Cambridge University Press
978-1-316-51887-8 — Geometry of the Phase Retrieval Problem
Alexander H. Barnett , Charles L. Epstein , Leslie Greengard , Jeremy Magland
Excerpt
More Information

1

Introduction

In a wide range of applications, including crystallography, astronomy, coherent diffraction imaging (CDI), and ptychography, physical measurements are made that can be interpreted as samples of the magnitude of the Fourier transform, $\{|\widehat{\rho}(\boldsymbol{\xi})| : \boldsymbol{\xi} \in \Xi\}$, of an X-ray scattering density $\rho(\boldsymbol{x})$.[1] Here, Ξ denotes a finite discrete set of sample locations. Given this Fourier magnitude data, the problem of interest is the recovery of $\rho(\boldsymbol{x})$ itself or, more precisely, a bandlimited version of ρ sampled on a regular grid. This recovery entails first determining the phases of $\{\widehat{\rho}(\boldsymbol{\xi}) : \boldsymbol{\xi} \in \Xi\}$. Hence, this is called the *phase retrieval problem*. It is clear that the problem, as stated, is not solvable without some additional constraints on ρ, which we call *auxiliary information*. In CDI, it is typically known that $\rho(\boldsymbol{x})$ is supported in a compact set $S \subset \mathbb{R}^d$. As discussed in the Preface, in applications the space dimension d is typically 2 or 3, yet much of our analysis will apply to general d.

In this book, we study a discrete model of classical, phase retrieval, which, in the limit of high resolution, tends to the continuum problem. The discrete analysis brings many advantages mathematically (see Section 1.5) and allows clean numerical experiments (involving a single discrete Fourier transform, or DFT) that avoid complications of data interpolation or regridding. Yet the model preserves most of the interesting pathologies of the continuum case and is highly relevant because almost all numerical solution schemes ultimately rely on DFTs.

The model centers around a DFT of size $N_1 \times \cdots \times N_d$, i.e., a d-dimensional rectangular (cuboid) grid. More abstractly this grid will be denoted by $J \subset \mathbb{Z}^d$, a subset of the integer lattice, with vector indices $\boldsymbol{j} = (j_1, \ldots, j_d) \in J$. In the

[1] At the energies we consider in this book, the function $\rho(\boldsymbol{x})$ is connected to the refractive index, $n(\boldsymbol{x})$, of the material being imaged, by the relation $\rho(\boldsymbol{x}) = k^2(1 - n^2(\boldsymbol{x}))$. At X-ray energies $n(\boldsymbol{x}) = 1 - \delta(\boldsymbol{x}) - i\beta(\boldsymbol{x})$, with $|\delta(\boldsymbol{x}) + i\beta(\boldsymbol{x})| \ll 1$ (see Jacobsen, 2019, §3.3).

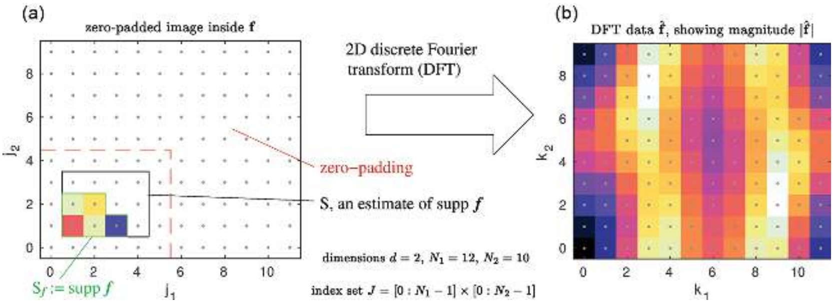

Figure 1.1 Setup for the discrete model analyzed in this book. A 2D example of DFT size 12×10 is shown, with index set J with $|J| = 120$. Part (a) is the image data $\boldsymbol{f} \in \mathbb{R}^J$, and part (b) is the resulting DFT data $\widehat{\boldsymbol{f}}$ which lives on the same index set J. The $\boldsymbol{k} = (0,0)$ index corresponds to zero spatial frequency. The constraint set S is a set containing the true support $S_{\boldsymbol{f}}$, which, in this case, comprises only five pixels.

real case, the DFT (see (1.7)) takes an image vector $\boldsymbol{f} \in \mathbb{R}^J$, meaning a grid of real values living on the rectangle J, to its Fourier coefficient vector $\widehat{\boldsymbol{f}} \in \mathbb{R}^J$. The "measured data" is then the set of magnitudes $|\widehat{f}(\boldsymbol{k})|$ for $\boldsymbol{k} \in J$, where $\boldsymbol{k} = (k_1, \ldots, k_d)$ is a frequency index. This data comprises $N_1 \ldots N_d$ values, although in the real case, by symmetry, there are about half this number of independent values.

For reasons we will see shortly, for phase retrieval to stand a chance of success, the *unknown image* must live inside a rectangle at most *half* the size of J in each direction; the rest of the $\boldsymbol{f}$ vector contains zeros. For example, in the 2D case, the unknown image must live in a subrectangle of at most $N_1/2$ by $N_2/2$ pixels within J with the remaining pixel values zeros. An example is shown in Figure 1.1, where the nonzero pixels lie in the bottom-left region. In the physics literature this is often referred to as "oversampling" the Fourier coefficients of the image, but in fact is simply a matter of collecting Fourier data that contains information about support of the image within the field of view.

Unless stated, we allow J to be an arbitrary finite rectangular subset of a regular grid. The correspondence with the continuous problem it is easiest to describe in the "square" special case: $N_1 = \cdots = N_d = N$, where N sets the size of the $\boldsymbol{f}$ array. In this case, one may think of $\boldsymbol{f}$ as having elements

$$f_{\boldsymbol{j}} = \rho\left(\frac{\boldsymbol{j}}{N}\right) \quad \text{for } \boldsymbol{j} \in J, \tag{1.1}$$

where $\boldsymbol{j}/N$ means the vector $(j_1/N, \dots, j_d/N)$ and $\rho : \mathbb{R}^d \to \mathbb{R}$ is a continuous intensity function that vanishes outside of $[0, 1/2]^d$. This last condition ensures that the resulting $\boldsymbol{f}$ has support in a rectangular subset with sidelengths at most $N/2$. Then, fixing ρ and a nonnegative index vector $\boldsymbol{k}$, the DFT coefficient $\hat{f}_{\boldsymbol{k}}$ of $\boldsymbol{f}$ tends (disregarding a constant prefactor) as $N \to \infty$ to the Fourier transform $\hat{\rho}(\boldsymbol{k}) := \mathscr{F}(\rho)(\boldsymbol{k})$, using the definition (1). This is true because, when properly scaled, the DFT is a Riemann sum for the Fourier integral. This is discussed in greater detail in Section 1.5. In other words, the data in the discrete problem approximates the Fourier transform magnitudes for wavevectors Ξ on the integer lattice. (Note that the negative indices are to be found periodically "wrapped" into the other corners of the set J, as usual for the DFT, as shown in Figure 1.1.)

In the continuum case, the magnitude Fourier data is unchanged if $\rho(\boldsymbol{x})$ is replaced by a translate $\rho(\boldsymbol{x} - \boldsymbol{v})$, for a fixed vector $\boldsymbol{v}$, or by the inversion $\rho(-\boldsymbol{x})$, if ρ is real valued; for complex valued ρ, this operation is replaced by $\overline{\rho(-\boldsymbol{x})}$. These are called "trivial associates" of ρ. There is a similar phenomenon in the discrete problem. A more detailed discussion of the relationship between these two problems is given in Section 1.5.

We call the problem of reconstructing $\boldsymbol{f}$ from a knowledge of $|\widehat{\boldsymbol{f}}|$, plus some auxiliary information, the discrete, *classical*, phase retrieval problem, as the data are samples, $\{|\widehat{f}(\boldsymbol{k})|\}$, of the moduli of the Fourier transform itself, and what is sought are the phases of the complex numbers $\{\widehat{f}(\boldsymbol{k})\}$. For definiteness we define the *phase* of a nonzero complex number, z, to be the point on the unit circle in $\mathbb{C}$ given by

$$e^{i\theta} = \frac{z}{|z|}. \tag{1.2}$$

The phase of 0 is not defined, and the phase, as an S^1-valued function on $\mathbb{C} \setminus \{0\}$, cannot be extended continuously to $z = 0$. At times we are sloppy and refer to the *angle* θ as the phase; it is only defined up to the addition of multiples of 2π.

This problem is a member of a more general class of inverse problems where similar algorithms are used, generically referred to as "phase retrieval problems." Recently a great deal of effort has been directed toward the development and analysis of algorithms for solving these generalized problems (see Candès et al. 2015a, 2015b; Cahill et al. 2016; Alaifari et al. 2019). In this larger class of problems, the samples of the magnitude of the DFT data and information about its support are replaced with collection of absolute values of inner products

$$\mathfrak{M} = \{|\langle \boldsymbol{f}, \boldsymbol{m}_l \rangle| : l = 1, \dots, L\}. \tag{1.3}$$

Here, $\boldsymbol{f} = (f(\boldsymbol{x}_1), \ldots, f(\boldsymbol{x}_K))^t$ is the unknown sample vector and $\{\boldsymbol{m}_l : l = 1, \ldots, L\}$ is a collection of "measurement vectors," which define a frame for an underlying Hilbert space, $\mathscr{H}$, of dimension $K < \infty$.

It has been shown that if $\mathscr{H}$ is complex and $L \geq 4K - 4$, or if $\mathscr{H}$ is real and $L > 2K$, then, for a generic set of measurement vectors, the measurements $\mathfrak{M}$ *uniquely* determine $\boldsymbol{f}$ (see Conca et al. 2015; Marchesini et al. 2016). We do not consider such generalized problems but restrict our attention to the classical phase retrieval problem, as its solution is required in many applications. While that solution is not unique, the nonuniqueness is (under broad conditions) a consequence of simple translation invariance or inversion symmetry. The well-posedness of the problem, however, is more subtle to analyze and there are many properties of the Fourier transform that lead to subtle geometric features in the problem, which can be both analyzed and probed numerically.

Remark 1.1 In current experimental practice, there has been a substantial shift of interest from CDI to ptychography (Rodenburg et al. 2007; Dierolf et al. 2010; Pfeiffer 2018; Thibault et al. 2008). Rather than using a single plane wave to scatter from the entire sample, one selectively illuminates a small subregion using either a mask or a more focused beam. This introduces a more complicated forward model but yields vastly more data – rendering the inverse problem much better posed and more easily solved. Recent work has achieved remarkably high resolution in this manner (see, for example, Guizar-Sicairos et al. 2014). Unfortunately, this technique requires rastering across the sample, longer acquisition times, and more complicated experimental protocols. Each of these leads to interesting new algorithmic challenges but for the sake of simplicity, we restrict our attention in this book to classical CDI itself. There are many important physical applications where one is *constrained* to solve the classical problem, especially when the sample is involved in a dynamic process and the global structure is desired over the smallest possible time window. Bragg CDI, which is used to study strain in crystals, is another inverse problem of the same general type as CDI (see Jacobsen, 2019, §10.3.8). We believe that the mathematical analysis presented here could also play a role in a deeper understanding of ptychography.

As in the continuum case, it is quite clear that magnitude DFT data alone does not suffice to determine any object, even up to translations and inversion, and therefore some auxiliary information is needed. Fixing the magnitudes of the DFT data determines a torus, which we denote by $\mathbb{A}$. This torus is typically very high dimensional; in the Fourier representation it is a Cartesian product of round circles, lying in complex lines, one corresponding to the unknown

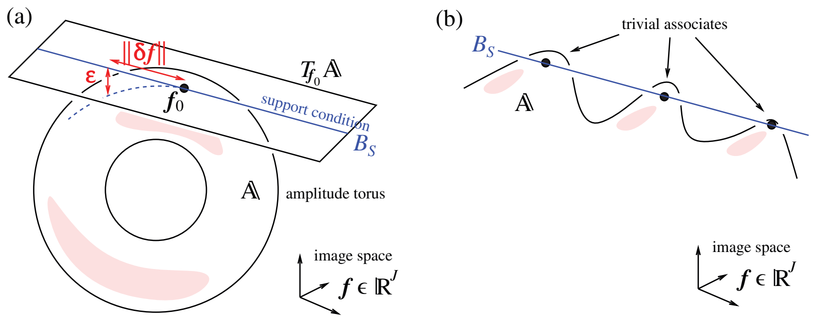

Figure 1.2 This figure illustrates (in a low dimensional cartoon) the full "image space" $\mathbb{R}^J$ in which f lives. In reality the ambient dimension, $|J|$, is thousands or more, so the 3D sketch misses many features. A valid image f_0 must be at an intersection of the torus $\mathbb{A}$ (defined by given Fourier magnitude data) and the support condition B_S (the subset here shown as a line). Part (a) shows a nontransversal such intersection where $B_S \cap T_{f_0}\mathbb{A} \neq \{f_0\}$. The distance between B_S and $\mathbb{A}$ grows as $\epsilon \propto (\delta f)^2$. Part (b) shows a collection of trivial associates (separate intersection points between the two sets).

phase of each magnitude DFT value measured. As they are intrinsically flat, these tori do *not* metrically resemble the familiar "donut" embedded in $\mathbb{R}^3$, as shown in Figure 1.2. The simplest such flat torus is the set

$$\{(r_1 e^{i\theta_1}, r_2 e^{i\theta_2}) : \theta_1, \theta_2 \in [0, 2\pi)\} \text{ with } r_1, r_2 > 0,$$

sitting in $\mathbb{C}^2$.

Functions that satisfy the auxiliary information belong to a different subset, denoted by B. Thus, phase retrieval problems can be understood as the problem of searching for the intersections between $\mathbb{A}$ and B. This point of view, championed by Bauschke, Borwein, Combettes, Elser, Fienup, Luke, and many others is the basis for most of the results in this book (see Fienup 1987; Bauschke and Borwein 1993, 1996; Bauschke et al. 2002; Elser 2003; Elser et al. 2007; Gravel and Elser 2008; Borwein and Sims 2011; Borwein 2012). The set B is often taken to be the linear subspace of objects with support in a given set. If the set B is chosen well, then the set $\mathbb{A} \cap B$ consists of finitely many points. The major theoretical contributions of this book are concerned with the analysis of the geometry of the two subsets *near* to points in $\mathbb{A} \cap B$. This geometry, in particular the transversality properties of these intersections, has a decisive influence on the intrinsic difficulty of this problem, and the behavior of any practical algorithm for solving it.

In the DFT representation, the fact that the set $\mathbb{A}$ is a product of round circles, lying in coordinate planes, is clear; for a real image it is therefore a smooth, *nonlinear, nonconvex* subset of a Euclidean space, $\mathbb{R}^N$. We let $T\mathbb{A}$ denote the tangent bundle of $\mathbb{A}$. For $\boldsymbol{f} \in \mathbb{A}$, the fiber of the tangent bundle at $\boldsymbol{f}$, $T_{\boldsymbol{f}}\mathbb{A}$, is naturally identified with the affine subspace of $\mathbb{R}^N$ that is the best linear approximation to $\mathbb{A}$ near to $\boldsymbol{f}$, with a similar definition for $T_{\boldsymbol{f}}B$. For a complex image, $\mathbb{A}$ is a subset of $\mathbb{C}^N$ and $T_{\boldsymbol{f}}\mathbb{A}$ is the real affine subspace of $\mathbb{C}^N$ that is the best linear approximation to $\mathbb{A}$.

The natural linearization of the problem of finding points in $\mathbb{A} \cap B$, is to find points in $T_{\boldsymbol{f}}\mathbb{A} \cap T_{\boldsymbol{f}}B$. This linearization is a good model for the original problem *if the intersection between* $\mathbb{A}$ *and* B *is transversal*: that is, $T_{\boldsymbol{f}}\mathbb{A} \cap T_{\boldsymbol{f}}B$ consists only of the point $\boldsymbol{f}$ itself. If the intersection is not transversal, then $T_{\boldsymbol{f}}\mathbb{A} \cap T_{\boldsymbol{f}}B$ is positive dimensional, which vastly complicates the analysis and dramatically affects the behavior of algorithms for finding these intersections. This latter fact is already evident in 2-dimensional examples.

For illustration, consider locating the zero of the function $y = x^p$, for $p \in \mathbb{N}$. This can be understood as finding the intersection between the sets

$$A_p = \{(x, y) : y = x^p\} \text{ and } B = \{(x, y) : y = 0\}.$$

If $p > 1$, then $T_{(0,0)}A_p = B$ and therefore this intersection is nontransversal. Newton's algorithm, with a starting point $0 < x_0$, gives iterates

$$x_n = \left(\frac{p-1}{p}\right)^n x_0 \text{ if } 1 < p. \tag{1.4}$$

As is well-known, these converge *geometrically*, with higher-order contact producing a slower rate of convergence. This should be compared with the result of using Newton's algorithm to find the roots of $y = x^2 - 1$, starting with an $x_0 \in (1, 2)$. In this case the intersection is transversal and the iterates, $x_{n+1} = \frac{1}{2}(x_n + x_n^{-1})$ satisfy

$$|x_n - 1| \leq \frac{|x_0 - 1|^{2^n}}{2^{2^{n-1}}}, \tag{1.5}$$

and therefore, converge super-geometrically. Stated differently, higher-order contact between A_p and B, leads to a weaker attraction toward the fixed points of the iterated map.

Moreover, higher-order contact (nontransversal intersection) causes a more fundamental issue: it renders the *problem* of finding points in $A_p \cap B$ ill conditioned. This is a property independent of any algorithm used to attempt its solution. Intuitively, any small uncertainty in the specification of the set A_p leads to a strongly amplified error in the intersection point. In fact, this amplification factor can be arbitrarily large, as we now illustrate with an example where A_p is the graph of $y = x^p$. Supposing that an error shifts this set to the graph of $y = x^p - \epsilon$, for some small $\epsilon > 0$, then an intersection with B occurs (at least) at $x = \epsilon^{1/p}$, rather than the true value $x = 0$. In this example the error in the specification of the data has been amplified by factor $\epsilon^{(1-p)/p}$, which is clearly unbounded as $\epsilon \to 0$. This is known as an infinite *absolute* condition number: formally, the problem is infinitely sensitive to its input data.

In practice, intersection points do not usually lie at the origin; for a more generic variant of the above example one can take for A_p the graph of $y = (x-a)^p + b$, and for B the line $y = b$, for nonzero constants a and b. If B is perturbed to the line $y = b - \epsilon$, then the intersection (a, b) is replaced by $(a + \epsilon^{\frac{1}{p}}, b - \epsilon)$. In general, it is the *relative* condition number that is of interest. This is the largest possible ratio between the the *relative* error in output and the *relative* error in data that specifies the problem. For the second example this would be $(\epsilon^{\frac{1}{p}}/a)/(\epsilon/b) = b\epsilon^{(1-p)/p}/a$, which is again unbounded as $\epsilon \to 0$. This is further discussed in Appendix 1.B. Note that having an infinite relative condition number does not prevent the problem from being solved accurately; it merely implies some loss of digits. For instance, in the above example, if the problem "data" (b) is known to D digits of accuracy, then the Hölder continuity implies that the solution (x-value of an intersection point) can only be inferred, in principle, to around D/p digits of accuracy. This may be adequate for imaging purposes but is an important constraint on any solution algorithm.

Taking advantage of the special structure of the classical phase retrieval problem, we are able to analyze the tangent bundle to the torus and show that it usually meets the set B nontransversally. Regardless of the algorithm, this renders the problem of finding points in $\mathbb{A} \cap B$ somewhat ill posed. Most algorithms for finding these points involve iterating some map. The failure of transversality usually weakens the attraction toward the fixed-point sets of these maps, often leading to very complicated, nonconvergent dynamics. As we shall see, it also makes a definitive analysis of these maps on the fixed-point set very complicated to carry out in practice, and so one must resort to numerical experimentation. Finally, our analysis points to ways of understanding improved methods for *collecting data* that will allow for better image reconstruction.

In the remainder of this rather long introduction, we present the definitions used throughout the book, and briefly describe the main results. The book is divided into three parts. Part I contains theoretical results. Part II investigates the behavior of standard algorithms used in phase retrieval and Part III provides statistical analyses of the behavior of "hybrid iterative maps," as well as suggestions for more robust reconstruction methods.

1.1 Discrete, Phase Retrieval Problems

We now turn to a description of the finite model problem that is studied in this book. As noted above, the unknowns consist of finitely many samples of $\rho(\boldsymbol{x})$ for $\boldsymbol{x}$ lying on a uniformly spaced lattice, for example,

$$\left\{f_{\boldsymbol{j}} = \rho\left(\frac{\boldsymbol{j}}{N}\right) : \boldsymbol{j} \in \mathbb{Z}^d\right\}, \tag{1.6}$$

where it is assumed that $\operatorname{supp}\rho$ is contained in a relatively compact subrectangle of $[0,1]^d$, with side-lengths less than $\frac{1}{2}$. The images we study in this book are often real valued, though many of our results, suitably modified, also hold for complex valued images. This is explained as required. Unless explicitly stated otherwise, real valued images can take both signs. We model the measurements as samples of the modulus of the DFT of this finite sequence. It is important to note that these measurements are not samples of the modulus of the continuous Fourier transform, $|\widehat{\rho}(\boldsymbol{\xi})|$, described above, but lead to a related problem that is easier to analyze. As noted above, when properly scaled, samples of the modulus of the DFT tend to uniformly spaced samples of $|\widehat{\rho}|$ in the limit $N \to \infty$.

For integers $n < N$, we define $[n : N] \overset{d}{=} \{n, n+1, \ldots, N-1, N\}$. A subset of $\mathbb{Z}^d$ of the form $[n_1 : N_1] \times \cdots \times [n_d : N_d]$ is called a *rectangular* subset. We call real or complex sequences indexed by such a set "images," or sometimes d-dimensional images. To define the discrete analogue of the classical phase retrieval problem we let $\boldsymbol{f}$ be an image $\{f_{\boldsymbol{j}} : \boldsymbol{j} \in J\}$, labeled by J a rectangular subset of $\mathbb{Z}^d$. We use the notation $\mathbb{R}^J$ to denote $\mathbb{R}^{|J|}$ with this indexing; respectively $\mathbb{C}^J$ to denote $\mathbb{C}^{|J|}$. Typically J takes the form $[0 : N]^d$, or $[-N : N]^d$. This approach to indexing images retains information about their underlying dimensions, and their organization in space.

Up to a rigid translation, any rectangular subset is equivalent, by translation, to a set of the following type: $J = [0 : N_1 - 1] \times \cdots \times [0 : N_d - 1]$. For an image $\boldsymbol{f} \in \mathbb{C}^J$, the DFT is defined, for $\boldsymbol{k} \in J$, to be

$$\mathscr{F}(\boldsymbol{f})_{\boldsymbol{k}} = \widehat{f}_{\boldsymbol{k}} \overset{d}{=} \sum_{j_1=0}^{N_1-1} \cdots \sum_{j_d=0}^{N_d-1} f_{\boldsymbol{j}} \prod_{l=1}^{d} \exp\left(-\frac{2\pi i j_l k_l}{N_l}\right). \tag{1.7}$$

The DFT could be defined for $\boldsymbol{k}$ in a different index set, as is sometimes done in the imaging literature; in this book we always take the domain of a d-dimensional image and its DFT to be the same rectangular subset of $\mathbb{Z}^d$. The formula for the DFT defines an extension of $\widehat{f}_{\boldsymbol{k}}$ from $\boldsymbol{k} \in J$ to all of $\mathbb{Z}^d$, as an image that is periodic in every index

$$\widehat{f}_{\boldsymbol{k}+(l_1 N_1, \ldots, l_d N_d)} = \widehat{f}_{\boldsymbol{k}} \quad \text{for all } (l_1, \ldots, l_d) \in \mathbb{Z}^d. \tag{1.8}$$

Definition 1.2 If $\boldsymbol{f} \in \mathbb{C}^J$, with J equivalent to $[0:N_1-1]\times\cdots\times[0:N_d-1]$, then $\boldsymbol{f}$, extended to all of $\mathbb{Z}^d$ by setting

$$f_{\boldsymbol{j}+(l_1N_1,\ldots,l_dN_d)} = f_{\boldsymbol{j}} \quad \text{for } \boldsymbol{j} \in J \text{ and all } (l_1,\ldots,l_d) \in \mathbb{Z}^d, \tag{1.9}$$

is said to be J*-periodically extended*. This agrees with the J-periodic extension of $\boldsymbol{f}$ given by inverse DFT.

The measurements in the discrete, classical, phase retrieval problem consist of a collection of nonnegative real numbers, which are interpreted as the magnitudes of the DFT of an unknown image $\boldsymbol{f}_0$:

$$a_{\boldsymbol{k}} = |\widehat{f}_{0\boldsymbol{k}}| \quad \text{for } \boldsymbol{k} \in J. \tag{1.10}$$

We denote the set of images with the given magnitude DFT data by $\mathbb{A}_{\boldsymbol{a}}$:

$$\mathbb{A}_{\boldsymbol{a}} = \{\boldsymbol{f} \in \mathbb{R}^J : |\widehat{f}_{\boldsymbol{k}}| = a_{\boldsymbol{k}} \quad \text{for all } \boldsymbol{k} \in J\} \tag{1.11}$$

and call this the *magnitude torus* defined by $\boldsymbol{a}$. Geometrically, this is a real torus, i.e., a product of round circles, sitting in $\mathbb{C}^J$. We use $\widehat{\mathbb{A}}_{\boldsymbol{a}}$ to denote its representation in the DFT domain

$$\widehat{\mathbb{A}}_{\boldsymbol{a}} = \{(a_{\boldsymbol{k}}e^{i\theta_{\boldsymbol{k}}}) : \theta_{\boldsymbol{k}} \in (-\pi,\pi] \quad \text{for all } \boldsymbol{k} \in J\}. \tag{1.12}$$

As in the continuum case, there are two operations on a (periodic) image that leave the magnitude DFT data invariant.

Definition 1.3 Let $\boldsymbol{f}$ be an image indexed by $J \subset \mathbb{Z}^d$, which is extended periodically to $\mathbb{Z}^d$.

(i) The *inversion* of $\boldsymbol{f}$ is defined by

$$\check{f}_{\boldsymbol{j}} \overset{d}{=} f_{-\boldsymbol{j}}. \tag{1.13}$$

Note that if $\boldsymbol{f}$ is real, then $\mathscr{F}(\check{\boldsymbol{f}})_{\boldsymbol{k}} = \overline{\mathscr{F}(\boldsymbol{f})_{\boldsymbol{k}}}$.

(ii) For $\boldsymbol{v} \in \mathbb{Z}^d$ the *translation* of $\boldsymbol{f}$ by $\boldsymbol{v}$, $\boldsymbol{f}^{(\boldsymbol{v})}$, is defined by

$$f^{(\boldsymbol{v})}_{\boldsymbol{j}} \overset{d}{=} f_{\boldsymbol{j}-\boldsymbol{v}}. \tag{1.14}$$

If $J = [1:N]^d$, then

$$\mathscr{F}(\boldsymbol{f}^{(\boldsymbol{v})})_{\boldsymbol{k}} = e^{\frac{2\pi i \boldsymbol{k}\cdot\boldsymbol{v}}{N}} \mathscr{F}(\boldsymbol{f})_{\boldsymbol{k}}. \tag{1.15}$$

(iii) If $\boldsymbol{f}$ is real valued, then $-\boldsymbol{f}$ has the same magnitude DFT data. If $\boldsymbol{f}$ is complex valued, then the images $\{e^{i\theta}\boldsymbol{f} : \theta \in [0,2\pi)\}$ all have the same magnitude DFT data.

The set of images generated from $\boldsymbol{f}$ by these operations are called its *trivial associates.*

We let $S_{\boldsymbol{f}} \subset J$ denote the *support of* $\boldsymbol{f}$

$$\boldsymbol{j} \in S_{\boldsymbol{f}} \text{ if and only if } f_{\boldsymbol{j}} \neq 0. \tag{1.16}$$

We assume that we have an estimate for $S_{\boldsymbol{f}}$, encoded as a subset $S \subset J$, such that $S_{\boldsymbol{f}} \subset S$. The goal of the discrete, classical, phase retrieval problem, using support as the auxiliary information, is to find an image $\boldsymbol{f}_1$, indexed by J, with the given magnitude DFT data

$$\begin{aligned} |\widehat{f}_{1\boldsymbol{k}}| = a_{\boldsymbol{k}} \quad &\text{for all } \boldsymbol{k} \in J, \text{ and} \\ S_{\boldsymbol{f}_1} \subset S.& \end{aligned} \tag{1.17}$$

That is, we search for points in $\mathbb{A}_{\boldsymbol{a}} \cap B_S$, where

$$B_S = \{\boldsymbol{f} \in \mathbb{R}^J : f_{\boldsymbol{j}} = 0 \text{ if } \boldsymbol{j} \notin S\}. \tag{1.18}$$

The solution to the phase retrieval problem is unique, *up to trivial associates*, if the set of points $\mathbb{A}_{\boldsymbol{a}} \cap B_S$ consist of trivial associates of one such point. From the perspective of applications, this sort of nonuniqueness does not pose any problems, and it is the strongest uniqueness statement that may be true.

Building on earlier work of Bruck and Sodin (1979), Hayes (1982) proved that, if the true support of $\boldsymbol{f}_0$, $S_{\boldsymbol{f}_0}$, is contained in a rectangle $R \subset J$ with the length of each side of R at most half the length of the corresponding side of J, then "generically" the solution of the phase retrieval problem is unique up to trivial associates. The meaning of generic in this context is explained below. If $S_{\boldsymbol{f}_0}$ is contained in such a rectangle, then we say that $\boldsymbol{f}_0$ has *small support*.

The Z-transform of $\boldsymbol{f}$ is defined by

$$\boldsymbol{X}(z) = \sum_{j \in J} f_{\boldsymbol{j}} \boldsymbol{z}^{-\boldsymbol{j}}; \tag{1.19}$$

using multi-index notation: $\boldsymbol{z} = (z_1, \ldots, z_d)$ and $\boldsymbol{z}^{\boldsymbol{j}} = z_1^{j_1} \ldots z_d^{j_d}$ is a monomial of degree $|\boldsymbol{j}| = j_1 + \cdots + j_d$. Note that the degree of a monomial can be either positive or negative. A polynomial is a finite sum of monomials of nonnegative degrees

$$p(z) = \sum_{\boldsymbol{j} \in \mathscr{J}} a_{\boldsymbol{j}} \boldsymbol{z}^{\boldsymbol{j}}, \text{ with } \mathscr{J} \text{ a finite set.} \tag{1.20}$$

Its degree is defined as the

$$\deg p \stackrel{d}{=} \max\{|\boldsymbol{j}| : \boldsymbol{j} \in \mathscr{J} \text{ and } a_{\boldsymbol{j}} \neq 0\}.$$