern extends to $q_{\text{max}} = \pi (42 \text{ } \mu\text{m})^{-1}$ in the vertical (narrower) direction, giving real-space pixels a dimension of 21 $\mu\text{m}$. The data is a composite of 75 shots taken at 3 different exposure times (4 ms, 88 ms and 1 s). The black central area saturated the detector and was therefore not measured. The very fine nearly horizontal lines are consistent with the orientation of the spider silk that was holding the wing. (b) Autocorrelation. A high pass filter was used to reduce the effect of the missing central data. The gray scale is linear but the data range was truncated to allow low-intensity features to be seen.
identified as those having a final 0 intensity and 0 exposure time. These central missing pixels are shown in black in the center of Fig. 5.12(a). Fortunately, their area is small enough so that no unconstrained mode analysis was needed.

Scattering from the thin silk threads is visible in Figure 5.12(a) as very thin, nearly horizontal streaks in the diffraction pattern. These threads were extended in space and are not oversampled. For this reason, they cannot be reconstructed and will thus act as a systematic error in the diffraction pattern, limiting the compatibility of the data with the support constraint.

A good diagnostic of the quality of the diffraction pattern is obtained by simply looking at the autocorrelation, shown in figure 5.12(b). It appears as three disconnected objects. The central feature is the correlation of the specimen with itself. An image of the insect wing is apparent in the two weaker objects. These “ghosts”, similar to those encountered in the yeast cell reconstruction (fig. 5.2) result from the interference between the light scattered by the wing and a small additional scatterer also present in the field of view. Although this scatterer was not intentional in our case, its presence can be beneficial, as discussed at the end of chapter 2. The scatterer was reconstructed with the wing and can be seen in the top right corner of the reconstruction (Fig. 5.13).

The difference map reconstruction of the wing is shown in figure 5.13. Mis-centering of the diffraction pattern (see section 4.2.1) caused a phase ramp to be present within the wing [fig. 5.13(a)]. No a priori information is available to determine how this phase variation should be corrected. Figure 5.13(b) shows the same reconstruction with a corrected phase ramp. This was adjusted arbitrarily to make the interior of the wing look more uniform. The support used to obtain this reconstruction was obtained from one of the “ghosts” in the autocorrelation,
Figure 5.13: Complex-valued reconstruction of the insect wing from its diffraction pattern (a) before and (b) after removal of a phase ramp. The reconstruction is an average of 50 estimates, taken every 50 iterations. The linear object in the upper right is also part of the reconstruction and is the cause of the “ghosts” in the autocorrelation (Fig. 5.12(b)).
and was refined using a drawing program, very much the same way as for the yeast cell. The reconstruction is the average of 50 Fourier estimates.

We mentioned in chapter 4 that the resolution of a reconstruction depends on many elements, including the contribution of the algorithm. As in the yeast cell, the algorithmic transfer function can give an idea of the resolution of the reconstruction. This curve is given in Figure 5.14(a). For this dataset, an even more useful quantity is the full map of the ratio of reconstructed intensities to the measured intensities [see equation (4.27), page 95]. This fidelity map is shown in Figure 5.14(b). A striking feature of this image is the high contrast between the horizontal region, where intensities are almost completely averaged to zero, and the vertical region with high fidelity up to the edge of the dataset. Comparing with the diffraction pattern, Figure 5.12(a), the cancellation of the lateral intensities can probably be explained by incompatibility between the spider silk scattering and the support constraint.

The goal of the first wing reconstruction was to demonstrate the feasibility of the experiment for teaching purposes. After this, a second setup was made in a research laboratory, with the more precise objective of testing the method for the measurement of wing thickness (the ultimate goal being to get an estimate of the wing’s mass distribution – useful for insect flight simulations for instance). A very similar setup was then reproduced, the only difference being that a smaller pinhole was used. In particular, the same detector and a very similar laser source was used.

This second specimen is the wing of a fruit fly (Drosophila melanogaster). Its preparation was different than the previous wing; it was held on a vertical glass slide (thanks to naturally occurring electrostatic forces) and within a window formed
Figure 5.14: (a) Algorithmic transfer function. The abscissa is labelled with the full period, equal to $2\pi/q$. (b) Fidelity map [cf. (4.27)] for the insect wing reconstruction. The dark regions on the left and the right of the center coincide with the region where noticeable scattering from the spider silk is present in the diffraction pattern, suggesting that this is the main cause of the higher fluctuation of the phases in these regions.
on this slide with adhesive tape. This type of specimen preparation has important advantages. First, the mask is a guarantee that no unwanted alien scatterers are present. Also, the mask, if in the same plane (focus depth) as the specimen, provides a very sharp boundary, known to help convergence of complex-valued reconstructions (see section 4.2.3 and appendix B). Finally, the reconstructed image represents the true wave amplitude, instead of the difference between the incident wave amplitude and the exit wave (by the “Babinet principle” – see the discussion on support projection, page 73). As will be seen below, the main benefit of reconstructing the wave amplitude is the possibility of detecting phase modulations (such as the phase ramp introduced by miscentering).

Figures 5.15(a) and 5.15(b) show the diffraction pattern and the autocorrelation of the *Drosophila* wing. The data collection method was identical as that described above, with 4 instead of 3 different exposure times (2.8 ms, 22.4 ms, 0.25 s and 2 s). As above, the few pixels that saturated the CCD for all exposure times were small enough in number to make unconstrained mode analysis unnecessary.

The *Drosophila* wing reconstruction is shown in figure 5.16(a). It is an average of 50 estimates taken 50 iterations apart. The support was refined using the same method as above. An interesting aspect of this reconstruction, when compared to Fig. 5.13, is that the amplitude of the wavefront itself can be probed using the region surrounding the wing. This allows for the substraction of possible phase aberrations. Figure 5.16(b) shows the wing reconstruction when a phase factor $e^{i\varphi}$, with $\varphi = q_0 \cdot r + Br^2$ is factored out of the original reconstruction. The three free parameters in this phase term correspond to the miscentering of the diffraction pattern ($q_0$) and to a beam divergence or convergence ($B$), which can also be caused by the CCD not being perfectly in the objective’s back focal plane.
Figure 5.15: (a) Diffraction pattern of a *Drosophila* wing. Its aspect ratio makes the resolution in $y$ (76 $\mu$m) significantly lower than in $x$ (54 $\mu$m). The black central area is the missing data region. (b) Autocorrelation. A high-pass filter was used to reduce the effect of the missing central data. The large square shape gives the outline of the autocorrelation from the window. The gray scale is linear but the data range was truncated to allow low-intensity features to be seen. The weak modulations surrounding the autocorrelation support are caused by the sudden truncation of the missing data region.
To generate figure 5.16(b), these parameters were fitted to points surrounding the wing. The origin of the weak amplitude oscillations still visible around the wing is unknown, but could come from interference caused by reflection in the glass slide.

Figure 5.16: Complex-valued reconstruction of the *Drosophila* wing from its diffraction pattern (a) before and (b) after removal of a quadratic phase factor. The reconstruction is an average of 50 estimates, taken every 50 iterations.

The correction of the large-scale phase modulations of the exit wave makes possible a quantitative study of the wing contrast. Figure 5.17 shows the distribution, in the complex plane, of all the pixels in the support. The approximate spiral shape is probably an indication that the optical thickness of the wing can be modeled in the framework of the Eikonal approximation [eq. (2.46), page 18]. The gap between 0 and $\pi/2$ suggests that no portion of the wing has an optical thickness smaller than $\pi/2$. If the average refractive index in the wing is between 1.1 and 1.3, the wing has a minimum thickness of about 0.5 to 1.5 $\mu$m.

It was suggested that measurements such as the *Drosophila* wing could give accurate density maps of insect wings. A probable limitation in the current setting
Figure 5.17: Distribution of the pixel values in the complex plane for the wing reconstruction 5.16. The large cloud at 0 phase is made of pixels of the undisturbed wave (in the window, around the wing). The pixels from the wing are located in the elongated cloud centered around a phase angle of $\pi$. This plot should be compared with the simulated scatter plots of a model yeast cell (Figure 5.22(b), page 170).
is that slightly thicker wings make the phase wrap back to 0, as can already be seen in the thicker parts of the current reconstruction. To reduce phase winding, one could resort to using longer wavelengths, although using light invisible to the eye certainly adds practical complications. Another possibility is matching the index of the wing by immersing it in a transparent liquid (like water). This would require a change of geometry (horizontal specimen and vertical beam), but could give good results.

The method presented here could in principle be applied to shorter wavelength experiments. In particular, placing a specimen in a window could reduce the intensity of the incoming beam enough to remove the need for a beamstop. Probably as a result of specimen preparation difficulties, this approach has not been widely used up to now.

5.3 Quasicrystal reconstruction

We now turn to a very different application of iterative phase retrieval. In this section, we present the reconstruction of the electron density of a quasicrystal from x-ray data.

A complete introduction to the field of quasicrystals is beyond the objectives of this thesis. More than sufficient information can be found, for instance, in Hippert and Gratias (1994). We will restrict the following discussion to the minimum needed to make clear the reconstruction steps. In particular, we will avoid as much as possible the specialized terminology of this field.

A quasicrystal is a structure exhibiting strong Bragg peaks in its diffraction pattern (like a regular crystal) but lacking translational symmetry (unlike a regular crystal). In a quasicrystal, Bragg peaks have positions defined in terms of lattice
vectors $a_i$

$$q = \sum_{i}^{N} n_i a_i,$$  \hspace{1cm} (5.9)

where $n_i$ are integers. The important difference with regular crystals is that the lattice has more than 3 basis vectors; for quasicrystals with icosahedral symmetry, $N = 6$ basis vectors are needed.

Equation (5.9) suggests that, although the crystal is not periodic in three dimensions, reflections can be assumed to lie on a regular lattice in an $N$-dimensional space; the basis $a_i$ is then seen as a projection on a three-dimensional Fourier space of linearly independent lattice vectors in $N$ dimensions. This higher-dimensional description of quasicrystal diffraction data turns out to be extremely convenient since it makes it conceptually identical to regular crystallography. The problem, then, consists in solving the phase problem for a $N$-dimensional unit cell. Since equation (5.9) expresses a projection onto the physical space (often called parallel space), the counterpart in direct space is a cut trough the $N$-dimensional unit cell. The long range order of quasicrystals originates from the periodicity of the higher-dimensional unit cell, and the lack of translational symmetry comes from the incommensurate cut through the cell.\textsuperscript{5} The orthogonal complement of the parallel space is called the perpendicular space. Whereas, in the parallel space the electron density has an atomistic character, the perpendicular space has extended features called atomic surfaces.

In addition to the early Patterson maps, various phase retrieval methods have been used for the structure determination of quasicrystals (see for instance Yamamoto, 1996; Brown et al., 2000; Takakura et al., 2001, 2007). Because of the

\textsuperscript{5}There exist systems very similar to quasicrystals which cut the high-dimensional unit cell with a rational slope, giving so-called “approximants” to the quasicrystal.
extended atomic surfaces, crystallographic direct methods (which rely on atomicity) cannot be used. Instead, most of the methods use in one way or another the fact that the $N$-dimensional unit cell is filled with large voids occupying most of the volume. Thus, a reconstruction has the important constraint that large regions should be nearly 0.

The approach taken here is consistent with the general formalism of the difference map. One constraint is provided by the Fourier data and the other is that a predetermined volume should have 0 electron density. Projection onto the former is the usual modulus projection (see section 3.4.1) and the projection on the latter is the dynamic support projection (projection 10, page 75).

We will first give a description of the reconstruction procedure. We then show our results, which are in agreement with results obtained by other methods. We also show how the reconstruction method presented here allows for an algorithmic probing of a long-standing question about the centrosymmetry of icosahedral quasicrystals (de Boissieu et al., 1994; Lee et al., 1996).

5.3.1 Reconstruction details

We present the reconstruction of an icosahedral AlMnPd quasicrystal electron density, from x-ray data measured by Boudard et al. (1992) and kindly provided by one of the authors (M. de Boissieu). The crystal has a body-centered cubic bravais lattice in direct space (and thus a face-centered cubic lattice in reciprocal space).

Implementation of the projections described above on a computer is straightforward in principle. In practice, it is nevertheless made difficult by storage requirements and by the book-keeping required to satisfy the symmetry operations.
The first issue is especially important because of the high-dimensionality of the problem. Holding all the experimental data in its original form would require a six-dimensional cubic grid having 21 points in each direction, taking 686 MB of memory for single-precision complex numbers. An implementation of the difference map needs at least 4 independent arrays of this size. At the time of writing this thesis, this type of memory requirement is at the limit of what can be done on a non-parallel computing systems. This problem can be circumvented by first noting that reflections are indexed in a cubic cell system, which means that half of the indices are 0 because of the bcc lattice extinction rules. Transforming the cell to a native indexing is therefore the first obvious step to do to contain the data in a more compact way. The problem then becomes one of finding the appropriate basis vectors that hold the measured reflections on the most compact cell. In a mathematical language, this is the problem of finding the optimal unimodular transformation for this lattice. Similar lattice reduction problems have been thoroughly studied; in the present case the transformation could be obtained by inspection.

The optimal reduced lattice was found to hold all the reflections on a $13^6$ array. The reconstruction was then chosen to take place in a unit cell sampled on a $12^6$ grid. The original data contains 600 independent reflections. The icosahedral group has 120 elements (including reflections through the origin), resulting in some of these reflections having up to 120 identical copies in reciprocal space. These copies are “orbits” of the symmetry group, and the number of elements in a given orbit is called the multiplicity. Upon application of the symmetry operations

---

\( ^6 \)As we will show below, reflection through the origin (centrosymmetry) is not necessarily a symmetry of the crystal, although it is a symmetry of the reflection data because the density is real (Friedel symmetry).
of this group, 36 reflections had reciprocal-space vectors lying outside the selected 12\(^6\) cell. These reflections were eliminated from the reconstruction, leaving 564 independent reflections used in the reconstruction.

After symmetry operations, only 45874 of the intensities at the 12\(^6\) lattice points are known, that is, less than 1.6%. Leaving untouch all the other Fourier amplitudes in the reconstruction is obviously not possible since we need a sufficiently overconstrained system. One constraint that can be applied to these unmeasured Fourier amplitudes is that they also satisfy the symmetry requirement. We found that 2296 additional orbits were completely contained within the reciprocal cell. Since most of the scattered power is contained in the measured reflections, we simply set to 0 the amplitudes at all lattice points belonging to incomplete orbits. The 2296 complete orbits were not set to 0, but a total power bound was applied to them to prevent unphysically high values. Two unmeasured orbits were however treated differently: one is the origin and thus contains only one element; the other was a missing measurement found to lie close to the origin, which means that, although not observed, it could be a strong reflection. These two amplitudes were allowed to float freely.

To our knowledge, all quasicrystal reconstructions in the literature use the assumption that the structure is centrosymmetric, forcing all the phases to be 0 or \(\pi\) (in the standard centering). We have decided to not use this assumption; all orbits having a multiplicity of 120 were thus allowed to have a complex phase. Because of icosahedral symmetry all other orbits are forced to have a real phase factor (these are called the “centrosymmetric orbits” below).

Each Fourier projection would then consist in applying the following steps

1. Compute the 6D FFT of the iterate
2. Set to 0 the Fourier amplitudes having incomplete orbits

3. Reset the amplitude of all measured centrosymmetric orbits, preserving the sign of the real part and discarding the imaginary part.

4. Compute the average of the complex amplitudes (taking care of the conjugation of Friedel pairs) of measured non-centrosymmetric orbits. Keep this phase and restore the measured amplitude.

5. Replace also all non-measured complete orbits with the amplitude averaged over the orbit elements, discarding the imaginary part for centrosymmetric orbits. Then, rescale all the amplitudes if their total power is higher than a given value.\(^7\)

6. Inverse Fourier transform.

Comparatively, the direct-space projection is much simpler. Its formulation is exactly as stated on page 75 (the “area-based dynamic support projection”): in short, it involves sorting all voxel values in the unit cell and setting the \(N_d\) lowest to 0.

Overall, the reconstruction requires two adjustable parameters: the power bound on the unmeasured orbits, and the volume fraction occupied by the atoms. We have explored various values of these parameters and found little variation in the reconstructions as long as these were within a reasonable range. For the following reconstructions, the power bound is 8\% of the total power, and the volume fraction of the non-zero region in the unit cell is 25\%.

The reconstruction shown here is the average of 5 individual reconstructions

\(^7\)This rescaling can be shown to be a distance minimizing operation.
Figure 5.18: Two-dimensional sections of the reconstructed atomic surfaces of the AlMnPd quasicrystal. The axes are in units of the six-dimensional lattice constant. The four atomic surfaces are centered on the four inequivalent centers of icosahedral symmetry. Facet-like features can be seen, as expected. Note the low amplitude of the scale in (b); only three sites have a high occupancy.
Figure 5.19: Reconstructed electron density in a parallel-space plane having icosahedral mirror symmetry. The units are in angstrom.

from independent runs, each being the average of 475 Fourier estimates taken 20 iterations apart. The combined average is thus made of 2375 Fourier estimates.

Figure 5.18 shows two-dimensional sections of the reconstructed atomic surfaces. The chosen plane is the plane in perpendicular space having icosahedral mirror symmetry. These reconstructed atomic surfaces are very similar to those reconstructed from a quasicrystal of different composition by Takakura et al. (2001).

A two-dimensional slice of the electron density in parallel (physical) space is shown on Figure 5.19. This figure agrees very well with a previous reconstruction of the same dataset (Brown et al., 2000).

Because Bragg reflections come from a large monocrystal, the reconstructed electron density is that of the average structure of the quasicrystal. As is always the case in crystallography, phonons produce slight atomic displacements that result in
blurred electron densities, whose effect is characterized as a resolution-dependent attenuation of the reflections. The attenuation factor is the Debye-Waller factor, well known to crystallographers. For quasicrystals, a similar attenuation is also present in the perpendicular direction — the associated long-wavelength perturbations analogous to phonons are then called phasons. Phasons are responsible for the soft edges of the atomic surfaces shown on Figure 5.18. In parallel space, they manifest themselves as low-occupation sites that have unphysical distances with neighboring atoms.

5.3.2 Algorithmic test of the centrosymmetry assumption

Whether quasicrystals deviate from centrosymmetry is still a question subject to debate. Three-beam measurements (from which one can deduce phase triplet information), made by Lee et al. (1996) seem to indicate that icosahedral quasicrystals are non-centrosymmetric. However, measurements by de Boissieu et al. (1994) using anomalous scattering close to the Pd K edge did not show any sign of departure from centrosymmetry. While the authors of the latter mention twining as a possible explanation for their negative result, three-beam measurements can be very sensitive to multiple scattering effects, and other experimental techniques should be used to confirm these results.

We looked for signs of non-centrosymmetry by monitoring the phases of the brightest reflections not forced to have centrosymmetric phases. Figure 5.20 shows the histogram distribution of these phases. Without imposing any constraint on these phases, they spontaneously oscillate very close to 0, and our results do not show statistically significant signs of bias away from 0.
Figure 5.20: Histogram distribution of the phase angles of the 6 brightest reflections not forced to be real by symmetry. The intensity given for each reflection is to be compared with the brightest measured reflection, $I = 295.9$. No significant deviation from 0 (or $\pi$) is observed.
5.4 Simulations

5.4.1 Propagation through a yeast cell model

We present here propagation simulations through a model cell. This simulation was used to arrive at a better understanding of the contrast mechanism in the yeast cell experiment (section 5.1). The freeze dried yeast cell model is a 3 µm sphere made of lipids and proteins. All the lipids are concentrated in the 50 nm thick cell membrane. The protein material inside the cell is modeled as a binary-valued distribution occupying 25% of the cell volume. The exact spatial distribution of the protein material is random, but the envelope of its power spectrum was forced to follow a power law of the form $|F(q)| \propto q^{-\alpha}$. The exponent of this power law controls the relative contributions of high and low spatial frequencies. Figure 5.21 shows the protein material in one slice of the cell when generated with our choice of exponent, $\alpha = 1.6$. For the simulations, the refractive index of the material inside the cell was sampled on a grid with $(25 \text{ nm})^3$ voxels.

The refractive indices $n = 1 - \delta - i\beta$, of a model lipid (62.3% H, 31.4% C and 6.3% O as number fraction) and protein (48.6% H, 32.9% C, 8.9% N, 8.9% O and 0.6% S) were calculated using the data of Henke et al. (1993). These are given in table 5.1 for the 750 eV x-ray energy of the yeast-cell experiment.

<table>
<thead>
<tr>
<th></th>
<th>$\delta$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>lipid</td>
<td>$4.23 \times 10^{-4}$</td>
<td>$6.88 \times 10^{-5}$</td>
</tr>
<tr>
<td>protein</td>
<td>$5.44 \times 10^{-4}$</td>
<td>$1.05 \times 10^{-4}$</td>
</tr>
</tbody>
</table>
From table 5.1 we see that the wave acquires a phase shift of $\Delta \varphi = 2\pi t\delta/\lambda$ or about $\pi/2$ as it passes through the center of a 3 $\mu$m diameter cell with a quarter of its volume filled with protein so that $t = 0.75$ $\mu$m. The specimen is thus optically thick.

Figure 5.22(a) shows a typical exit wave obtained with the multislice method of section 2.2.4. The exit wave was propagated back to the center of the cell, as shown in Figure 5.24 below. The plane going through the center of the cell is the plane selected by the support constraint and is thus the plane of the reconstruction. To the eye, the difference between the wave at the central plane and the wave at the exit of the cell (not shown) is unnoticeable. Figure 5.22(b) shows the distribution of the exit wave values (in the complex plane), back propagated to the central plane. This distribution is itself useful in that it might be used to constrain the reconstruction of the object/exit-wave. That these values lie on an approximate
Figure 5.22: (a) Exit wave for the model yeast cell (750 eV x-rays), back-propagated to the center of the cell. The wave at the exit of the cell (not shown) is almost identical to the eye. (b) Complex plane distribution of the pixel values in the exit-wave shown in (a). The spiral shape is a sign that the cell is optically thick, and broadening of the spiral is explained by Ewald sphere lift-off effects.

The spiral is explained by the Eikonal approximation [equation (2.45), page 18]. The distribution of exit wave values $\Psi$ is simply related to the distribution, transverse to the beam, of projected protein thicknesses $t$ by

$$\Psi \propto \exp(ik t \delta n).$$

(5.10)

Broadening of the spiral is due to Ewald sphere lift-off effects, that is, terms of higher order in the transverse spatial frequency, $q/k$.

Figure 5.23 illustrates the simulated perturbation of the incident wave as it propagates through the cell. At the scale of this image, the lateral dispersion of the wavefield is unnoticeable. Figure 5.24(a) shows the wavefield propagated (in free space) both forward and backward from the exit plane, with the axis in the direction of propagation compressed by a factor 5. As emphasized in figure 5.24(b), there is a noticeable spread of the field perturbation. The “waist” is located at
Figure 5.23: Longitudinal section of the simulated wave perturbation through the model cell (outlined in white).

Figure 5.24: (a) Longitudinal section of the exit-wave of the model cell propagated forward and backward. (b) Same as (a) but with an intensity threshold to emphasize the weak perturbations of the wavefield. The propagation axis has been compressed by a factor 5.
the center of the cell, the plane of the reconstruction. This figure helps explain
the blurring effect caused by the averaging of many estimates in a reconstruction.
In particular, a loose support allows a wider range of focus planes and results in a
less well-defined reconstruction (see appendix B).

5.4.2 Reconstructions involving many datasets

Section 2.4.2 contains a list of many recent methods which have in common the
collection of many datasets. The increase in redundancy makes easier the recon-
structions and removes the requirement that the specimen be isolated. All but one
of these methods (Rodenburg et al., 2007a) use the same algorithm to extract in-
formation from the multiple datasets: a many-projection version of the alternating
scheme, of the form

\[ x_{n+1} = P_N \circ \cdots \circ P_2 \circ P_1(x_n) \]  

(5.11)

In the reported work, this scheme seems to be sufficient and successful. Never-
theless, it is known that such a combination of non-convex projections can stagnate
far from the solution. Also, there is no simple metric to monitor the evolution of
this iteration scheme. It seems, therefore, that using the difference map formalism
to tackle these problems would provide a greater insight and help clear doubts on
reproducibility.

As mentioned in chapter 3, the difference map is formulated to find the intersec-
tion of only two constraint sets. We have also mentioned that adding constraints
is possible through the increase of the search space. In what follows, we will show
how various problems involving many datasets can be posed to make possible the
use of the difference map. The “replica” method introduced in section 3.2.1 will
serve as an inspiration, although we will make adaptations for each problem.
A close look at each of the multiple-measurement methods indicates that the mathematical structure of the problem to solve is always the same. We will first express the problem in an abstract, “neutral”, notation. It will then suffice to identify each symbol for the particular applications discussed later. The problem, in general, is that of finding a two-dimensional complex function \( x(u) \) subject to \( n \) constraints of the form

\[
a_i(v) = \left| \mathcal{F}[y_i(u)x(u)] \right|^2, \tag{5.12}
\]

where \( u \) and \( v \) are Fourier conjugate coordinates and the functions \( y_i(u) \) are known complex-valued functions.

The approach taken by most authors is as follows. Starting with an estimate of \( x \), first compute \( \mathcal{F}[y_1x] \) and apply the constraint in Fourier space (with the usual modulus projection). Inverse Fourier transform the result and divide by \( y_1 \) to obtain a new estimate of \( x \). Repeat the procedure for all the measurements and repeat this cycle until some convergence criterion is reached. This procedure is obviously of the type (5.11).

The difference map can solve this type of problem by enlarging the search space in a specific way. Let the new iterate be the direct product of \( n \) two-dimensional complex-valued functions \( z_i(u) \) (in practice always sampled on a grid):

\[
z = z_1 \otimes z_2 \otimes \cdots \otimes z_n. \tag{5.13}
\]

With these auxiliary variables, the problem (5.12) can now be decoupled into two constraints. The constraint from the measured data becomes

\[
a_i(v) = |\tilde{z}_i(v)|^2, \tag{5.14}
\]

where \( \tilde{z}_i = \mathcal{F}z_i \) as usual. The second constraint enforces the fact that all \( z_i \) are
coupled through a single function $x(u)$:

$$z_i(u) = y_i(u)x(u). \quad (5.15)$$

The projection onto the first constraint set is trivial, as all datasets are decoupled:

$$P_D(z) = P_1(z_1) \otimes P_2(z_2) \otimes \cdots \otimes P_n(z_n), \quad (5.16)$$

where the projections $P_i$ are projections to the modulus constraints (5.14) (see section 3.4.1). $P_D$ is the “direct product” projection.

The projection onto the constraint expressed by the $n$ equations (5.15) can be found by finding a new set of $z'_i(u)$ that minimize the distance

$$d^2 = \sum_{i=1}^{n} \sum_{u} |z_i(u) - z'_i(u)|^2 \quad (5.17)$$

while also satisfying (5.15). Lagrange multipliers are not needed here, as the problem reduces to minimizing

$$d^2 = \sum_{i=1}^{n} \sum_{u} |y_i(u)x(u) - z_i(u)|^2 \quad (5.18)$$

with respect to $x$. The solution is

$$x(u) = \frac{\sum_{i=1}^{n} y_i^*(u)z_i(u)}{\sum_{i=1}^{n} |y_i(u)|^2}. \quad (5.19)$$

and the projection is therefore given by

$$P_A(z) = y_1x \otimes y_2x \otimes \cdots \otimes y_nx, \quad (5.20)$$

with $x$ given by (5.19).\footnote{\textit{PA} stands for “average”, in analogy with the method described in section 3.2.1.}

Both projections are well-defined and easy to compute. In terms of practical implementation, the main drawback of this method is its memory requirement: a
large collection of measurements means a large iterate as well. On the other hand, the form of \( P_D \) makes it simple to implement for parallel computing. The difference map approach can be beneficial since tools such as averaging can be used in the usual way, and thus reproducibility issues can be addressed on firmer grounds.

In the following, we will cover three applications of the above framework. We will first consider \textit{aperture scanning methods}, which require the use of masks (for oversampling) moved upstream or downstream from the specimen. The second case considered is the classical \textit{focal series} reconstruction, useful in electron microscopy for instance, where it is possible to form highly magnified direct space images with a good control of defocus. A third method, \textit{astigmatic diffraction}, which uses phase perturbations in the incident wave, will be seen to have a formulation identical to the aperture scanning one.

We will also give a brief account on possible applications in the field of crystallography, where multiple dataset collection is very common. This section will be concluded by a discussion of future research directions in the algorithm development for multiple dataset measurements.

\textbf{Aperture scanning methods}

In the last few years, two versions of aperture scanning methods have been proposed by the same group of researchers. The first proposal (hereinafter the “downstream method”) involves the measurement of diffraction patterns from radiation coming out of a mask downstream from a specimen.\textit{(Faulkner and Rodenburg, 2004)} The other method (the “upstream method”) simply reverses the geometry and places the mask upstream \textit{(Rodenburg et al., 2007a,b)}. In addition to allowing any extended specimen, both cases provide the benefit of reducing the intensity of the
central beam, possibly enough to remove the necessity of a beamstop. Considering radiation damage, the upstream method is certainly more advantageous as all the radiation incident upon the specimen generates useful information.

For a given position $r_i$ of a mask, a diffraction pattern in the downstream method is given by

$$I(q) = \left| \mathcal{F}[M(r - r_i)\Psi(r)] \right|^2,$$

(5.21)

where $M(r)$ is the function of the mask – 0 outside a given support and typically 1 inside – and $\Psi(r)$ is the (forward-propagated) exit-wave from the specimen. As long as the mask can be considered as an optically thin object inside the support, this expression is valid for all scattering regimes.

In the upstream case, the wave incident on the specimen is the forward-propagated exit wave of the mask $M_z(r)$ (the “illumination profile”). The wave at the exit of the specimen depends on its scattering properties. Assuming small scattering angles, the incident illumination profile can be factored out of the exit wave and the diffraction pattern is given by

$$I(q) = \left| \mathcal{F}[M_z(r - r_i)\Psi(r)] \right|^2,$$

(5.22)

where now $\Psi(r)$ is given by the Eikonal approximation [eq. (2.45), page 18] or its approximation linear in the scattering strength (the projection approximation).

Equations (5.21) and (5.22) have the structure of equation (5.12). In the current notation, the first projection is

$$P_D(\psi) = P_1(\psi_1) \otimes P_2(\psi_2) \otimes \cdots \otimes P_n(\psi_n),$$

(5.23)

where $\rho$ takes the role of $z$ in the above discussion, and the projections $P_i$ are modulus projections [projection 1, page 66] for the individual datasets $I_i$. 
The second projection is:

\[ P_A(\psi) = \left( M(r - r_1)\Psi_A(r) \right) \otimes \cdots \otimes \left( M(r - r_n)\Psi_A(r) \right), \]  

(5.24)

with

\[ \Psi_A(r) = \frac{\sum_{i=1}^{n} M^*(r - r_i)\psi_i(r)}{\sum_{i=1}^{n} |M(r - r_i)|^2}. \]  

(5.25)

Following the example of Rodenburg et al. (2007a), it is adviseable to add to the denominator of this expression a small positive number \( \delta \) for numerical stability.

We now give a simple demonstration of the upstream method for only two aper-

---

Figure 5.25: Simulation of an aperture scanning experiment (upstream version). A circular aperture is moved to two positions in front of the specimen (a), resulting in the diffraction patterns (b) and (c). A low noise level has been added to the diffraction patterns.
Table 5.2: Physical parameters in the aperture scanning simulation (Figures 5.25, 5.26) assuming red laser illumination (632.8 nm).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mask diameter</td>
<td>0.75 mm</td>
</tr>
<tr>
<td>Mask to specimen distance</td>
<td>2 mm</td>
</tr>
<tr>
<td>Specimen to detector distance</td>
<td>10 cm</td>
</tr>
<tr>
<td>Detector pixel size</td>
<td>20 μm</td>
</tr>
<tr>
<td>Image resolution</td>
<td>12 μm</td>
</tr>
</tbody>
</table>

The image reconstructed using the projections described above is shown on Figure 5.26. The high redundancy in the two datasets makes it a computationally easy problem and convergence is almost immediate.

---

9We suggest, in passing, that this approach could improve dramatically the reconstruction shown in Fig. 3 of the paper by Rodenburg et al. (2007a)
Figure 5.26: Reconstruction from the diffraction patterns shown on Figures 5.25(b)–(c).
Focal series reconstruction

We now consider the case of phase retrieval from a series of intensity measurements in the near field (or in the conjugate plane corresponding to near-field).

As we have seen in chapter 2 (in page 14), propagation by a distance $z$ simply amounts to a multiplication by a phase factor in Fourier space. Thus, intensity measurements of defocussed images are given by

$$I_i = \left| \mathcal{F}^{-1} [ \widetilde{P}_z(q) \bar{\Psi}(q) ] \right|^2,$$

where $\widetilde{P}_z(q) = \exp(-\frac{i}{2} q^2 / k)$ in the paraxial approximation. Again, equation (5.26) has the same form as (5.12). $P_D$ is as above (5.23) and, because $P_z$ is a pure phase object, $P_A$ takes a simpler form, this time very close to an average:

$$P_A(\tilde{\psi}) = P_{z_1} \bar{\Psi}_A \otimes P_{z_2} \bar{\Psi}_A \otimes \cdots \otimes P_{z_n} \bar{\Psi}_A,$$

with

$$\Psi_A = \frac{1}{n} \sum_{i=1}^{n} \tilde{P}_{-z_i} \tilde{\psi}_i.$$

A simple simulation of a focal series reconstruction, using this formalism, is shown below. Figure 5.27 is a simulated focal series of a pure phase specimen. Noise was added to the datasets to demonstrate that the method is robust.

A reconstruction using all the datasets in Figure 5.27 is shown on Figure 5.28(a). For comparison, one estimate, taken of the 50 used in the average, is shown in Figure 5.28(b). The average, reproducible and of a higher quality, should certainly be preferred to a single estimate as the final reconstruction.

The above demonstration should be seen more as a proof of principle than as an actual demonstration of a fully developed technique. In particular, the specimen was conveniently assumed to be isolated and surrounded by a flat region.
Figure 5.27: Simulated focal series measurements of a pure phase object. In the specimen, the maximum phase shift is $\pi/2$. Poisson noise (such that relative error is 5% on the brightest pixel) has been added. The focal distances correspond to radiation of wavelength 632.8 nm; detector pixels have dimensions 10 $\mu$m $\times$ 10 $\mu$m.

Although this situation is possible, the requirement that a specimen is isolated is always limiting, and multi-measurement methods have been developed, certainly in part, to overcome this limitation. A more realistic reconstruction involves the measurement of an extended object. Fortunately, the technique developed in this section can be adapted to this situation, by relaxing appropriately the constraints.

**Astigmatic diffraction**

*Astigmatic diffraction* methods were introduced by a group based in Melbourne (Nugent et al., 2003, 2005; McBride et al., 2005). In these experiments, one measures multiple far-field diffraction patterns from an isolated specimen (or a selectively illuminated region of a specimen), with controlled variations of the phase profile of the incident illumination. Assuming that the incident wavefield $\Psi_0^{(i)}$ can
be factored out of the exit wave, the measured intensities are

\[ I_i(q) = |\mathcal{F} \left[ \Psi_0^{(i)}(r) \rho(r) \right] |^2, \tag{5.29} \]

where \( \rho \) is the wave modulation. This type of constraint is not only of the form (5.12), but is also very close conceptually to the upstream version of the aperture scanning methods. As such, the same kind of analysis results, and a simulation example will not be provided (we have verified that it indeed works as expected).

**Discussion**

The examples shown above are not the only instances of experimental techniques based on the measurement of multiple datasets. In crystallography, almost all phasing methods are based on multiple measurements. Here is a quick overview of the most common methods.

**Isomorphous replacement** (Green et al., 1954; Blow and Rossmann, 1961) is a technique where diffraction data is taken from a native crystal and from a crystal in which heavy atoms have been added (either by soaking the crystal in a heavy atom solution or by a replacement of atoms within the pro-
tein structure itself – typically Sulfur atoms replaced with Selenium). The term “isomorphous” expresses the requirement that this heavy atom addition does not change the structure of the protein and the crystal parameters. “Multiple isomorphous replacement” (MIR) is also often used to add further constraints.

**Anomalous scattering diffraction** techniques, either in “single” (SAD) or in “multiple” (MAD) form, involve the measurement of diffraction data at a wavelength close to an absorption edge of a heavy scatterer (see for instance Hendrickson, 1991). Absorption has the effect of making the scattering factor effectively complex-valued, which lead to the disappearance of the Friedel symmetry in the measurements. Anomalous scattering can be combined with the isomorphous replacement method; this is the “single isomorphous replacement with anomalous scattering” (SIRAS).

**Radiation induced phasing** (RIP) is a more recent method where multiple diffraction datasets are taken from a specimen having endured various x-ray doses (Ravelli et al., 2003). The technique relies on the fact that, for a moderate dose, radiation damage has only localized effects (disulfide bonds and carboxyle groups are especially affected, see Weik et al., 2000).

All these methods have in common that a specific combination of datasets gives a density of isolated peaks (from the heavy or anomalous scatterers, or the localized damages) on which the atomicity constraint can be used successfully (see projection 11 on page 77). This type of information, combined with the diffraction measurements, could be sufficient to obtain the structure of a crystal. In practice, information contained in crystallographic diffraction data is much scarcer than
in oversampled datasets, which makes reconstructions generally harder. A more robust reconstruction method should include even more constraints, as those based on chemistry or on the knowledge of the protein sequence.

An approach similar to the multiple-dataset technique could also be necessary for the reconstruction of three-dimensional optically thick specimens. Outside the region of validity of the Born approximation, two-dimensional diffraction patterns cannot be assembled into a three-dimensional dataset [as described by equation (2.66), page 21]. If other analytic expressions (from the Eikonal or Rytov approximations) are still valid, the three-dimensional density of the specimen might still be reconstructed from the diffraction data. As before, the iterate would be formed by a direct product of two-dimensional arrays, this time corresponding to each orientation of the specimen. One of the projection would be the “direct product” projection, as defined above, where each individual projection is the usual modulus projection. The second projection would combine the individual views using tomographic methods, and apply the three-dimensional support constraint on the generated density. The problem could also be approached differently with a three-dimensional array as the iterate, and use a “tomographic projection” that enforces the constraint from the measured diffraction patterns. In both cases, the relatively complicated structure of the Rytov approximation equation make unlikely the definition of an exact projection.
Chapter 6

Conclusion

Many diffractive imaging techniques have now shown their potential as useful probes for scientists. Yet, many obstacles still need to be overcome. On the experimental side, the requirements are often stringent and numerous. We have enumerated a few that have direct consequences on the quality of measured data (high coherence, low noise, complete dataset), but many other technical aspects have not been addressed. Like other developing experimental techniques, diffractive imaging relies on the very most recent technological advances.

The object of this thesis was to give an account of the recent progress of the algorithmic component of diffractive imaging techniques, especially of diffraction microscopy. We have described the theoretical background of the physical processes involved in diffraction experiments, and of the framework that we judge most appropriate to describe iterative algorithms. The formalism, defined in terms of constraint sets and projections onto them, allow for accurate interpretation of the search dynamics and convergence characteristics involved in iterative algorithms. Still, many fundamental questions remain unexplored or unanswered. Diagnostic tools that can help assess the quality or reconstructibility of a dataset prior to the reconstruction are still rudimentary. Analysis of the autocorrelation is already a well known way of observing the sharpness of the support. As we have seen, it also allows an evaluation of the transverse coherence length. Other diagnostics could be based on expected speckle statistics, Friedel symmetry violation, or background level evaluation, to name a few. Reconstruction methods could be improved on many fronts. Of course, new problems will impede the definition of new constraint
sets and projections; alternatives to averaging for a reproducible solution could be found to be more appropriate; more information on the search dynamics could be obtained from statistics accumulated as the iteration progresses; finally, an algorithm optimized to search for the intersection between more than two constraints could broaden substantially the range of applications of the difference map.

Among future applications of diffraction microscopy, one of the most ambitious is the reconstruction of single macromolecules by measurement and assembly of their diffraction patterns from the very short pulses provided by the x-ray free electron lasers currently being constructed (Neutze et al., 2000). This project involves measuring, sorting and assembling thousands of noisy diffraction patterns from identical copies of macromolecules. The intense x-ray pulse will destroy the molecule in a few femtoseconds, but simulations show that short enough pulses can give usable diffraction data. Reconstruction of a larger specimen using such a short pulse has already been reported by Chapman et al. (2006a).
Appendix A

Retrieved phases for the AlMnPd quasicrystal x-ray dataset

Table A.1: Retrieved phases for the AlMnPd quasicrystal from x-ray data by Boudard et al. (1992). Indexing convention is from Cahn et al. (1986). Included in the table are the length of the reciprocal space vector in parallel ($Q_\parallel$) and perpendicular ($Q_\perp$) spaces, measured intensities ($I$), and the phases retrieved by Brown et al. (2000) with the same dataset ($\varphi'$). Reflections that can be complex-values are marked with a *. The first entry in this table is the retrieved origin (equal to the total charge in the unit cell).

<table>
<thead>
<tr>
<th>$h$</th>
<th>$h'$</th>
<th>$k$</th>
<th>$k'$</th>
<th>$l$</th>
<th>$l'$</th>
<th>$Q_\parallel$</th>
<th>$Q_\perp$</th>
<th>$I$</th>
<th>$\varphi$</th>
<th>$\varphi'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>117.0150</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0.44</td>
<td>1.14</td>
<td>0.005</td>
<td>0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0.64</td>
<td>1.04</td>
<td>0.0390</td>
<td>$\pi$</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.71</td>
<td>0.71</td>
<td>3.1379</td>
<td>$\pi$</td>
<td>$\pi$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.74</td>
<td>1.20</td>
<td>0.0019</td>
<td>0</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1.03</td>
<td>1.39</td>
<td>0.0000</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.04</td>
<td>0.64</td>
<td>0.4665</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1.14</td>
<td>0.44</td>
<td>0.6896</td>
<td>$\pi$</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.20</td>
<td>0.74</td>
<td>0.1256</td>
<td>$\pi$</td>
<td>-</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1.28</td>
<td>1.36</td>
<td>0.0000</td>
<td>$\pi$</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1.36</td>
<td>1.28</td>
<td>0.0000</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.40</td>
<td>1.03</td>
<td>0.0308</td>
<td>$\pi$</td>
<td>-</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1.41</td>
<td>1.41</td>
<td>0.3409</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>1.41</td>
<td>1.41</td>
<td>0.0269</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>1.59</td>
<td>0.98</td>
<td>0.0259</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1.66</td>
<td>0.86</td>
<td>0.6418</td>
<td>$\pi$</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1.69</td>
<td>0.40</td>
<td>9.3907</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1.70</td>
<td>1.05</td>
<td>0.0557</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

| $h$ | $h'$ | $k$ | $k'$ | $l$ | $l'$ | $Q_{||}$ | $Q_{\perp}$ | $I$ | $\varphi$ | $\varphi'$ |
|-----|-----|-----|-----|-----|-----|------|-------|-----|------|-------|
| 0   | 3   | -1  | 1   | 0   | 0   | 1.82 | 1.48  | 0.0014 | 0   | -     |
| 3   | 1   | 0   | 0   | 0   | 1   | 1.82 | 1.48  | 0.0000 | $\pi$ | -     |
| 0   | 3   | 0   | 0   | 1   | 0   | 1.84 | 1.27  | 0.0499 | $\pi$ | -     |
| 1   | 2   | 0   | 0   | 1   | 1   | 1.85 | 0.27  | 80.1753| 0    | 0     |
| 2   | 2   | 0   | 0   | 0   | 0   | 1.95 | 0.46  | 10.6861| 0    | 0     |
| 2   | 2   | -1  | 1   | 1   | 0   | 2.00 | 1.23  | 0.0000 | $\pi$ | -     |
| 2   | 2   | 0   | 1   | -1  | 1   | 2.05 | 1.14  | 0.1022 | 0    | -     |
| 2   | 2   | 1   | 0   | 0   | 1   | 2.07 | 0.84  | 0.5165 | $\pi$ | 0     |
| 2   | 2   | 0   | 0   | 0   | 0   | 2.08 | 1.29  | 0.0503 | $\pi$ | -     |
| 2   | 2   | 0   | 0   | 2   | 0   | 2.08 | 1.29  | 0.0000 | $\pi$ | -     |
| 1   | 3   | -1  | 1   | -1  | 1   | 2.20 | 1.47  | 0.0000 | 0    | -     |
| 1   | 3   | 1   | 0   | 0   | 0   | 2.21 | 0.79  | 0.5492 | $\pi$ | 0     |
| 1   | 3   | 0   | 0   | 0   | 1   | 2.26 | 0.63  | 5.3459 | $\pi$ | 0     |
| 2   | 2   | 0   | 0   | 0   | 2   | 2.29 | 0.87  | 0.0335 | $\pi$ | -     |
| 1   | 3   | 0   | 1   | 1   | 0   | 2.29 | 0.87  | 0.0345 | 0    | -     |
| 1   | 3   | 1   | 0   | 2   | 0   | 2.33 | 1.44  | 0.0000 | $\pi$ | -     |
| 3   | 2   | 0   | 0   | -1  | 1   | 2.33 | 1.44  | 0.0244 | $\pi$ | -     |
| 1   | 3   | -1  | 2   | 0   | 0   | 2.33 | 1.44  | 0.1815 | 0    | -     |
| 3   | 2   | 1   | 0   | 1   | 0   | 2.38 | 1.36  | 0.0086 | 0    | -     |
| 1   | 3   | -1  | 1   | 1   | 1   | 2.40 | 1.12  | 0.1106 | 0    | -     |
| 3   | 2   | 0   | 1   | 0   | 0   | 2.40 | 1.12  | 0.0326 | $\pi$ | -     |
| 0   | 4   | 0   | 0   | 0   | 0   | 2.41 | 1.49  | 0.0847 | 0    | -     |
| 3   | 2   | -1  | 1   | 0   | 1   | 2.41 | 1.49  | 0.0000 | $\pi$ | -     |
| 3   | 2   | 0   | 0   | 0   | 1   | 2.51 | 1.09  | 0.1607 | 0    | 0     |
| 1   | 3   | 0   | 0   | 2   | 1   | 2.56 | 0.98  | 0.0952 | $\pi$ | -     |
| 2   | 3   | -1  | 1   | 0   | 0   | 2.56 | 0.98  | 0.0411 | $\pi$ | -     |
| 2   | 3   | 0   | 0   | 1   | 0   | 2.58 | 0.61  | 9.6018 | $\pi$ | 0     |
| 2   | 3   | 1   | 0   | -1  | 1   | 2.58 | 1.15  | 0.0000 | $\pi$ | -     |
| 2   | 3   | 0   | 0   | -1  | 2   | 2.68 | 1.35  | 0.0116 | 0    | -     |
| 2   | 3   | 2   | 0   | 1   | 0   | 2.68 | 1.35  | 0.0315 | $\pi$ | -     |
| 2   | 3   | 1   | 1   | 0   | 1   | 2.69 | 0.53  | 1.4646 | 0    | 0     |
| 2   | 3   | 1   | 1   | 0   | 0   | 2.73 | 0.25  | 12.6164| 0    | 0     |
| 2   | 3   | 0   | 1   | 0   | 1   | 2.75 | 0.65  | 1.2552 | $\pi$ | $\pi$ |
| 1   | 4   | 0   | 0   | -1  | 1   | 2.79 | 1.32  | 0.0000 | $\pi$ | -     |
| 1   | 4   | 1   | 0   | 1   | 0   | 2.83 | 1.23  | 0.1031 | 0    | 0     |
| 2   | 3   | -1  | 1   | 0   | 0   | 2.83 | 1.23  | 0.0711 | 0    | -     |
| 1   | 4   | 1   | 0   | 0   | 0   | 2.84 | 0.96  | 0.2525 | $\pi$ | 0     |
| 1   | 4   | -1  | 1   | 0   | 0   | 2.85 | 1.37  | 0.0091 | 0    | -     |
| 3   | 3   | 1   | 0   | 0   | 0   | 2.94 | 0.92  | 0.1006 | 0    | -     |
| 1   | 4   | 0   | 0   | 1   | 1   | 2.94 | 0.92  | 0.0572 | $\pi$ | -     |
| 3   | 3   | 0   | 0   | 0   | 1   | 2.98 | 0.78  | 3.8666 | $\pi$ | 0     |
| 2   | 3   | 0   | 0   | 1   | 2   | 3.00 | 0.17  | 290.2840| 0    | 0     |
| 3   | 3   | 0   | 1   | 1   | 0   | 3.00 | 0.99  | 0.1902 | 0    | 0     |
| 3   | 3   | 2   | 0   | 0   | 1   | 3.07 | 1.43  | 0.0000 | 0    | -     |
| 3   | 3   | 1   | 1   | -1  | 1   | 3.09 | 1.22  | 0.1072 | 0    | -     |
| 3   | 3   | -1  | 1   | 1   | 1   | 3.09 | 1.22  | 0.0419 | 0    | -     |
| 3   | 3   | 1   | 1   | -1  | 1   | 3.09 | 1.22  | 0.0419 | 0    | -     |
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

| $h$ | $h'$ | $k$ | $k'$ | $l$ | $l'$ | $Q_{||}$ | $Q_{\perp}$ | $I$ | $\varphi$ | $\varphi'$ |
|-----|-----|-----|-----|-----|-----|---------|---------|-----|----------|----------|
| 2   | 4   | 0   | 0   | 0   | 0   | 3.15    | 0.28    | 193.8670 | 0  | 0        |
| 3   | 3   | 1   | 0   | 0   | 2   | 3.18    | 1.18    | 0.2305    | 0  | $\pi$    |
| 2   | 4   | -1  | 1   | 1   | 0   | 3.18    | 1.18    | 0.0086    | 0  | $-\ast$  |
| 2   | 4   | 0   | 1   | -1  | 1   | 3.21    | 1.08    | 0.0143    | $\pi$ | $-\ast$  |
| 3   | 3   | 0   | 0   | 2   | 1   | 3.21    | 1.08    | 0.1847    | 0  | 0        |
| 2   | 4   | 1   | 0   | 0   | 1   | 3.23    | 0.76    | 1.9273    | $\pi$ | $\pi$ $\ast$ |
| 2   | 4   | 2   | 0   | 0   | 0   | 3.24    | 1.24    | 0.0000    | 0  | $-\ast$  |
| 2   | 4   | 0   | 0   | 2   | 0   | 3.24    | 1.24    | 0.0524    | $\pi$ | $-\ast$  |
| 4   | 3   | 0   | 0   | 1   | 0   | 3.31    | 1.42    | 0.1137    | 0  | $-\ast$  |
| 2   | 4   | -1  | 2   | 0   | 1   | 3.31    | 1.42    | 0.0038    | 0  | $-\ast$  |
| 2   | 4   | 1   | 1   | 1   | 0   | 3.32    | 0.70    | 0.0000    | 0  | $-\ast$  |
| 2   | 4   | 0   | 1   | 1   | 1   | 3.35    | 0.52    | 3.2888    | $\pi$ | $\pi$    |
| 2   | 4   | 0   | 0   | 0   | 2   | 3.37    | 0.80    | 0.0076    | $\pi$ | $-\ast$  |
| 2   | 4   | 0   | 2   | 0   | 2   | 3.37    | 0.80    | 0.4498    | $\pi$ | $-\ast$  |
| 4   | 3   | 0   | 1   | 0   | 1   | 3.40    | 1.39    | 0.0067    | $\pi$ | $-\ast$  |
| 1   | 5   | 1   | 0   | 0   | 0   | 3.40    | 1.39    | 0.0323    | $\pi$ | $-\ast$  |
| 3   | 3   | 0   | 0   | 0   | 3   | 3.43    | 1.31    | 0.5033    | $\pi$ | $-\ast$  |
| 1   | 5   | 0   | 0   | 0   | 1   | 3.43    | 1.31    | 0.2585    | 0  | 0        |
| 4   | 3   | 1   | 1   | 0   | 0   | 3.43    | 1.31    | 0.0000    | $\pi$ | $-\ast$  |
| 2   | 4   | 1   | 0   | 2   | 1   | 3.44    | 1.07    | 0.0048    | 0  | $-\ast$  |
| 1   | 5   | 0   | 1   | 1   | 0   | 3.45    | 1.44    | 0.0000    | 0  | $-\ast$  |
| 4   | 3   | 1   | 0   | 1   | 1   | 3.45    | 1.44    | 0.0176    | $\pi$ | $-\ast$  |
| 2   | 4   | -1  | 1   | 1   | 2   | 3.53    | 1.02    | 0.0176    | $\pi$ | $-\ast$  |
| 3   | 4   | 0   | 0   | -1  | 1   | 3.53    | 1.02    | 0.0807    | 0  | $-\ast$  |
| 3   | 4   | 1   | 0   | 1   | 0   | 3.56    | 0.91    | 0.3888    | $\pi$ | 0 $\ast$ |
| 3   | 4   | 0   | 1   | 0   | 0   | 3.57    | 0.49    | 0.7713    | $\pi$ | 0 $\ast$ |
| 3   | 4   | -1  | 1   | 0   | 1   | 3.58    | 1.09    | 0.0517    | 0  | $-\ast$  |
| 4   | 3   | 0   | 0   | 1   | 2   | 3.65    | 1.30    | 0.0000    | 0  | $-\ast$  |
| 3   | 4   | 0   | 1   | 2   | 0   | 3.65    | 1.30    | 0.0133    | 0  | $-\ast$  |
| 3   | 4   | 0   | 0   | 1   | 1   | 3.65    | 0.39    | 13.8646   | 0  | 0        |
| 2   | 4   | 0   | 0   | 2   | 2   | 3.70    | 0.54    | 1.1535    | $\pi$ | $\pi$    |
| 3   | 4   | 1   | 1   | 0   | 1   | 3.70    | 0.54    | 0.2100    | 0  | 0        |
| 3   | 4   | 0   | 2   | -1  | 1   | 3.73    | 1.26    | 0.0000    | $\pi$ | $-\ast$  |
| 3   | 4   | 2   | 0   | 1   | 1   | 3.73    | 1.26    | 0.0332    | 0  | $-\ast$  |
| 2   | 5   | -1  | 1   | 0   | 0   | 3.76    | 1.17    | 0.1763    | 0  | 0        |
| 2   | 5   | 0   | 0   | 1   | 0   | 3.77    | 0.89    | 0.0299    | $\pi$ | $-\ast$  |
| 3   | 4   | 2   | 1   | 0   | 0   | 3.77    | 0.89    | 0.3431    | $\pi$ | $-\ast$  |
| 3   | 4   | 0   | 1   | 0   | 2   | 3.77    | 0.89    | 0.0147    | $\pi$ | $-\ast$  |
| 2   | 5   | 1   | 0   | -1  | 1   | 3.78    | 1.32    | 0.0076    | $\pi$ | $-\ast$  |
| 3   | 4   | -1  | 1   | 2   | 1   | 3.78    | 1.32    | 0.0005    | $\pi$ | $-\ast$  |
| 2   | 5   | 2   | 0   | 1   | 0   | 3.84    | 1.50    | 0.0062    | 0  | $-\ast$  |
| 2   | 5   | 0   | 0   | -1  | 2   | 3.84    | 1.50    | 0.0606    | $\pi$ | $-\ast$  |
| 2   | 5   | 0   | 1   | 0   | 1   | 3.85    | 0.84    | 0.1837    | $\pi$ | 0 $\ast$ |
| 2   | 5   | 1   | 1   | 0   | 0   | 3.88    | 0.69    | 5.2985    | $\pi$ | 0        |
| 3   | 4   | 1   | 0   | 1   | 2   | 3.88    | 0.69    | 2.6812    | $\pi$ | $\pi$    |
| 2   | 5   | 1   | 0   | 1   | 1   | 3.89    | 0.92    | 0.0483    | 0  | $-\ast$  |
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

| $h$ | $h'$ | $k$ | $k'$ | $l$ | $l'$ | $Q_{||}$ | $Q_{\perp}$ | $I$ | $\varphi$ | $\varphi'$ |
|-----|------|-----|------|-----|------|---------|------------|-----|---------|---------|
| 4   | 4    | 0   | 0    | 0   | 0    | 3.89    | 0.92       | 0.3057 | 0       | 0       |
| 3   | 4    | 0   | 0    | 3   | 1    | 3.92    | 1.47       | 0.0208 | $\pi$   | $-$     |
| 4   | 4    | $-1$| 1    | 1    | 0    | 3.92    | 1.47       | 0.0000 | 0       | $-$     |
| 4   | 4    | 0   | 1    | $-1$| 1    | 3.95    | 1.39       | 0.0080 | $\pi$   | $-$     |
| 2   | 5    | 1   | 1    | 2   | 0    | 3.95    | 1.39       | 0.0033 | $\pi$   | $-$     |
| 2   | 5    | $-1$| 1    | 0   | 2    | 3.95    | 1.39       | 0.0000 | 0       | $-$     |
| 4   | 4    | 1   | 0    | 0   | 1    | 3.96    | 1.16       | 0.0719 | $\pi$   | $0$     |
| 2   | 5    | 0   | 2    | 1   | 0    | 3.96    | 1.16       | 0.0596 | $-$     | $-$     |
| 2   | 5    | 0   | 0   | 1   | 2    | 4.07    | 0.67       | 4.2071 | $\pi$   | $0$     |
| 3   | 5    | 1   | 0    | 0   | 0    | 4.14    | 0.60       | 0.0467 | $-$     | $-$     |
| 3   | 4    | 0   | 0   | 1   | 3    | 4.14    | 0.60       | 0.7252 | $\pi$   | $0$     |
| 3   | 5    | 0   | 0   | 0   | 1    | 4.17    | 0.38       | 1.0493 | 0       | $-$     |
| 3   | 5    | 0   | 1   | 1   | 0    | 4.18    | 0.71       | 0.1642 | $\pi$   | $\pi$   |
| 3   | 5    | 1   | 0   | 0   | 2    | 4.31    | 0.96       | 0.0000 | 0       | $-$     |
| 3   | 5    | 0   | 2   | 0   | 1    | 4.34    | 0.83       | 0.2427 | $\pi$   | $\pi$   |
| 3   | 5    | 0   | 0   | 2   | 1    | 4.34    | 0.83       | 0.9844 | $\pi$   | $0$     |
| 3   | 5    | 1   | 1   | 1   | 1    | 4.35    | 0.33       | 31.7931 | 0       | 0       |
| 3   | 5    | 1   | 2   | 0   | 0    | 4.41    | 0.15       | 58.6973 | 0       | 0       |
| 3   | 5    | 0   | 1   | 1   | 2    | 4.45    | 0.40       | 20.9346 | 0       | 0       |
| 4   | 5    | 0   | 0   | 1   | 0    | 4.51    | 0.81       | 0.7577 | $\pi$   | $0$     |
| 3   | 5    | 1   | 0   | 2   | 2    | 4.57    | 0.76       | 0.0423 | $\pi$   | $-$     |
| 4   | 5    | 0   | 1   | 0   | 1    | 4.57    | 0.76       | 0.3911 | $\pi$   | $\pi$   |
| 4   | 5    | 1   | 1   | 0   | 0    | 4.60    | 0.59       | 0.9008 | $\pi$   | $\pi$   |
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

| $h$ | $h'$ | $k$ | $k'$ | $l$ | $l'$ | $Q_{||}$ | $Q_{\perp}$ | $I$ | $\varphi$ | $\varphi'$ |
|-----|------|-----|------|-----|------|--------|---------|----|--------|--------|
| 3   | 7    | 1   | 0    | 0   | 0    | 5.34   | 1.00    | 0.2133 | 0      | $\pi$  |
| 4   | 6    | -1  | 1    | 1   | 2    | 5.34   | 1.00    | 0.0554 | $\pi$  | $-\ast$|
| 4   | 6    | 2   | 1    | 1   | 1    | 5.36   | 0.88    | 0.1453 | $\pi$  | 0      |
| 3   | 7    | 0   | 0    | 0   | 1    | 5.36   | 0.88    | 0.7743 | $\pi$  | $\pi$  |
| 4   | 6    | 1   | 2    | 0   | 1    | 5.37   | 0.44    | 0.3892 | 0      | 0      |
| 4   | 6    | 1   | 1    | 1   | 2    | 5.42   | 0.32    | 17.7992| 0      | 0      |
| 4   | 6    | 0   | 0    | 2   | 2    | 5.46   | 0.49    | 1.6188 | 0      | 0      |
| 4   | 6    | 2   | 2    | 0   | 0    | 5.46   | 0.49    | 0.6315 | 0      | $\pi$  |
| 3   | 7    | 1   | 1    | 1   | 1    | 5.50   | 0.86    | 0.1878 | $\pi$  | $0\ast$|
| 5   | 6    | 0   | 1    | 0   | 0    | 5.50   | 0.86    | 0.1013 | $\pi$  | $-$    |
| 3   | 7    | 1   | 2    | 0   | 0    | 5.55   | 0.81    | 0.3404 | $\pi$  | $\pi$  |
| 3   | 6    | 0   | 0    | 3   | 3    | 5.55   | 0.81    | 0.1036 | 0      | $-$    |
| 5   | 6    | 0   | 0    | 1   | 1    | 5.55   | 0.81    | 0.0539 | $\pi$  | $-$    |
| 4   | 6    | 0   | 1    | 1   | 3    | 5.57   | 0.66    | 2.9595 | $\pi$  | $\pi$  |
| 3   | 7    | 0   | 1    | 1   | 2    | 5.59   | 0.89    | 0.0125 | 0      | $-\ast$|
| 5   | 6    | 1   | 1    | 0   | 1    | 5.59   | 0.89    | 0.0162 | 0      | $-\ast$|
| 5   | 6    | 1   | 2    | 1   | 0    | 5.70   | 0.99    | 0.0153 | $\pi$  | $-\ast$|
| 4   | 7    | -1  | 1    | 1   | 0    | 5.70   | 0.99    | 0.0032 | $\pi$  | $-$    |
| 5   | 6    | 1   | 0    | 1   | 2    | 5.70   | 0.99    | 0.0107 | $\pi$  | $-\ast$|
| 4   | 7    | 0    | 0    | 1   | 0    | 5.71   | 0.63    | 1.1373 | $\pi$  | 0      |
| 4   | 6    | 1    | 0    | 2    | 3    | 5.71   | 0.63    | 4.1409 | $\pi$  | 0      |
| 4   | 7    | 0    | 1    | 0    | 1    | 5.76   | 0.56    | 0.5770 | 0      | 0      |
| 4   | 7    | 1    | 1    | 0    | 0    | 5.78   | 0.30    | 0.5324 | 0      | 0      |

| $h$ | $h'$ | $k$ | $k'$ | $l$ | $l'$ | $Q_{||}$ | $Q_{\perp}$ | $I$ | $\varphi$ | $\varphi'$ |
|-----|------|-----|------|-----|------|--------|---------|----|--------|--------|
| 4   | 7    | 1   | 0    | 0   | 1    | 5.79   | 0.67    | 0.2574 | $\pi$  | $\pi$  |
| 4   | 7    | 0   | 2    | 1   | 0    | 5.83   | 0.98    | 0.0153 | 0      | $-\ast$|
| 5   | 6    | 0   | 1    | 2   | 2    | 5.83   | 0.98    | 0.0000 | $\pi$  | $-\ast$|
| 5   | 6    | 0   | 0    | 1   | 3    | 5.88   | 0.93    | 0.0185 | $\pi$  | $-$    |
| 4   | 7    | 0   | 1    | 2   | 1    | 5.88   | 0.93    | 0.0000 | 0      | $-\ast$|
| 4   | 7    | 2   | 1    | 0   | 1    | 5.88   | 0.93    | 0.0000 | 0      | $-\ast$|
| 3   | 7    | 0   | 0    | 2   | 3    | 5.90   | 0.80    | 1.6879 | $\pi$  | 0      |
| 4   | 7    | 1   | 1    | 0   | 2    | 5.90   | 0.80    | 0.6050 | $\pi$  | $\pi\ast$|
| 4   | 7    | 0   | 0    | 1   | 2    | 5.91   | 0.24    | 66.6811| 0      | 0      |
| 4   | 6    | 0   | 0    | 2   | 4    | 5.99   | 0.33    | 57.0894| 0      | 0      |
| 4   | 7    | 1   | 2    | 1   | 1    | 5.99   | 0.33    | 49.9870| 0      | 0      |
| 4   | 7    | 2   | 2    | 1   | 0    | 6.03   | 0.78    | 0.4304 | $\pi$  | 0      |
| 4   | 7    | 0   | 2    | 1   | 2    | 6.03   | 0.78    | 0.4387 | $\pi$  | 0      |
| 5   | 7    | 1   | 0    | 0   | 0    | 6.08   | 0.73    | 0.0456 | 0      | $-$    |
| 4   | 7    | 1   | 1    | 2    | 2    | 6.10   | 0.55    | 0.0456 | $\pi$  | $-$    |
| 4   | 7    | 1   | 3    | 0    | 0    | 6.10   | 0.55    | 4.8214 | $\pi$  | $\pi$  |
| 5   | 7    | 0    | 0    | 0    | 1    | 6.10   | 0.55    | 3.8116 | $\pi$  | $\pi$  |
| 4   | 7    | 0    | 1    | 0    | 3    | 6.11   | 0.81    | 0.0138 | 0      | $-\ast$|
| 5   | 7    | 0    | 1    | 1    | 0    | 6.11   | 0.81    | 0.0000 | 0      | $-\ast$|
| 5   | 7    | 0    | 2    | 0    | 1    | 6.22   | 0.93    | 0.2503 | $\pi$  | 0      |
| 5   | 7    | 0    | 0    | 2    | 1    | 6.22   | 0.93    | 0.0000 | 0      | $-$    |
| 5   | 7    | 1    | 1    | 1    | 1    | 6.22   | 0.52    | 1.7449 | $\pi$  | $\pi\ast$|
| 5   | 7    | 1    | 2    | 0    | 0    | 6.27   | 0.43    | 9.0592 | 0      | 0      |
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

| $h$ | $h'$ | $k$ | $k'$ | $l$ | $l'$ | $Q_{||}$ | $Q_{\perp}$ | $I$ | $\varphi$ | $\varphi'$ |
|-----|-----|-----|-----|-----|-----|--------|---------|-----|----------|----------|
| 4   | 7   | 0   | 1   | 2   | 3   | 6.27   | 0.43    | 2.0398 | 0         | 0         |
| 4   | 8   | 0   | 0   | 0   | 0   | 6.30   | 0.57    | 0.3161 | $\pi$     | 0         |
| 5   | 7   | 0   | 1   | 1   | 2   | 6.30   | 0.57    | 0.0487 | $\pi$     | *         |
| 4   | 7   | 0   | 0   | 1   | 4   | 6.34   | 0.91    | 0.0234 | $\pi$     | *         |
| 4   | 8   | 1   | 0   | 0   | 1   | 6.34   | 0.91    | 0.0115 | $\pi$     | *         |
| 5   | 7   | 1   | 2   | 0   | 2   | 6.38   | 0.86    | 0.0542 | $\pi$     | *         |
| 5   | 7   | 1   | 0   | 2   | 2   | 6.38   | 0.86    | 0.1996 | $\pi$ $\pi$ | *         |
| 4   | 8   | 1   | 1   | 1   | 1   | 6.38   | 0.86    | 0.0000 | 0         | *         |
| 5   | 7   | 2   | 2   | 0   | 1   | 6.40   | 0.72    | 2.6114 | $\pi$ $\pi$ |         |
| 4   | 8   | 0   | 1   | 1   | 1   | 6.40   | 0.72    | 1.6884 | $\pi$ 0 *  |         |
| 4   | 7   | 1   | 0   | 3   | 3   | 6.41   | 0.94    | 0.0161 | $\pi$     |         |
| 4   | 8   | 0   | 0   | 0   | 2   | 6.41   | 0.94    | 0.0179 | 0         | *         |
| 4   | 8   | 0   | 2   | 0   | 0   | 6.41   | 0.94    | 0.0899 | 0         | *         |
| 5   | 7   | 2   | 2   | 1   | 2   | 6.41   | 0.94    | 0.0064 | 0         | *         |
| 4   | 7   | 1   | 2   | 0   | 1   | 6.52   | 0.69    | 1.3308 | $\pi$ 0 *  |         |
| 5   | 7   | 1   | 1   | 1   | 3   | 6.52   | 0.69    | 1.6651 | $\pi$ $\pi$ |         |
| 4   | 8   | 1   | 1   | 1   | 2   | 6.57   | 0.63    | 0.3075 | $\pi$ $\pi$ | *         |
| 5   | 7   | 0   | 0   | 2   | 3   | 6.58   | 0.41    | 0.1848 | 0         | 0         |
| 4   | 8   | 2   | 2   | 0   | 0   | 6.59   | 0.73    | 0.3522 | $\pi$ $\pi$ |         |
| 4   | 8   | 0   | 0   | 2   | 2   | 6.59   | 0.73    | 0.0403 | 0         | *         |
| 5   | 8   | 0   | 0   | −1  | 1   | 6.67   | 0.97    | 0.0649 | $\pi$     | −         |
| 4   | 8   | 1   | 3   | 1   | 0   | 6.67   | 0.97    | 0.0288 | $\pi$     | −         |
| 4   | 8   | 0   | 1   | 1   | 3   | 6.69   | 0.85    | 0.2809 | $\pi$ 0 *  |         |

| $h$ | $h'$ | $k$ | $k'$ | $l$ | $l'$ | $Q_{||}$ | $Q_{\perp}$ | $I$ | $\varphi$ | $\varphi'$ |
|-----|-----|-----|-----|-----|-----|--------|---------|-----|----------|----------|
| 5   | 8   | 1   | 0   | 1   | 0   | 6.69   | 0.85    | 0.2092 | $\pi$ 0 *  |         |
| 4   | 7   | 0   | 0   | 3   | 4   | 6.70   | 0.37    | 8.7933 | 0         | 0         |
| 5   | 8   | 0   | 1   | 0   | 0   | 6.70   | 0.37    | 2.5066 | 0         | 0         |
| 5   | 8   | 0   | 0   | 1   | 1   | 6.74   | 0.23    | 19.5604 | 0         | 0         |
| 5   | 8   | 1   | 1   | 0   | 1   | 6.77   | 0.44    | 5.6730 | 0         | *         |
| 4   | 8   | 1   | 0   | 2   | 3   | 6.80   | 0.83    | 0.3077 | $\pi$ 0 *  |         |
| 5   | 8   | 2   | 1   | 0   | 0   | 6.80   | 0.83    | 0.3414 | $\pi$     | 0         |
| 5   | 8   | 0   | 1   | 0   | 2   | 6.80   | 0.83    | 0.1395 | $\pi$ $\pi$ | *         |
| 5   | 8   | 0   | 2   | 1   | 1   | 6.85   | 0.78    | 0.1358 | $\pi$ 0 *  |         |
| 5   | 7   | 1   | 0   | 2   | 4   | 6.85   | 0.78    | 0.0337 | $\pi$ −    |         |
| 5   | 8   | 1   | 2   | 1   | 0   | 6.86   | 0.62    | 1.8139 | $\pi$ 0 *  |         |
| 5   | 8   | 1   | 0   | 1   | 2   | 6.86   | 0.62    | 1.1959 | $\pi$ 0 *  |         |
| 4   | 9   | 1   | 1   | 0   | 0   | 6.97   | 0.97    | 0.0000 | $\pi$ −    |         |
| 4   | 8   | 0   | 1   | 3   | 3   | 6.97   | 0.97    | 0.0259 | $\pi$ −    |         |
| 5   | 8   | 0   | 1   | 2   | 2   | 6.97   | 0.59    | 2.1570 | $\pi$ $\pi$ | *         |
| 5   | 8   | 2   | 2   | 1   | 1   | 7.02   | 0.51    | 0.6737 | 0         | 0         |
| 5   | 8   | 0   | 0   | 0   | 1   | 3     | 7.02   | 0.51    | 0.2521 | $\pi$ 0    |         |
| 5   | 8   | 2   | 2   | 1   | 2   | 7.03   | 0.20    | 2.8201 | 0         | 0         |
| 6   | 8   | 0   | 0   | 0   | 0   | 7.04   | 0.63    | 0.0000 | 0         | −         |
| 4   | 8   | 0   | 0   | 2   | 4   | 7.04   | 0.63    | 0.0845 | $\pi$ −    |         |
| 5   | 8   | 1   | 3   | 0   | 1   | 7.04   | 0.63    | 0.2139 | $\pi$ $\pi$ |         |
| 6   | 8   | 1   | 0   | 0   | 1   | 7.08   | 0.95    | 0.0041 | $\pi$ −    | *         |
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

<table>
<thead>
<tr>
<th>$h$</th>
<th>$h'$</th>
<th>$k$</th>
<th>$k'$</th>
<th>$l$</th>
<th>$l'$</th>
<th>$Q_{\parallel}$</th>
<th>$Q_{\perp}$</th>
<th>$I$</th>
<th>$\varphi$</th>
<th>$\varphi'$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>7.08</td>
<td>0.95</td>
<td>0.0000</td>
<td>$\pi$</td>
<td>$-$</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>7.08</td>
<td>0.95</td>
<td>0.0032</td>
<td>$0$</td>
<td>$-$</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>7.12</td>
<td>0.90</td>
<td>0.0880</td>
<td>$\pi$</td>
<td>$0$</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>7.12</td>
<td>0.90</td>
<td>0.0100</td>
<td>$0$</td>
<td>$-$</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>7.13</td>
<td>0.77</td>
<td>0.5807</td>
<td>$\pi$</td>
<td>$0$</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>7.13</td>
<td>0.77</td>
<td>1.5237</td>
<td>$\pi$</td>
<td>$0$</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7.14</td>
<td>0.09</td>
<td>115.7430</td>
<td>$0$</td>
<td>$-$</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>7.14</td>
<td>0.98</td>
<td>0.0449</td>
<td>$0$</td>
<td>$-$</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>$-1$</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>7.14</td>
<td>0.98</td>
<td>0.0440</td>
<td>$0$</td>
<td>$-$</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>7.14</td>
<td>0.98</td>
<td>0.0630</td>
<td>$\pi$</td>
<td>$-$</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>7.14</td>
<td>0.98</td>
<td>0.0254</td>
<td>$0$</td>
<td>$-$</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>7.21</td>
<td>0.25</td>
<td>68.5241</td>
<td>$0$</td>
<td>$-$</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>7.24</td>
<td>0.75</td>
<td>0.8290</td>
<td>$\pi$</td>
<td>$\pi$</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>7.28</td>
<td>0.69</td>
<td>0.0000</td>
<td>$0$</td>
<td>$-$</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>7.28</td>
<td>0.69</td>
<td>0.0158</td>
<td>$0$</td>
<td>$-$</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7.28</td>
<td>0.69</td>
<td>1.1549</td>
<td>$\pi$</td>
<td>$\pi$</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7.30</td>
<td>0.50</td>
<td>0.1212</td>
<td>$0$</td>
<td>$0$</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>7.31</td>
<td>0.78</td>
<td>0.0000</td>
<td>$0$</td>
<td>$-$</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>7.31</td>
<td>0.78</td>
<td>0.0705</td>
<td>$\pi$</td>
<td>$-$</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>7.31</td>
<td>0.78</td>
<td>0.0000</td>
<td>$0$</td>
<td>$-$</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>7.40</td>
<td>0.90</td>
<td>0.0966</td>
<td>$\pi$</td>
<td>$-$</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>7.40</td>
<td>0.90</td>
<td>0.0632</td>
<td>$\pi$</td>
<td>$0$</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>7.40</td>
<td>0.90</td>
<td>0.3100</td>
<td>$\pi$</td>
<td>$\pi$</td>
</tr>
</tbody>
</table>
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

| $h$ | $h'$ | $k$ | $k'$ | $l$ | $l'$ | $Q_{||}$ | $Q_{\perp}$ | $I$ | $\varphi$ | $\varphi'$ |
|-----|------|-----|------|-----|------|---------|-----------|-----|----------|---------|
| 6   | 9    | 2   | 1    | 1   | 0    | 0.95    | 0.0000   | 0   | $-\pi$   | 0       |
| 5   | 9    | 1   | 4    | 0   | 0    | 0.95    | 0.5632   | $\pi$| 0        | $\pi$   |
| 5   | 9    | 0   | 2    | 2   | 3    | 0.82    | 1.4478   | $\pi$| 0        | $\pi$   |
| 6   | 9    | 1   | 1    | 0   | 2    | 0.82    | 0.5873   | $\pi$| 0        | $\pi$   |
| 6   | 9    | 0   | 0    | 1   | 2    | 0.30    | 15.4622  | 0   | 0        | $\pi$   |
| 5   | 8    | 0   | 0    | 3   | 5    | 0.06    | 21.8775  | 0   | 0        | $\pi$   |
| 6   | 9    | 1   | 2    | 1   | 1    | 0.38    | 13.2323  | 0   | 0        | $\pi$   |
| 5   | 9    | 1   | 1    | 1   | 3    | 0.80    | 0.0961   | 0   | $-\pi$   | 0       |
| 5   | 10   | 0   | 1    | 0   | 0    | 0.80    | 0.5781   | $\pi$| 0        | $\pi$   |
| 6   | 9    | 2    | 2    | 1   | 0    | 0.80    | 0.1873   | $\pi$| $\pi$    | $\pi$   |
| 6   | 9    | 0    | 2    | 1   | 2    | 0.80    | 0.5594   | $\pi$| $\pi$    | $\pi$   |
| 5   | 10   | 0    | 0    | 1   | 1    | 0.80    | 0.2752   | $\pi$| 0        | $\pi$   |
| 5   | 9    | 1    | 0    | 2    | 4    | 0.75    | 0.1448   | $\pi$| 0        | $\pi$   |
| 6   | 9    | 1    | 1    | 2    | 2    | 0.58    | 2.8360   | $\pi$| $\pi$    | $\pi$   |
| 6   | 9    | 1    | 3    | 0    | 0    | 0.58    | 0.2667   | $\pi$| 0        | $\pi$   |
| 5   | 10   | 1    | 1    | 0    | 1    | 0.80    | 0.0518   | $\pi$| $-\pi$   | $\pi$   |
| 6   | 9    | 1    | 0    | 1    | 3    | 0.83    | 0.0143   | 0   | $-\pi$   | 0       |
| 5   | 10   | 1    | 0    | 1    | 2    | 0.94    | 0.0000   | $\pi$| $-\pi$   | $\pi$   |
| 7   | 9    | 0    | 0    | 0    | 1    | 0.94    | 0.1200   | $\pi$| $\pi$    | $\pi$   |
| 6   | 9    | 1    | 3    | 0    | 2    | 0.94    | 0.0000   | $\pi$| $-\pi$   | $\pi$   |
| 5   | 10   | 1    | 2    | 1    | 0    | 0.94    | 0.1993   | $\pi$| $\pi$    | $\pi$   |
| 6   | 9    | 2    | 2    | 1    | 2    | 0.94    | 1.6493   | $\pi$| $\pi$    | $\pi$   |
| 6   | 9    | 0    | 1    | 2    | 3    | 0.94    | 2.1734   | 0   | 0        | $\pi$   |
| 5   | 10   | 1    | 1    | 2    | 3    | 0.94    | 0.0000   | 0   | $-\pi$   | $\pi$   |
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

<table>
<thead>
<tr>
<th>h</th>
<th>h'</th>
<th>k</th>
<th>k'</th>
<th>l</th>
<th>l'</th>
<th>Q_∥</th>
<th>Q_⊥</th>
<th>I</th>
<th>φ</th>
<th>φ'</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>10</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>8.33</td>
<td>0.75</td>
<td>0.0851</td>
<td>π</td>
<td>−</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>8.33</td>
<td>0.75</td>
<td>0.0262</td>
<td>π</td>
<td>−</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8.40</td>
<td>0.99</td>
<td>0.0000</td>
<td>0</td>
<td>− *</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>8.40</td>
<td>0.99</td>
<td>0.1231</td>
<td>π</td>
<td>π</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>−1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>8.40</td>
<td>0.99</td>
<td>0.0000</td>
<td>π</td>
<td>− *</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8.41</td>
<td>0.87</td>
<td>0.2417</td>
<td>π</td>
<td>π *</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>8.41</td>
<td>0.87</td>
<td>0.0505</td>
<td>π</td>
<td>−</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>8.42</td>
<td>0.41</td>
<td>0.3886</td>
<td>π</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>8.42</td>
<td>0.41</td>
<td>0.7834</td>
<td>0</td>
<td>0 *</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>8.45</td>
<td>0.29</td>
<td>7.9406</td>
<td>0</td>
<td>0 *</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>8.47</td>
<td>0.47</td>
<td>3.8846</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>8.47</td>
<td>0.47</td>
<td>0.7321</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>8.50</td>
<td>0.85</td>
<td>0.1087</td>
<td>π</td>
<td>0 *</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>8.50</td>
<td>0.85</td>
<td>0.0415</td>
<td>π</td>
<td>− *</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>8.54</td>
<td>0.80</td>
<td>0.1863</td>
<td>π</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>8.54</td>
<td>0.80</td>
<td>0.0609</td>
<td>π</td>
<td>0 *</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>8.55</td>
<td>0.64</td>
<td>0.4476</td>
<td>π</td>
<td>π *</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>8.55</td>
<td>0.64</td>
<td>0.1810</td>
<td>0</td>
<td>−</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>8.56</td>
<td>0.88</td>
<td>0.0190</td>
<td>0</td>
<td>− *</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>8.56</td>
<td>0.88</td>
<td>0.0079</td>
<td>0</td>
<td>− *</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>8.63</td>
<td>0.98</td>
<td>0.0476</td>
<td>0</td>
<td>− *</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>2</td>
<td>3</td>
<td>−1</td>
<td>1</td>
<td>8.63</td>
<td>0.98</td>
<td>0.0533</td>
<td>π</td>
<td>− *</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>8.63</td>
<td>0.98</td>
<td>0.0088</td>
<td>π</td>
<td>− *</td>
</tr>
</tbody>
</table>
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

| \(h\) | \(h'\) | \(k\) | \(k'\) | \(l\) | \(l'\) | \(Q_||\) | \(Q_{\perp}\) | \(I\) | \(\varphi\) | \(\varphi'\) |
|-------|-------|-------|-------|-------|-------|--------|--------|-------|-------|-------|
| 6     | 11    | 0     | 0     | 1     | 0     | 8.86   | 0.77   | 0.1289 | \(\pi\) | 0     |
| 7     | 10    | 0     | 1     | 2     | 2     | 8.86   | 0.77   | 0.2412 | \(\pi\) | \(\pi\) | *     |
| 6     | 11    | 0     | 1     | 0     | 1     | 8.89   | 0.71   | 0.0140 | \(\pi\) | \(-\) * |
| 7     | 10    | 0     | 0     | 1     | 3     | 8.89   | 0.71   | 0.3892 | \(\pi\) | \(\pi\) |       |
| 5     | 10    | 0     | 0     | 3     | 5     | 8.89   | 0.71   | 0.5930 | \(\pi\) | 0     |       |
| 7     | 10    | 2     | 2     | 1     | 1     | 8.89   | 0.71   | 0.1464 | \(\pi\) | \(\pi\) | *     |
| 6     | 11    | 1     | 1     | 0     | 0     | 8.90   | 0.53   | 1.5611 | \(\pi\) | 0     |       |
| 7     | 10    | 1     | 2     | 1     | 2     | 8.90   | 0.53   | 0.0618 | \(\pi\) | \(\pi\) | *     |
| 6     | 11    | 1     | 0     | 1     | 2     | 8.91   | 0.80   | 0.0184 | \(\pi\) | \(-\) * |
| 6     | 10    | 0     | 2     | 2     | 4     | 8.91   | 0.80   | 0.0574 | 0     | \(-\) * |
| 7     | 10    | 1     | 3     | 0     | 1     | 8.91   | 0.80   | 0.0552 | 0     | \(-\) * |
| 6     | 11    | 1     | 1     | 0     | 2     | 8.98   | 0.91   | 0.1054 | \(\pi\) | \(\pi\) |
| 6     | 11    | 0     | 0     | 1     | 2     | 8.99   | 0.50   | 0.0577 | \(\pi\) |       |
| 6     | 9     | 0     | 0     | 3     | 6     | 8.99   | 0.50   | 0.0682 | 0     | \(-\) * |
| 7     | 10    | 2     | 3     | 0     | 0     | 8.99   | 0.50   | 0.2502 | \(\pi\) |       |
| 6     | 10    | 1     | 1     | 3     | 4     | 9.02   | 0.41   | 3.8882 | 0     |       |
| 6     | 11    | 2     | 1     | 1     | \(\times\) | 9.04   | 0.55   | 0.2940 | 0     | \(0\) * |
| 7     | 10    | 1     | 1     | 2     | \(\times\) | 9.04   | 0.55   | 0.0721 | \(\pi\) | \(0\) * |
| 6     | 10    | 1     | 0     | 2     | 5     | 9.07   | 0.90   | 0.0000 | \(\pi\) | \(-\) * |
| 6     | 11    | 2     | 2     | 1     | 0     | 9.07   | 0.90   | 0.0878 | \(\pi\) | \(0\) * |
| 6     | 11    | 0     | 2     | 1     | 2     | 9.07   | 0.90   | 0.0258 | 0     | \(-\) * |
| 7     | 10    | 2     | 3     | 0     | 2     | 9.07   | 0.90   | 0.0000 | \(\pi\) |       |
| 7     | 10    | 0     | 0     | 3     | 3     | 9.10   | 0.85   | 0.0689 | \(\pi\) |       |
| 6     | 11    | 1     | 0     | 1     | 0     | 8.90   | 0.85   | 0.0689 | \(\pi\) |       |
| 7     | 10    | 2     | 2     | 1     | 3     | 9.10   | 0.85   | 0.0633 | \(\pi\) | \(-\) * |
| 6     | 11    | 1     | 3     | 0     | 0     | 9.11   | 0.70   | 1.6525 | \(\pi\) | \(\pi\) |
| 6     | 11    | 1     | 1     | 2     | 2     | 9.11   | 0.70   | 0.2080 | \(\pi\) | \(0\) * |
| 6     | 10    | 0     | 0     | 4     | 4     | 9.12   | 0.93   | 0.0153 | 0     | \(-\) * |
| 6     | 11    | 1     | 0     | 1     | 3     | 9.12   | 0.93   | 0.0288 | \(\pi\) | \(-\) * |
| 7     | 10    | 0     | 1     | 2     | 4     | 9.19   | 0.68   | 0.4663 | \(\pi\) | \(\pi\) |
| 6     | 11    | 2     | 2     | 1     | 2     | 9.19   | 0.68   | 1.2595 | \(\pi\) | \(\pi\) |
| 6     | 11    | 0     | 1     | 2     | 3     | 9.23   | 0.61   | 0.1266 | \(\pi\) | \(\pi\) |
| 6     | 11    | 2     | 3     | 0     | 1     | 9.23   | 0.61   | 0.2271 | 0     | \(0\) * |
| 7     | 11    | 1     | 0     | 0     | 0     | 9.23   | 0.61   | 0.0144 | 0     | \(-\) * |
| 7     | 11    | 0     | 0     | 0     | 1     | 9.24   | 0.39   | 0.2101 | 0     | \(\pi\) |
| 7     | 11    | 1     | 0     | 1     | 3     | 9.24   | 0.39   | 0.1231 | \(\pi\) | \(\pi\) |
| 7     | 11    | 0     | 0     | 0     | 1     | 9.25   | 0.72   | 0.0000 | \(\pi\) | \(-\) * |
| 7     | 11    | 2     | 1     | 1     | 3     | 9.25   | 0.72   | 0.0000 | 0     | \(-\) * |
| 6     | 11    | 1     | 2     | 1     | 2     | 9.30   | 0.96   | 0.0026 | \(\pi\) | \(-\) * |
| 7     | 11    | 1     | 0     | 0     | 2     | 9.30   | 0.96   | 0.0117 | \(\pi\) | \(-\) * |
| 6     | 11    | 2     | 3     | 2     | 1     | 9.30   | 0.96   | 0.0000 | 0     | \(-\) * |
| 7     | 11    | 0     | 0     | 2     | 1     | 9.32   | 0.84   | 0.0851 | \(\pi\) | \(\pi\) |
| 6     | 11    | 3     | 3     | 0     | 0     | 9.32   | 0.84   | 0.4567 | \(\pi\) | 0     |
| 7     | 11    | 0     | 0     | 2     | 1     | 9.32   | 0.84   | 0.2744 | \(\pi\) | \(\pi\) |
| 7     | 11    | 1     | 3     | 2     | 2     | 9.32   | 0.84   | 0.1441 | \(\pi\) | 0     |
| 7     | 10    | 1     | 0     | 3     | 4     | 9.32   | 0.84   | 0.3925 | \(\pi\) | \(0\) * |
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

\begin{table}[h]
\centering
\begin{tabular}{cccccccccccc}
\hline
$h$ & $h'$ & $k$ & $k'$ & $l$ & $l'$ & $Q_{||}$ & $Q_{\perp}$ & $I$ & $\varphi$ & $\varphi'$ & \\
\hline
7 & 11 & 1 & 1 & 1 & 1 & 9.32 & 0.35 & 2.2070 & 0 & 0 & * \\
7 & 11 & 1 & 2 & 0 & 0 & 9.35 & 0.19 & 11.5073 & 0 & 0 & \\
7 & 11 & 0 & 1 & 1 & 2 & 9.37 & 0.42 & 2.4959 & 0 & 0 & * \\
6 & 10 & 1 & 0 & 4 & 5 & 9.40 & 0.82 & 0.1521 & $\pi$ & $\pi$ & \\
6 & 11 & 2 & 4 & 1 & 0 & 9.40 & 0.82 & 0.0598 & $\pi$ & $-$ & \\
7 & 11 & 1 & 2 & 0 & 2 & 9.43 & 0.77 & 0.1325 & $\pi$ & $\pi$ & * \\
7 & 11 & 1 & 0 & 2 & 2 & 9.43 & 0.77 & 0.0000 & 0 & $-$ & * \\
7 & 11 & 2 & 2 & 0 & 1 & 9.44 & 0.60 & 0.7395 & $\pi$ & $\pi$ & * \\
6 & 11 & 1 & 1 & 2 & 4 & 9.44 & 0.60 & 0.7828 & $\pi$ & 0 & * \\
6 & 10 & 0 & 0 & 0 & 0 & 9.45 & 0.85 & 0.0736 & $-$ & $-$ & \\
7 & 11 & 2 & 1 & 1 & 2 & 9.45 & 0.85 & 0.0134 & $\pi$ & $-$ & * \\
7 & 10 & 1 & 1 & 2 & 5 & 9.45 & 0.85 & 0.0814 & $\pi$ & $-$ & \\
6 & 11 & 1 & 2 & 3 & 3 & 9.45 & 0.85 & 0.0814 & 0 & $-$ & \\
6 & 12 & 0 & 0 & 0 & 0 & 9.45 & 0.85 & 0.0736 & $-$ & $-$ & \\
7 & 11 & 2 & 2 & 2 & 1 & 9.52 & 0.96 & 0.0575 & $\pi$ & 0 & * \\
6 & 11 & 0 & 0 & 3 & 4 & 9.52 & 0.58 & 0.0981 & $\pi$ & $-$ & \\
7 & 11 & 1 & 1 & 3 & 5 & 9.52 & 0.58 & 0.4996 & $\pi$ & 0 & * \\
7 & 11 & 1 & 3 & 1 & 1 & 9.52 & 0.58 & 1.2260 & $\pi$ & $\pi$ & * \\
7 & 10 & 0 & 0 & 3 & 5 & 9.55 & 0.50 & 1.0722 & 0 & $\pi$ & \\
7 & 11 & 1 & 2 & 2 & 2 & 9.55 & 0.50 & 0.2323 & 0 & $-$ & * \\
7 & 11 & 0 & 0 & 2 & 3 & 9.56 & 0.15 & 1.8594 & 0 & 0 & \\
7 & 11 & 2 & 3 & 1 & 0 & 9.57 & 0.62 & 0.0000 & 0 & $-$ & * \\
8 & 11 & 0 & 0 & 1 & 0 & 9.60 & 0.94 & 0.0341 & $\pi$ & $-$ & \\
6 & 12 & 1 & 2 & 0 & 1 & 9.60 & 0.94 & 0.0000 & $-$ & $-$ & \\
6 & 11 & 0 & 2 & 3 & 4 & 9.60 & 0.94 & 0.0367 & $\pi$ & $-$ & \\
8 & 11 & 0 & 1 & 0 & 1 & 9.63 & 0.89 & 0.0233 & $\pi$ & $-$ & * \\
6 & 12 & 1 & 1 & 1 & 2 & 9.63 & 0.89 & 0.0000 & $-$ & $-$ & \\
6 & 11 & 0 & 1 & 2 & 5 & 9.63 & 0.89 & 0.0000 & 0 & $-$ & * \\
7 & 11 & 1 & 4 & 0 & 0 & 9.63 & 0.89 & 0.0030 & $\pi$ & $-$ & \\
8 & 11 & 1 & 1 & 0 & 0 & 9.64 & 0.76 & 0.0366 & $\pi$ & $-$ & \\
7 & 11 & 0 & 2 & 2 & 3 & 9.64 & 0.76 & 0.5607 & $\pi$ & $\pi$ & * \\
8 & 11 & 1 & 0 & 1 & 1 & 9.65 & 0.97 & 0.0142 & 0 & $-$ & * \\
6 & 12 & 0 & 0 & 2 & 2 & 9.65 & 0.97 & 0.0030 & $-$ & $-$ & \\
6 & 12 & 2 & 2 & 0 & 0 & 9.65 & 0.97 & 0.0000 & $-$ & $-$ & \\
7 & 11 & 0 & 1 & 1 & 4 & 9.65 & 0.97 & 0.0000 & $\pi$ & $-$ & * \\
6 & 10 & 0 & 0 & 4 & 6 & 9.69 & 0.21 & 45.9949 & 0 & 0 & \\
7 & 11 & 2 & 3 & 1 & 2 & 9.69 & 0.21 & 37.7341 & 0 & 0 & \\
8 & 11 & 0 & 0 & 1 & 2 & 9.72 & 0.74 & 0.1389 & $\pi$ & 0 & \\
7 & 11 & 1 & 1 & 3 & 3 & 9.72 & 0.74 & 0.4642 & $\pi$ & 0 & * \\
7 & 11 & 3 & 3 & 1 & 1 & 9.72 & 0.74 & 0.7059 & $\pi$ & $\pi$ & \\
7 & 11 & 1 & 3 & 1 & 3 & 9.72 & 0.74 & 0.2865 & $\pi$ & $\pi$ & \\
7 & 11 & 1 & 0 & 2 & 4 & 9.75 & 0.68 & 0.0000 & 0 & $-$ & * \\
7 & 11 & 2 & 2 & 2 & 3 & 9.76 & 0.48 & 0.2690 & $\pi$ & $\pi$ & \\
7 & 11 & 2 & 4 & 0 & 1 & 9.76 & 0.48 & 0.4766 & $\pi$ & $\pi$ & \\
6 & 11 & 1 & 0 & 3 & 5 & 9.77 & 0.77 & 0.0507 & $\pi$ & $-$ & * \\
8 & 11 & 1 & 2 & 1 & 1 & 9.77 & 0.77 & 0.0185 & 0 & $-$ & * \\
\hline
\end{tabular}
\end{table}
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

| $h$ | $h'$ | $k$ | $k'$ | $l$ | $l'$ | $Q_||$ | $Q_\perp$ | $I$  | $\varphi$ | $\varphi'$ |
|-----|------|-----|------|-----|------|--------|-----------|-----|-----------|-----------|
| 6   | 12   | 2   | 3    | 1   | 1    | 9.83   | 0.89     | 0.2823 | $-\pi$    |           |
| 8   | 11   | 1   | 3    | 0    | 0    | 9.83   | 0.89     | 0.0146 | 0         | $-\pi$    |
| 8   | 11   | 1   | 1    | 2    | 2    | 9.83   | 0.89     | 0.1433 | $\pi$     | $\pi^*$    |
| 7   | 10   | 1   | 0    | 3    | 6    | 9.83   | 0.89     | 0.0601 | $\pi$     |           |
| 7   | 12   | 1   | 0    | 1    | 0    | 9.83   | 0.89     | 0.0000 | $\pi$     | $-\pi$    |
| 7   | 12   | 0   | 1    | 0    | 0    | 9.84   | 0.45     | 0.3406 | $\pi$     |           |
| 7   | 11   | 1   | 2    | 2    | 4    | 9.87   | 0.34     | 2.6427 | 0         |           |
| 7   | 12   | 0   | 0    | 1    | 1    | 9.87   | 0.34     | 6.7895 | 0         |           |
| 7   | 11   | 3   | 4    | 0    | 0    | 9.87   | 0.34     | 3.3607 | 0         |           |
| 7   | 12   | 1   | 1    | 0    | 1    | 9.89   | 0.50     | 0.7597 | 0         | $\pi^*$    |
| 7   | 11   | 0   | 1    | 3    | 4    | 9.89   | 0.50     | 0.3308 | 0         | $\pi^*$    |
| 7   | 12   | 2   | 1    | 0    | 0    | 9.91   | 0.87     | 0.0000 | $\pi$     |           |
| 7   | 12   | 0   | 1    | 0    | 2    | 9.91   | 0.87     | 0.0060 | $\pi$     | $-\pi$    |
| 6   | 12   | 1   | 2    | 2    | 3    | 9.91   | 0.87     | 0.0000 | $-\pi$    |           |
| 8   | 11   | 2   | 2    | 1    | 2    | 9.91   | 0.87     | 0.0608 | $\pi$     | $0^*$      |
| 6   | 11   | 0   | 1    | 4    | 5    | 9.94   | 0.82     | 0.0466 | $\pi$     |           |
| 7   | 12   | 0   | 2    | 1    | 1    | 9.94   | 0.82     | 0.2152 | $\pi$     | $\pi^*$    |
| 8   | 11   | 2    | 3    | 0    | 1    | 9.94   | 0.82     | 0.0000 | 0         | $-\pi$    |
| 8   | 11   | 0    | 1    | 2    | 3    | 9.94   | 0.82     | 0.0000 | 0         | $-\pi$    |
| 7   | 11   | 0   | 0    | 2    | 5    | 9.95   | 0.67     | 1.1349 | $\pi$     |           |
| 7   | 12   | 1   | 0    | 1    | 2    | 9.95   | 0.67     | 1.7599 | $\pi$     | $\pi^*$    |
| 7   | 12   | 1   | 2    | 1    | 0    | 9.95   | 0.67     | 0.0000 | 0         | $-\pi$    |
| 7   | 12   | 1   | 1    | 2    | 1    | 9.96   | 0.90     | 0.0000 | 0         | $-\pi$    |

| $h$ | $h'$ | $k$ | $k'$ | $l$ | $l'$ | $Q_||$ | $Q_\perp$ | $I$  | $\varphi$ | $\varphi'$ |
|-----|------|-----|------|-----|------|--------|-----------|-----|-----------|-----------|
| 6   | 12   | 0   | 0    | 2    | 4    | 9.96   | 0.90     | 0.0222 | $-\pi$    |           |
| 7   | 11   | 2   | 1    | 3    | 4    | 9.96   | 0.90     | 0.0000 | $\pi$     |           |
| 6   | 12   | 2   | 4    | 0    | 0    | 9.96   | 0.90     | 0.0475 | $-\pi$    |           |
| 8   | 11   | 1   | 2    | 1    | 3    | 9.96   | 0.90     | 0.0090 | $\pi$     | $-\pi^*$    |
| 7   | 11   | 0   | 2    | 2    | 5    | 10.03  | 1.00     | 0.0013 | $\pi$     |           |
| 8   | 11   | 3    | 3    | 0    | 0    | 10.03  | 1.00     | 0.0256 | $\pi$     |           |
| 8   | 11   | 1    | 3    | 2    | 2    | 10.03  | 1.00     | 0.0764 | $-\pi$    |           |
| 7   | 12   | 0   | 1    | 2    | 2    | 10.03  | 0.64     | 0.4560 | $\pi$     | $\pi^*$    |
| 6   | 11   | 0   | 0    | 3    | 6    | 10.03  | 0.64     | 0.8459 | $\pi$     |           |
| 7   | 12   | 0   | 0    | 1    | 3    | 10.06  | 0.57     | 0.9217 | 0         |           |
| 7   | 12   | 2   | 2    | 1    | 1    | 10.06  | 0.57     | 0.0000 | $\pi$     | $-\pi^*$    |
| 7   | 12   | 1   | 2    | 1    | 2    | 10.07  | 0.32     | 0.2547 | 0         | $0^*$      |
| 7   | 12   | 1   | 3    | 0    | 1    | 10.08  | 0.68     | 0.0994 | $\pi$     | $\pi^*$    |
| 7   | 11   | -1  | 1    | 3    | 5    | 10.10  | 0.98     | 0.0286 | $\pi$     | $-\pi^*$    |
| 7   | 12   | 2    | 1    | 2    | 2    | 10.10  | 0.98     | 0.0388 | 0         | $-\pi^*$    |
| 8   | 11   | 2    | 4    | 1    | 0    | 10.10  | 0.98     | 0.0000 | $-\pi^*$    |
| 6   | 12   | 1    | 1    | 3    | 4    | 10.13  | 0.94     | 0.0000 | $-\pi^*$    |
| 7   | 10   | 0    | 0    | 3    | 7    | 10.13  | 0.94     | 0.4737 | $\pi$     |           |
| 7   | 12   | 0    | 2    | 1    | 3    | 10.13  | 0.94     | 0.0000 | 0         | $-\pi^*$    |
| 8   | 11   | 1    | 1    | 2    | 4    | 10.14  | 0.81     | 0.1516 | $\pi$     | $\pi^*$    |
| 7   | 11   | 1    | 1    | 3    | 5    | 10.14  | 0.26     | 6.0460 | 0         |           |
| 7   | 12   | 2    | 3    | 0    | 0    | 10.14  | 0.26     | 9.7305 | 0         |           |
| 7   | 12   | 1    | 1    | 2    | 3    | 10.19  | 0.35     | 7.8875 | 0         | $0^*$      |
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

| $h \, h'$ | $k \, k'$ | $l \, l'$ | $Q_{||}$ | $Q_{\perp}$ | $I$ | $\varphi$ | $\varphi'$ |
|------------|-----------|-----------|----------|-----------|-----|---------|---------|
| 8 12 0    | 0 0 0 0   | 10.19 0.35 | 8.5450 0 0 |
| 8 12 1    | 0 0 1 1   | 10.22 0.79 | 0.2147 $\pi$ $\pi$ * |
| 8 11 0    | 0 3 4 4   | 10.22 0.79 | 0.1390 $\pi$ 0 |
| 7 12 2    | 3 0 2 2   | 10.22 0.79 | 0.0939 0 $\pi$ * |
| 7 12 0    | 0 3 3 3   | 10.24 0.73 | 0.1007 0 $-$ |
| 7 12 2    | 2 1 3 3   | 10.24 0.73 | 0.0646 0 $\pi$ * |
| 8 12 1    | 1 1 1 0   | 10.24 0.73 | 0.3296 $\pi$ $\pi$ * |
| 7 11 0    | 0 4 5 5   | 10.26 0.56 | 0.4561 $\pi$ $\pi$ |
| 8 12 0    | 1 1 1 1   | 10.26 0.56 | 0.4051 $\pi$ $\pi$ * |
| 8 12 0    | 0 0 0 2   | 10.26 0.82 | 0.0000 0 $-$ |
| 8 12 0    | 2 0 0 0   | 10.26 0.82 | 0.0017 $\pi$ $-$ |
| 7 12 3    | 3 3 0 1   | 10.26 0.82 | 0.0000 $-$ $-$ * |
| 7 12 1    | 2 1 4 4   | 10.33 0.93 | 0.0500 $\pi$ $-$ * |
| 8 12 2    | 1 1 1 1   | 10.33 0.93 | 0.0161 $\pi$ $-$ * |
| 7 12 1    | 4 1 2 2   | 10.33 0.93 | 0.1535 $\pi$ 0 |
| 6 12 0    | 1 3 5 5   | 10.33 0.93 | 0.0110 $-$ $-$ |
| 7 12 0    | 1 2 4 4   | 10.33 0.53 | 0.9677 $\pi$ $\pi$ * |
| 8 12 1    | 2 0 1 1   | 10.33 0.53 | 0.2433 $\pi$ $\pi$ * |
| 7 12 2    | 3 2 2 2   | 10.33 0.53 | 0.0000 0 $-$ |
| 8 12 1    | 1 1 2 2   | 10.36 0.44 | 1.0170 0 $\pi$ * |
| 7 12 2    | 4 1 1 1   | 10.36 0.44 | 1.0525 0 0 |
| 8 12 0    | 0 2 2 2   | 10.38 0.58 | 0.0000 $\pi$ $-$ |
| 7 11 0    | 1 3 6 6   | 10.38 0.58 | 0.0533 0 $-$ |

| $h \, h'$ | $k \, k'$ | $l \, l'$ | $Q_{||}$ | $Q_{\perp}$ | $I$ | $\varphi$ | $\varphi'$ |
|------------|-----------|-----------|----------|-----------|-----|---------|---------|
| 8 12 2    | 2 0 0 0   | 10.38 0.58 | 0.0000 0 $-$ |
| 7 12 1    | 3 2 3 3   | 10.38 0.58 | 0.1122 $\pi$ $-$ |
| 7 12 2    | 1 2 4 4   | 10.40 0.91 | 0.0000 $\pi$ $-$ * |
| 6 11 0    | 0 5 6 6   | 10.40 0.91 | 0.1934 0 $-$ |
| 8 12 1    | 2 2 1 1   | 10.40 0.91 | 0.0000 $-$ $-$ |
| 7 12 2    | 2 2 3 3   | 10.43 0.87 | 0.0000 0 $-$ |
| 8 12 1    | 3 1 1 0   | 10.43 0.87 | 0.0249 $-$ $-$ |
| 7 13 0    | 0 0 0 0   | 10.43 0.87 | 0.0457 $-$ $-$ |
| 7 13 0    | 0 0 0 1   | 10.44 0.73 | 0.3990 0 $-$ |
| 7 12 3    | 4 1 0 1   | 10.44 0.73 | 0.1120 $\pi$ $-$ |
| 8 12 0    | 1 1 1 3   | 10.44 0.73 | 0.7736 $\pi$ 0 $*$ |
| 8 12 1    | 0 3 4 4   | 10.44 0.73 | 0.7330 $\pi$ 0 $*$ |
| 8 11 1    | 0 3 5 5   | 10.45 0.94 | 0.0000 0 $-$ $*$ |
| 8 12 2    | 2 2 0 2   | 10.45 0.94 | 0.0000 $\pi$ $-$ $*$ |
| 8 12 0    | 2 2 2 2   | 10.45 0.94 | 0.0000 $-$ $-$ |
| 7 13 0    | 1 1 0 0   | 10.45 0.94 | 0.0000 $-$ $-$ |
| 7 13 1    | 1 1 1 1   | 10.51 0.70 | 0.3928 0 $-$ |
| 8 12 1    | 0 2 3 3   | 10.51 0.70 | 0.7422 $\pi$ 0 $*$ |
| 7 12 2    | 5 0 0 0   | 10.51 0.70 | 0.9097 $\pi$ 0 |
| 7 13 1    | 2 0 0 0   | 10.54 0.64 | 0.0131 $-$ $-$ |
| 7 11 1    | 0 4 6 6   | 10.54 0.64 | 0.0030 $\pi$ $-$ |
| 8 12 1    | 3 1 2 2   | 10.54 0.64 | 0.0898 0 0 $*$ |
| 7 12 1    | 2 3 4 4   | 10.55 0.43 | 0.3418 0 0 |
Table A.1: Retrieved phases for the AlMnPd quasicrystal (cont.)

| $h$ | $h'$ | $k$ | $k'$ | $l$ | $l'$ | $Q_{||}$ | $Q_{\perp}$ | $I$ | $\varphi$ | $\varphi'$ |
|-----|-----|-----|-----|-----|-----|---------|---------|-----|---------|---------|
| 8   | 12  | 2   | 3   | 1   | 1   | 10.55   | 0.43    | 0.2238 | $\pi$   | 0 - *   |
| 7   | 13  | 0   | 1   | 1   | 2   | 10.56   | 0.74    | 0.0548 | - -     |         |
| 7   | 12  | 1   | 1   | 2   | 5   | 10.56   | 0.74    | 0.0409 | 0 - *   |         |
| 8   | 12  | 2   | 2   | 2   | 2   | 10.56   | 0.74    | 0.0000 | $\pi$   | - *     |
| 7   | 13  | 1   | 2   | 0   | 2   | 10.61   | 0.98    | 0.0000 | - -     |         |
| 8   | 12  | 3   | 3   | 1   | 0   | 10.61   | 0.98    | 0.0000 | $\pi$   | - *     |
| 7   | 13  | 1   | 0   | 2   | 2   | 10.61   | 0.98    | 0.0000 | - -     |         |
| 8   | 11  | 1   | 4   | 5   | 10.61 | 0.98    | 0.0046 | -   |         |         |
| 8   | 12  | 1   | 1   | 1   | 4   | 10.61   | 0.98    | 0.0202 | 0 - *   |         |
| 9   | 12  | 0   | 0   | 1   | 1   | 10.61   | 0.98    | 0.1242 | $\pi$   | $\pi$   |
| 7   | 13  | 2   | 2   | 0   | 1   | 10.62   | 0.86    | 0.0917 | - $\pi$ |         |
| 8   | 12  | 0   | 1   | 3   | 3   | 10.62   | 0.86    | 0.1801 | - $\pi$ |         |
| 8   | 12  | 1   | 2   | 3   | 3   | 10.62   | 0.39    | 0.4691 | 0 0 - * |         |
| 7   | 12  | 0   | 0   | 3   | 5   | 10.65   | 0.26    | 5.1924 | 0 0     |         |
| 8   | 12  | 0   | 0   | 2   | 4   | 10.67   | 0.45    | 1.8886 | 0 0     |         |
| 8   | 12  | 2   | 4   | 0   | 0   | 10.67   | 0.45    | 1.8399 | 0 $\pi$ |         |
| 7   | 13  | 1   | 1   | 2   | 3   | 10.69   | 0.84    | 0.0987 | 0 - 0   |         |
| 8   | 11  | 0   | 0   | 3   | 6   | 10.69   | 0.84    | 0.0995 | $\pi$   | -       |
| 7   | 13  | 1   | 3   | 1   | 1   | 10.69   | 0.84    | 0.1403 | - $\pi$ |         |
| 7   | 12  | 0   | 2   | 3   | 5   | 10.72   | 0.79    | 0.3223 | $\pi$   | 0       |
| 8   | 12  | 3   | 3   | 1   | 2   | 10.72   | 0.79    | 0.0017 | 0 - 0   |         |
| 7   | 13  | 1   | 2   | 2   | 2   | 10.72   | 0.79    | 0.0000 | - -     |         |
| 7   | 13  | 0   | 0   | 2   | 3   | 10.73   | 0.63    | 0.8304 | - 0     |         |
Appendix B

The effect of averaging on the resolution of a reconstruction

We consider here the effect of the three translational symmetries of the specimen on the averaging method. The first calculation shows that the average over an ensemble of translated estimates in a plane perpendicular to the propagation direction can be modeled as a gaussian low-pass filtering operation. The second calculation shows how the resolution can be further affected by an averaging over many defocus planes. This defocus occurs only as a combined effect of the noise and complex-valuedness of the reconstruction.

B.1 Averaging over translation in the transverse plane

We assume for simplicity that the distance between the support and the boundary of the object is the same along its contour. Let \( r_0 \) be this distance, expressing the tightness of the support. The averaging is made over the family of estimates having Fourier components that differ by a phase tilt \( e^{i\mathbf{r} \cdot \mathbf{q}} \), with \( |\mathbf{r}| < r_0 \). Then, assuming that the distribution of \( \mathbf{r} \) is a gaussian of standard deviation \( r_0 \), we find that

\[
\left\langle \tilde{\Psi}_q e^{i\mathbf{r} \cdot \mathbf{q}} \right\rangle = \tilde{\Psi}_q \left\langle e^{i\mathbf{r} \cdot \mathbf{q}} \right\rangle = \tilde{\Psi}_q e^{-\frac{1}{2} r_0^2 \mathbf{q}^2}.
\] (B.1)

This results in an effective gaussian transfer function.

201
B.2 Averaging over defocus planes

We model one slice of the specimen being reconstructed as a uniform disk of radius $R$. The goal of the first part of this calculation is to find the approximate amplitude of the scattered field at distance $z$ downstream from the disk and a distance $r$ from the axis passing through its center. The Fourier transform of the wavefield at the plane $z$ is given by

$$\tilde{\Psi}_q(z) = \tilde{\Psi}_q(0)e^{iz\left(\sqrt{k^2-q^2}-k\right)},$$

(B.2)

where $\tilde{\Psi}_q(0)$ is the Fourier transform of the exit wave at $z = 0$. Since the specimen is assumed to be a uniform disk,

$$\Psi(r, z = 0) = \begin{cases} 
\Psi_0 & \text{if } |r| < R \\
0 & \text{otherwise.}
\end{cases}$$

(B.3)

Then $\tilde{\Psi}_q(0)$ is just an Airy disk:

$$\tilde{\Psi}_q(0) = \Psi_0 R^2 J_1(qR) \frac{q R}{q R}$$

(B.4)

The inverse Fourier transform of (B.2) can be written as:

$$\Psi(r, z) = \frac{1}{2\pi} \int \tilde{\Psi}_q(z) e^{ir \cdot q} d^2 q$$

$$\approx \frac{1}{2\pi} \int \left[ \Psi_0 R^2 J_1(Rq) \frac{q R}{q R} e^{-\frac{1}{2}izq^2/k} \right] e^{ir \cdot q} d^2 q$$

$$= \Psi_0 k R \int_0^\infty J_0(k r x) J_1(k R x) e^{-\frac{1}{2}ikzx^2} dx.$$  

(B.5)

Using the method of stationary phase, we find that, when $R \gg (r - R)$ and $z \gg (r - R)$,

$$\Psi(r, z) \sim \Psi_0 \frac{z}{k(r - R)^2},$$  

(B.6)

which is also, as it turns out, the asymptotic behavior of the Fresnel integral for a semi-infinite plane (that is, Fresnel diffraction of a plane wave by a straight edge).
As explained in Section 4.1.2, in the presence of noise in the Fourier data, the reconstruction algorithm enters in a steady-state regime. In the absence of convergence, the reconstruction is defined as an average over the Fourier estimates. Because of noise, the region outside the support in these estimates is never zero but fluctuates with an rms value $\varepsilon$ which scales like the difference map error. $\varepsilon$ can than be seen as the tolerance of the algorithm to fluctuations outside the support. The goal of this Appendix is to point out that this finite tolerance allows in turn a range of defocus values, as long as $\varepsilon > |\Psi/\Psi_0|$, that is:

$$\varepsilon > |\Psi/\Psi_0| \sim \frac{z}{k(R_S - R)^2}, \quad (B.7)$$

where we have set $r = R_S$, the radius of the circular support used in a reconstruction. It is then reasonable to assume that the extent of the defocus region $|z| < z_0$ tolerated by the algorithm is given by $z_0 \sim \varepsilon k(R_S - R)^2$.

The last step in this calculation consists in computing the effect of averaging on the ensemble of defocus planes. Taking $z_0$ as the standard deviation of the distribution of planes, we have

$$\langle \Psi_q e^{-\frac{1}{2}izq^2/k} \rangle = \tilde{\Psi}_q \left\langle e^{-\frac{1}{2}izq^2/k} \right\rangle$$

$$= \tilde{\Psi}_q e^{-\frac{1}{2}z(\varepsilon q^2/k)^2}$$

$$= \tilde{\Psi}_q e^{-\alpha e^2(R_S - R)^4 q^4}. \quad (B.8)$$

The averaging over defocus planes results in a low-pass filter both dependent on the noise (through $\varepsilon$) and the tightness of the support, $|R_S - R|$. 


S. Gravel. Personal communication. 2007.


<table>
<thead>
<tr>
<th>Reference</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allen and Oxley (2001)</td>
<td>42, 204</td>
</tr>
<tr>
<td>Allen et al. (2001)</td>
<td>42, 58, 71, 204</td>
</tr>
<tr>
<td>Ashcroft and Mermin (1976)</td>
<td>99, 204</td>
</tr>
<tr>
<td>Bates (1982)</td>
<td>38, 204</td>
</tr>
<tr>
<td>Bauschke and Borwein (1996)</td>
<td>45, 56, 204</td>
</tr>
<tr>
<td>Bauschke et al. (2002)</td>
<td>58, 204</td>
</tr>
<tr>
<td>Bauschke et al. (2003)</td>
<td>61, 204</td>
</tr>
<tr>
<td>Bauschke et al. (2004)</td>
<td>56, 204</td>
</tr>
<tr>
<td>Beetz and Jacobsen (2003)</td>
<td>93, 204</td>
</tr>
<tr>
<td>Beetz et al. (2005)</td>
<td>129, 204</td>
</tr>
<tr>
<td>Blow and Rossmann (1961)</td>
<td>182, 204</td>
</tr>
<tr>
<td>Born and Wolf (1999)</td>
<td>10, 14, 17, 22, 146, 204</td>
</tr>
<tr>
<td>Boudard et al. (1992)</td>
<td>160, 187, 205</td>
</tr>
<tr>
<td>Brames (1987)</td>
<td>109, 205</td>
</tr>
<tr>
<td>Brown et al. (2000)</td>
<td>159, 165, 187, 205</td>
</tr>
<tr>
<td>Brown Jr (1966)</td>
<td>24, 205</td>
</tr>
<tr>
<td>Bruck and Sodin (1979)</td>
<td>38, 205</td>
</tr>
<tr>
<td>Brégnan (1965)</td>
<td>58, 205</td>
</tr>
<tr>
<td>Cahn et al. (1986)</td>
<td>187, 205</td>
</tr>
<tr>
<td>Carrier (2002)</td>
<td>49, 205</td>
</tr>
<tr>
<td>Chao et al. (2005)</td>
<td>129, 205</td>
</tr>
<tr>
<td>Chapman et al. (2006a)</td>
<td>96, 111, 186, 205</td>
</tr>
<tr>
<td>Chapman et al. (2006b)</td>
<td>2, 89, 206</td>
</tr>
<tr>
<td>Cimmino (1938)</td>
<td>57, 206</td>
</tr>
<tr>
<td>Coene et al. (1996)</td>
<td>42, 206</td>
</tr>
<tr>
<td>Cooley and Tukey (1965)</td>
<td>9, 206</td>
</tr>
<tr>
<td>Cowley and Moodie (1957)</td>
<td>26, 206</td>
</tr>
<tr>
<td>Crimmins et al. (1990)</td>
<td>109, 110, 206</td>
</tr>
<tr>
<td>Devaney (1986)</td>
<td>24, 206</td>
</tr>
<tr>
<td>Eisebitt et al. (2004)</td>
<td>38, 206</td>
</tr>
<tr>
<td>Elder et al. (1947)</td>
<td>1, 206</td>
</tr>
<tr>
<td>Elser and Millane (2007)</td>
<td>35, 207</td>
</tr>
<tr>
<td>Elser (2003a)</td>
<td>47, 50, 79, 206</td>
</tr>
<tr>
<td>Elser (2003b)</td>
<td>77, 206</td>
</tr>
<tr>
<td>Elser (2003c)</td>
<td>50, 207</td>
</tr>
<tr>
<td>Elser (2005)</td>
<td>109, 207</td>
</tr>
<tr>
<td>Ewald (1913)</td>
<td>13, 207</td>
</tr>
<tr>
<td>Ewald (1940)</td>
<td>117, 207</td>
</tr>
<tr>
<td>Faulkner and Rodenburg (2004)</td>
<td>42, 175, 207</td>
</tr>
<tr>
<td>Feit and J. A. Fleck (1978)</td>
<td>26, 207</td>
</tr>
<tr>
<td>Fienup et al. (1982)</td>
<td>109, 207</td>
</tr>
<tr>
<td>Fienup (1978)</td>
<td>2, 58, 60, 73, 116, 207</td>
</tr>
<tr>
<td>Fienup (1982)</td>
<td>2, 60, 61, 116, 207</td>
</tr>
<tr>
<td>Fienup (1987)</td>
<td>111, 115, 207</td>
</tr>
<tr>
<td>Flåm and Zowe (1990)</td>
<td>56, 57, 207</td>
</tr>
<tr>
<td>Gabor (1948)</td>
<td>38, 207</td>
</tr>
<tr>
<td>Gabor (1975)</td>
<td>101, 207</td>
</tr>
<tr>
<td>Gerchberg and Saxton (1972)</td>
<td>2, 58–60, 208</td>
</tr>
<tr>
<td>Glauber (1963a)</td>
<td>83, 208</td>
</tr>
<tr>
<td>Glauber (1963b)</td>
<td>83, 208</td>
</tr>
<tr>
<td>Gori and Guattari (1973)</td>
<td>101, 208</td>
</tr>
<tr>
<td>Gravel (2007)</td>
<td>55, 208</td>
</tr>
<tr>
<td>Green et al. (1954)</td>
<td>182, 208</td>
</tr>
<tr>
<td>Grimm et al. (1998)</td>
<td>129, 208</td>
</tr>
<tr>
<td>Hauptman and Karle (1953)</td>
<td>59, 208</td>
</tr>
<tr>
<td>He et al. (2003)</td>
<td>2, 208</td>
</tr>
<tr>
<td>Hendrickson (1991)</td>
<td>183, 208</td>
</tr>
<tr>
<td>Henke et al. (1993)</td>
<td>168, 208</td>
</tr>
<tr>
<td>Hippert and Gratias (1994)</td>
<td>158, 208</td>
</tr>
<tr>
<td>Hoppe (1969)</td>
<td>127, 208</td>
</tr>
<tr>
<td>Howells et al. (2000)</td>
<td>38, 209</td>
</tr>
</tbody>
</table>
Howells et al. (2005), 92, 93, 209
Ibrahim Sezan (1992), 45, 209
Iijima (1977a), 127, 209
Iijima (1977b), 127, 209
Jackson (1975), 10, 209
Jacobsen (2005), 97, 209
Kaczmarz (1937), 58, 209
Karle and Hauptman (1950), 116, 209
Lagrange (1813), 43, 209
Lane (1987), 111, 209
Lee et al. (1996), 160, 166, 209
Le Gros et al. (2005), 129, 209
Lin and Fiddy (1992), 24, 210
Lin et al. (2003), 90, 210
Loh et al. (2007), 79, 210
Maleki and Devaney (1994), 42, 210
Mandel and Wolf (1965), 85, 86, 210
Marchesini et al. (2003), 76, 77, 111, 210
Marchesini et al. (2005), 117, 210
Marchesini (2004), 62, 210
McBride et al. (2004), 116, 210
McBride et al. (2005), 42, 181, 210
Mentes et al. (2002), 42, 210
Miao and Sayre (2000), 33, 78, 211
Miao et al. (1999), 2, 211
Miao et al. (2002), 2, 211
Miao et al. (2003), 2, 211
Millane (1990), 37, 59, 111, 211
Misell (1973), 41, 211
Nellist et al. (1995), 127, 211
Neutze et al. (2000), 186, 211
Nugent et al. (2003), 42, 58, 181, 211
Nugent et al. (2005), 42, 181, 211
Oristaglio (1985), 24, 211
Oszlányi and Sütő (2004), 62, 75, 212
Patterson (1939), 36, 117, 212
Peele et al. (2004), 49, 212
Pfeifer et al. (2006), 117, 121, 212
Pfeiffer et al. (2005), 90, 212
Plamann and Rodenburg (1998), 127, 212
Ravelli et al. (2003), 183, 212
Rayleigh (1900), 99, 212
Read and Kleywegt (2001), 58, 212
Robinson and Vartanyants (2001), 119, 212
Rodenburg et al. (2007a), 42, 143, 172, 175, 177, 178, 212
Rodenburg et al. (2007b), 42, 175, 212
Rose (1948), 93, 213
Russel Luke (2005), 51, 61, 213
Rytov (1937), 22, 213
Saxton (1978), 41, 213
Sayre and Chapman (1995), 92, 213
Sayre (1952a), 2, 31, 213
Sayre (1952b), 59, 116, 213
Sayre (2005), 30, 213
Schiske (1973), 41, 213
Shannon (1949), 2, 31, 213
Shapiro et al. (2005), 2, 128, 130, 140, 142, 213
Shen et al. (2004), 92, 93, 214
Shiono and Wolfson (1992), 58, 75, 214
Slepian and Pollack (1961), 101, 214
Spence et al. (2002), 111, 143, 214
Spence (2003), 12, 214
Stroke and Falconer (1964), 38, 214
Takakura et al. (2001), 159, 165, 214
Takakura et al. (2007), 159, 214
Thibault and Rankenburg (2007), 144,
algorithm
  charge flipping, 62
difference map, see difference map
error-reduction, 59
Gerchberg-Saxton, 59
hybrid input-output, 59
RAAR, 61
autocorrelation
defined, 8
averaging, 52, 94
beam propagation method, 26
Born approximation, 18–22
  relation with Ewald sphere, 21
coherence, 82–90
  spatial, 84
  temporal, 86
coherent diffractive imaging, 2
complex-valued
  colorwheel, 5
  reconstructions, 111
constraint ratio, 35
constraint sets, 44–45
  projection onto, 45–46
convergence, 47
convex feasibility problem, 44
convexity, 44
convolution, 8
cross-correlation, 8
difference map, 47–56
  convergence, 49
  error, 48
  estimates, 48
  fixed point, 48
  parameters, 50
diffraction microscopy, 30–38
diffractive imaging methods, 1
direct space
  defined, 7
eikonal approximation, 18
Ewald sphere, 13
Far-field diffraction
see Fraunhofer diffraction, 14
fixed point, 47
focal series, 41
Fourier holography, 38
Fourier space
defined, 7
Fourier transform
defined, 7
discrete, 9
Fraunhofer diffraction, 14
Fresnel number, 17
Fresnel propagator, 16
Friedel law, 7
Helmholtz equation, 12
Green function, 19
index of refraction, 12, 13
missing data, 96
multislice formula, 27
Nyquist frequency, 9
optical thickness, 18
overdetermination, see constraint ratio
oversampling, 10, 31–34
ratio, 33
paraxial approximation, 16
Parseval’s theorem, 7
phase problem
defined, 29
uniqueness, 37
phase ramp, 7
projection
average, 55
defined, see constraint sets
direct product, 55
modulus, 64–73
non-optimal, 54
pseudo-, 54
relaxed, 57
support, 74–77
value, 77–80
projection approximation, 18
radiation damage, 92
refractive index, see index of refraction
replica, 55
Rytov approximation, 22–24
sampling theorem, 2
speckles, see oversampling
support, 108
defined, 32
uniqueness, see phase problem
wave equation, see Helmholtz equation