Complete subsets of a diffraction pattern

2001 J. Phys. A: Math. Gen. 34 731

(http://iopscience.iop.org/0305-4470/34/4/304)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **34** (2001) 731–755 www.iop.org/Journals/ja PII: S0305-4470(01)13440-2

Complete subsets of a diffraction pattern

Antonio Cervellino[1] **and Salvino Ciccariello**[2,3]

[1] ETH Zentrum, Laboratory of Crystallography, CH-8092 Zurich, Switzerland
[2] Dipartimento di Fisica 'G Galilei' and Sez. INFM, via Marzolo 8, I-35131 Padova, Italy

Received 20 April 2000

Abstract

A positive atomic density $\rho(\boldsymbol{r}) = \sum_{j=1}^{N} n_j \delta(\boldsymbol{r} - \boldsymbol{r}_j)$ in a D-dimensional space can be exactly reconstructed from an appropriate finite subset (*complete set*) of its Fourier series coefficients $\{U_{\boldsymbol{h}}\}_{\boldsymbol{h}\in\mathbb{Z}^D}$ or even (limited to the support $\{\boldsymbol{r}_j\}_{j=1,\ldots,N}$) from a finite subset of *moduli* $|U_{\boldsymbol{h}}|$. It is necessary first to determine a complete set of Fourier coefficients $U_{\boldsymbol{h}}$ (possibly inside the unavoidable high-resolution cut-off $\|\boldsymbol{h}\| < L$) and then, by these coefficients, to determine the unknown density. We report some procedures which are able to single out complete sets. They are based on a property of Goedkoop's vector lattice $\{|A_{\boldsymbol{h}}\rangle\}_{\boldsymbol{h}\in\mathbb{Z}^D}$, defined so that $\langle A_{\boldsymbol{h}}|A_{\boldsymbol{k}}\rangle = U_{\boldsymbol{k}-\boldsymbol{h}}$. The property states that if the vector with index $\boldsymbol{h}^\star = (h_1^\star, \ldots, h_D^\star)$ is a linear combination of the vectors relevant to a set of indices with a particular shape, then all the vectors relevant to the hyperquadrant $\mathcal{Q}_{\boldsymbol{h}^\star} \equiv \{\boldsymbol{h} \mid h_\alpha \geqslant h_\alpha^\star,\ \alpha = 1, \ldots, D\}$ are linear combinations of the vectors relevant to $\mathcal{Q}_{\boldsymbol{0}} \setminus \mathcal{Q}_{\boldsymbol{h}}^\star$. Moreover, the determination of ρ from a complete set passes through the solution of a system of polynomial equations in D variables, whose roots determine the position vectors. We show how to convert this problem into the simpler problem of sequentially solving a set of polynomial equations in one variable.

PACS numbers: 6110D, 6110Y, 6112B

1. Introduction

This paper deals with a basic crystallographic problem which, for its implications in many other fields of modern science, can be considered a classical problem in mathematical physics. The problem consists in determining *exactly* an atomic density (cf definition 1) from an appropriate *finite* subset (*complete set*) of its Fourier coefficients. It has been shown (cf [1]) that its solution is the key step in solving the more awkward problem of determining the support of an atomic density from a finite subset of the *moduli* of Fourier coefficients.

The notion of a *complete set*, denoted by $\mathcal{C}$, is central. Its precise definition will be given in section 1.1. We recall that Fourier coefficients (or their moduli) are observable inside a preset high-resolution limit. Thus, to solve the aforementioned problem, we need to determine

[3] Author to whom correspondence should be addressed.

a *complete set* of Fourier coefficients (or of their moduli) which lies inside such limit, and then to set forth an explicit algorithm which is able to determine the unknown density. Solutions of these two points are reported in this work. For greater simplicity, in working out these results (sections 4–6), we refer to two-dimensional space ($D = 2$). Section 7 provides the generalization to arbitrary dimension D.

The plan of the paper is as follows. In section 1.1 of this introduction we give a general formulation of the problem that will also be split into a series of simpler questions, while section 1.2 demonstrates connections with the trigonometric moments theory. In section 2 the problem is reformulated in more explicit crystallographic terms and the question of the moduli is addressed. Then, section 3 introduces the solution procedure and the criteria to isolate a complete set $\mathcal{C}$. These criteria must be implemented with a search algorithm which is able to isolate complete sets having the desired characteristics. Section 4 illustrates one important result (property 1) which is useful in formulating search algorithms. Section 5 uses property 1 to find out *complete sets* at small resolution. Section 6 shows how to convert the knowledge of a complete set $\mathcal{C}$ into a set of polynomial equations in a single variable whose roots yield the atomic coordinates. These results are generalized to arbitrary dimension D in section 7 where some concluding remarks are also reported. Finally, some lengthy proofs are left to the appendix.

1.1. General formulation

Notation. Given a set $\mathcal{S} \subset \mathbb{Z}^D$ we will denote by $\mathcal{S}-\mathcal{S}$ the difference set $\{\boldsymbol{h} - \boldsymbol{h}' \,|\, \boldsymbol{h}, \boldsymbol{h}' \in \mathcal{S}\}$. For any $\boldsymbol{k} \in \mathbb{Z}^D$ and any set $\mathcal{S} \subset \mathbb{Z}^D$, $\mathcal{S} + \boldsymbol{k}$ denotes the $\boldsymbol{k}$-translated set $\{\boldsymbol{h} + \boldsymbol{k} | \boldsymbol{h} \in \mathcal{S}\}$. The integer D-plas $\boldsymbol{a}^*_\alpha$, $\alpha = 1, \ldots, D$ (i.e. $(1, 0, \ldots, 0)$, $(0, 1, \ldots, 0)$, $\ldots$, $(0, 0, \ldots, 1)$) denote the canonical basis of $\mathbb{R}^D$ (generators of $\mathbb{Z}^D$). Given $\boldsymbol{h} \in \mathbb{Z}^D$, we will denote by $\mathcal{N}_{\boldsymbol{h}} \equiv \{\boldsymbol{h} \pm \boldsymbol{a}^*_\alpha,\ \alpha = 1, \ldots, D\}$ the *first neighbourhood* of $\boldsymbol{h}$. Given a set $\mathcal{S} \subset \mathbb{Z}^D$, when no confusion is possible, we shall sometimes also denote by $\mathcal{S}$ the set of Fourier coefficients $\{U_{\boldsymbol{h}}\}$ of a fixed density $\rho(\boldsymbol{r})$ which have the indices $\boldsymbol{h}$ in $\mathcal{S}$.

Definition 1. *We define the* atomic density *ρ as a distribution having as its domain the unit hypercube*

$$\mathbb{U}^D \equiv \left\{ \boldsymbol{r} = (x_1, \ldots, x_D) \in \mathbb{R}^D \,\middle|\, x_\alpha \in [0, 1), \alpha = 1, \ldots, D \right\} \tag{1}$$

and constituted by a finite sum of N Dirac masses in the distinct *points $\{\boldsymbol{r}_j\}_{j=1,\ldots,N} \in \mathbb{U}^D$ (*scattering centres*) with* strictly positive *weights $\{n_j\}_{j=1,\ldots,N}$ (*charges*), so that*

$$\rho(\boldsymbol{r}) = \sum_{j=1}^{N} n_j \delta(\boldsymbol{r} - \boldsymbol{r}_j). \tag{2}$$

Remark 1. The normalization $\sum_{j=1}^{N} n_j = 1$ will be assumed throughout this work.

The Fourier series coefficients of ρ are, $\forall\ \boldsymbol{h} \in \mathbb{Z}^D$,

$$U_{\boldsymbol{h}} = \int_{\mathbb{U}^D} \mathrm{d}^D \boldsymbol{r}\, \rho(\boldsymbol{r})\, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{r}} = \sum_{j=1}^{N} n_j\, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{r}_j}. \tag{3}$$

The symbol $\cdot$ means the usual scalar product in $\mathbb{R}^D$. The set $\{U_{\boldsymbol{h}}\}_{\boldsymbol{h}\in\mathbb{Z}^D}$ will also be called the *infinite diffraction pattern*, or IDP. The inverse Fourier transform is

$$\rho(\boldsymbol{r}) = \sum_{\boldsymbol{h}\in\mathbb{Z}^D} U_{\boldsymbol{h}}\, \mathrm{e}^{-2\pi \mathrm{i} \boldsymbol{h}\cdot\boldsymbol{r}}. \tag{4}$$

Remark 2. In general, $U_h \in \mathbb{C}$. With ρ being real-valued, equation (3) yields $U_{-h} = \overline{U}_h$ (an overbar denotes the complex conjugate). Furthermore, normalizing ρ as in remark 1, it is $U_{\mathbf{0}} = 1$.

Remark 3. The restriction of the index $\boldsymbol{h}$ to $\mathbb{Z}^D$ corresponds to considering ρ as the restriction to $\mathbb{U}^D$ of a periodic function of period $\mathbb{U}^D$. In this sense equation (4) is valid for any $\boldsymbol{r} \in \mathbb{R}^D$. The coordinates of the scattering centres $\{\boldsymbol{r}_j\}_{j=1,\ldots,N}$ remain uniquely defined modulo 1.

Owing to the functional form of equation (2), ρ depends on a finite number of parameters, namely the coordinates of the scattering centres $\{\boldsymbol{r}_j\}_{j=1,\ldots,N}$ and the charges $\{n_j\}_{j=1,\ldots,N}$, while the IDP $\{U_h\}_{h\in\mathbb{Z}^D}$ is infinite. We will consider subsets $\{U_h\}_{\|h\|<L}$ (*limited diffraction pattern*, or LDP) where $L > 0$ is the high-resolution cut-off and $\|\boldsymbol{h}\|^2 \equiv \boldsymbol{h} \cdot \mathsf{M}^*\boldsymbol{h}$, where M^* is a positive-definite metric tensor (cf section 2.1).

The problem we deal with is the following:

Problem 1 (Crystallographic phase problem). *Knowing the* moduli $|U_h|$ *of the Fourier coefficient of an atomic density ρ belonging to a LDP $\{U_h\}_{\|h\|<L}$ of given radius L, and* with prior knowledge *of the number N of scattering centres and the values of the charges $\{n_j\}_{j=1,\ldots,N}$, is it possible to and, in the affirmative case, how do we determine the coordinates $\{\boldsymbol{r}_j\}_{j=1,\ldots,N}$ of the scattering centres?*

The physical aspects of this problem (cf also section 2) are clear. $|U_h|$ are the observables, L is an experimental high-frequency cut-off and ρ is the unknown atomic density to be determined. As shown in [1] (briefly summarized in section 2.2), problem 1 can be split into two sequential subproblems. The first is:

Problem 2 (Complex amplitude problem). *Knowing the Fourier coefficients U_h of an atomic density ρ belonging to a LDP $\{U_h\}_{\|h\|<L}$ of given radius L, is it possible to (and how do we) determine*

- *the number N of scattering centres,*
- *the coordinates $\{\boldsymbol{r}_j\}_{j=1,\ldots,N}$ of the scattering centres,*
- and *the charges $\{n_j\}_{j=1,\ldots,N}$?*

The second subproblem consists in deconvolving the density obtained by solving problem 2. As the deconvolution is essentially a discrete one, it can always be solved (see proposition 1 in section 2.2), provided the density is known precisely. In contrast, a simple truncated inverse Fourier transform based only on the coefficients with $\|\boldsymbol{h}\| < L$ is often insufficient in this respect.

Problem 2, which is important in its own right, is the main object of this paper. For a more detailed explanation, we will reformulate problem 2 as a sequence of questions. The first question is:

Question 1. *Given an atomic density ρ (cf definition 1) does there exist a* finite *subset $\mathcal{C} \subset \{U_h\}_{h\in\mathbb{Z}^D}$ of its IDP such that* all *the $U_h \notin \mathcal{C}$ can be determined from those in $\mathcal{C}$?*

For the one-dimensional case the answer has been known to be affirmative since 1927 thanks to the work of Ott [2]. For larger dimensions, the question was answered affirmatively only in recent years (cf [1] and references therein).

Definition 2. *For a given atomic density ρ a* finite *subset $\mathcal{C} \subset \{U_h\}_{h\in\mathbb{Z}^D}$ of its IDP such that:*

- all *the $U_h \notin \mathcal{C}$ can be determined from those in $\mathcal{C}$;*
- *none of its elements can be excluded without failing to determine part of the IDP $\{U_h\}_{h\in\mathbb{Z}^D}$*

will be called a complete set $\mathcal{C}$.

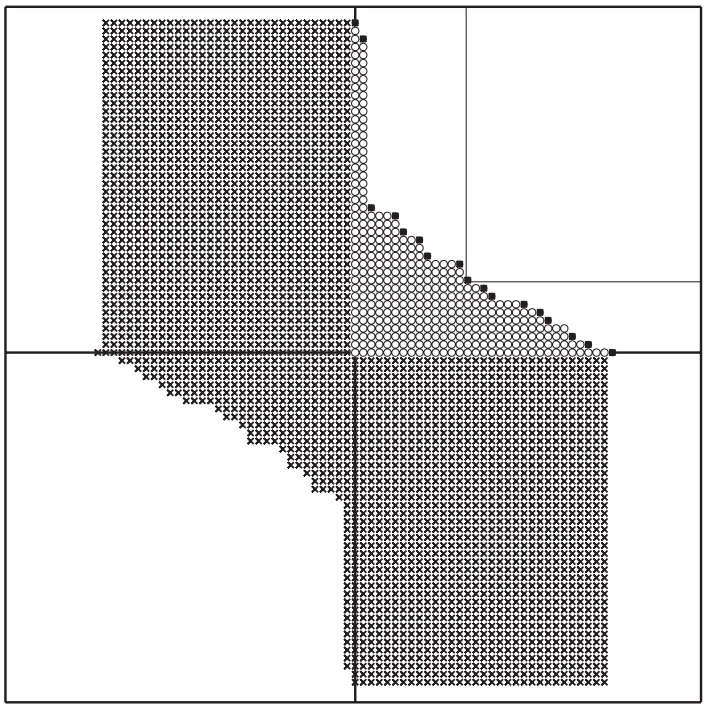

Figure 1. An illustration of a base set $\mathcal{B} = \mathcal{B}^{a_1^*}[m_1 \dots m_p]$ (○) obtained by the algorithm of section 3.2 and [1]. It is shown also the relevant complementary set $\mathcal{B}^c$ (●), and the rest of the resulting complete set $\mathcal{C}$ (×). $\mathcal{B}^c$ consists of the KH zeros found during the search. One of the quadrants originating on a KH zero and hence excluded from the search is also shown. $\mathcal{C}$ contains all the indices of the entries of the KH matrices used in the search. The density ρ is described in figure 3. We have $p = 41$ rows with lengths $[m_1 \dots m_p] = [32, 29, 27, 27, 24, 23, 21, 17, 16, 14, 14, 13, 9, 9, 8, 6, 6, 5, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1]$.

Remark 4. We list some general properties characterizing a complete set $\mathcal{C}$:

- Necessarily $U_{\mathbf{0}} \equiv U_{(0,0,\dots,0)} \in \mathcal{C}$ and, owing to remark 2, we lose no generality in assuming that, if $U_{\boldsymbol{h}} \in \mathcal{C}$, also $U_{-\boldsymbol{h}} = \overline{U}_{\boldsymbol{h}} \in \mathcal{C}$.
- Only atomic densities (cf definition 1) admit a finite complete set (cf [3], section 3).
- The fact that any set $\mathcal{C}$ is or not a complete set for an atomic density ρ *does not depend on the values of the (positive) charges* $\{n_j\}_{j=1,\dots,N}$, but only on the coordinates of the scattering centres $\{\boldsymbol{r}_j\}_{j=1,\dots,N}$ (cf [3], proposition 3.2).
- The set of indices $\{\boldsymbol{h}\}_{U_{\boldsymbol{h}} \in \mathcal{C}}$ of a complete set $\mathcal{C}$ cannot be completely contained in a proper subgroup of $\mathbb{Z}^D$ (cf [3], theorem 3.5).

We can now proceed with a further question.

Question 2. *Is the complete set $\mathcal{C}$ uniquely determined for a fixed atomic density ρ?*

In general, for a given ρ, *several different* complete sets $\mathcal{C}$ may exist. Even their cardinality may differ. An example is reported in figures 1 and 2.

Question 3. *Is there an algorithm which is able, for any given (unknown) ρ, to single out at least* one of the possible *complete sets $\mathcal{C}$?*

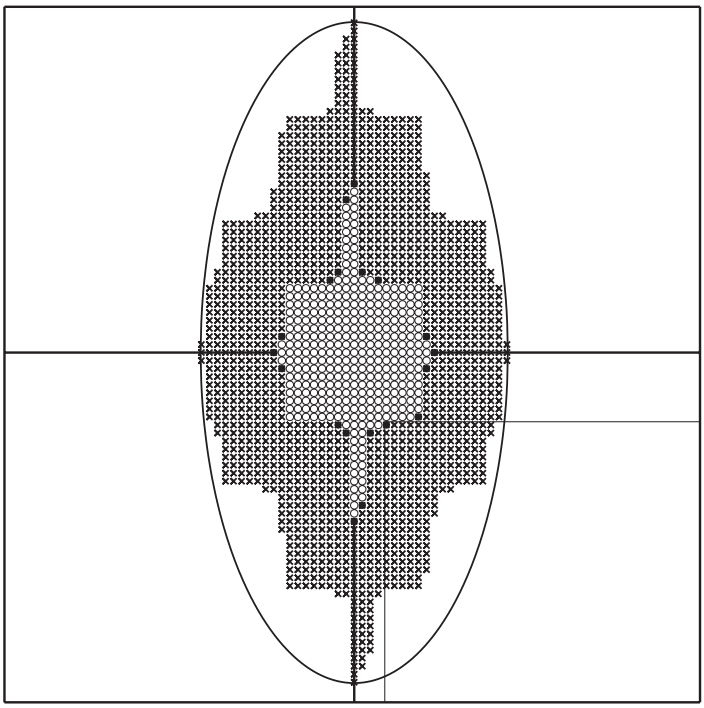

Figure 2. A base set $\mathcal{B}$ (○) isolated by the *centred square search* algorithm (cf section 5) is shown, together with the relevant complementary set $\mathcal{B}^c$ (●) and complete set $\mathcal{C}$ (×). Sets $\mathcal{B}$ and $\mathcal{C}$ are much more compact than those of figure 1. The ellipse (——) delimits a hypothetical high-resolution cut-off that contains $\mathcal{C}$. One of the excluded quadrants (——) originating on a KH zero is also shown.

For the one-dimensional case this algorithm has long been known [2]. For higher dimensions, such an algorithm was devised for the first time in [1] and is sketched in section 3.2. In sections 4 and 5 we present an improved algorithm.

Remark 5. Both algorithms yield complete sets $\mathcal{C}$ characterized by some further properties:

- *Construction*. The set of indices $\{\boldsymbol{h}\}_{U_{\boldsymbol{h}}\in\mathcal{C}}$ is a union of difference sets $\mathcal{S}_k - \mathcal{S}_k$ of a *finite* family of finite sets of indices $\mathcal{S}_1, \ldots, \mathcal{S}_M$ which satisfy (cf equation (23))
  - $\mathcal{S}_k \subset \{\boldsymbol{h}\}_{U_{\boldsymbol{h}}\in\mathcal{C}}$, $k = 1, \ldots, M$;
  - $\mathcal{S}_k \setminus (\mathcal{S}_k \cap \mathcal{S}_{k+1})$ is a singleton, $\mathcal{S}_{k+1} \setminus (\mathcal{S}_k \cap \mathcal{S}_{k+1}) \neq \emptyset$.

  The union of these singletons is denoted by $\mathcal{B}^c$ (i.e. $\mathcal{B}^c \equiv \bigcup_{k=1}^{M-1} \mathcal{S}_k \setminus (\mathcal{S}_k \cap \mathcal{S}_{k+1})$) and is called the *complementary set* of the *base set* $\mathcal{B} \equiv \left(\bigcup_{k=1}^{M-1} \mathcal{S}_k\right) \setminus \mathcal{B}^c$ (cf section 3.2).
- *Connectedness*. If $\boldsymbol{h} \in \mathcal{C}$, then its first neighbourhood contains other indices of $\mathcal{C}$: $\mathcal{N}_{\boldsymbol{h}} \cap \mathcal{C} \neq \emptyset$.
- *Boundedness*. $\mathcal{C} \subset \{\boldsymbol{h} \,|-N_\alpha \leqslant h_\alpha \leqslant N_\alpha,\ \alpha = 1, \ldots, D\}$, where N_α is the number of *distinct* αth coordinates $x_{j,\alpha}$ of the scattering centres $\{\boldsymbol{r}_j\}_{j=1,\ldots,N}$. See property A2 in the appendix.

Clearly, if we have a high-frequency cut-off L, it is of interest to single out a complete set fully contained in the LDP. Formally we must require

$$L_{\mathcal{C}} \equiv \max_{U_{\boldsymbol{h}}\in\mathcal{C}} \|\boldsymbol{h}\| < L. \tag{5}$$

Question 4. *Given an (unknown) ρ and a fixed high-frequency cut-off L, is there an algorithm which is able to single out a complete set $\mathcal{C}$ having $L_{\mathcal{C}} < L$?*

The answer of course depends on the value of L. A hyperparallelepiped bounding the set of indices of $\mathcal{C}$ is given in remark 5. However, it generally yields a very large bound for $L_{\mathcal{C}}$. In contrast, the number of unknown quantities present in equation (2) is $N(D+1)+1$ and this value is a lower bound for the cardinality of $\mathcal{C}$.

If L is too small the data will not be sufficient. Since the smallest cardinality of complete sets is *a priori* not known, question 4 is better reformulated as:

Question 5. *For any given (unknown) ρ, is there an algorithm which is able to single out the complete set $\mathcal{C}$ with* the smallest possible *$L_{\mathcal{C}}$?*

The answer is affirmative. The algorithm is discussed in section 5. It is based on an 'isotropic' inclusion algorithm which starts from $\boldsymbol{h} = \mathbf{0}$ and ensures the result by construction. Such an algorithm, in the case of insufficient data, would stop prematurely. In any other case it will yield a solution. In contrast, the algorithm reported in [1] is based on a *directional* inclusion algorithm with the preferred growth direction along one of the base vectors $\{\boldsymbol{a}^*_\alpha\}_{\alpha=1,\dots,D}$ and the resulting sets will turn out to be quite elongated along the chosen direction.

Finally, we have to ask:

Question 6. *Given any complete set $\mathcal{C}$ relevant to a given (unknown) ρ, is there an algorithm able to determine the associated ρ?*

The answer is again affirmative. It is possible to proceed in two different ways. The first is an iterative algorithm (cf [1]) which, starting from $\mathcal{C}$, is able to determine a number of Fourier coefficients sufficiently large to calculate ρ, with the desired accuracy, by using the inverse Fourier transform given by equation (4). The second, more analytical one, uses the knowledge of $\mathcal{C}$ to construct a system of polynomial equations in D variables, which is converted by the procedure explained in section 6 into a (finite) sequence of univariate polynomial equations whose solutions yield the coordinates $\{\boldsymbol{r}_j\}_{j=1,\dots,N}$ of the density support.

1.2. Connection with the trigonometric moments theory

Let

$$\mathbb{T}^D \equiv \left\{ \boldsymbol{z} = (z_1, \dots, z_D) \in \mathbb{C}^D \,\middle|\, |z_\alpha| = 1, \alpha = 1, \dots, D \right\}.$$

Given any atomic (positive) density $\rho(\boldsymbol{r})$ on $\mathbb{U}^D$ (cf definition 1), by the natural isomorphism

$$\tau : \boldsymbol{r} \longmapsto \boldsymbol{z} \equiv (z_1, \dots, z_D) = (\mathrm{e}^{2\pi \mathrm{i} x_1}, \dots, \mathrm{e}^{2\pi \mathrm{i} x_D}) \equiv \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{r}}$$

between $\mathbb{U}^D$ and $\mathbb{T}^D$ we can define $\mu(\boldsymbol{z}) = \rho\left(\tau^{-1}(\boldsymbol{z})\right)$ which turns out to be a (positive) Borel measure (cf [5]) on $\mathbb{T}^D$ having finite support $\{\mathrm{e}^{2\pi \mathrm{i} \boldsymbol{r}_j}\}_{j=1,\dots,D} \subset \mathbb{T}^D$. Note that the Fourier coefficients $U_{\boldsymbol{h}}$ of $\rho(\boldsymbol{r})$ are the trigonometric moments of $\mu(\boldsymbol{z})$.

The theory of truncated trigonometric moments deals with the relations between finite index sets $\mathcal{S} \subset \mathbb{Z}^D$, associate coefficients sets $\{U_{\boldsymbol{h}} \,|\, \boldsymbol{h} \in \mathcal{S}\}$ and positive measures $\mu(\boldsymbol{z})$ on $\mathbb{T}^D$. Consider first the *determinacy problem* (cf [3]): given a set $\mathcal{S} \subset \mathbb{Z}^D$ containing $\mathbf{0}$ and such that $\boldsymbol{h} \in \mathcal{S}$ iff $-\boldsymbol{h} \in \mathcal{S}$, and given an atomic density with finite support ρ, to find whether the subset of Fourier coefficients $\{U_{\boldsymbol{h}}\}_{\boldsymbol{h}\in\mathcal{S}}$ determine ρ uniquely. The solution of problem 2 (see sections 3–6) may give an important contribution to this field.

Another topic in the theory of trigonometric moments is the *extension problem* [4]. Some results in this field are particularly illuminating.

Consider a finite set $\mathcal{S} \subset \mathbb{Z}^D$ and its difference set $\mathcal{S} - \mathcal{S}$. Let M denote the cardinality of $\mathcal{S}$ and M' the cardinality of $\mathcal{S} - \mathcal{S}$. Let us order the elements of $\mathcal{S}$ arbitrarily as $\boldsymbol{h}_1, \dots, \boldsymbol{h}_M$. Consider now an arbitrary collection of complex numbers $\{U_{\boldsymbol{h}}\}_{\boldsymbol{h}\in\mathcal{S}-\mathcal{S}}$.

Definition 3. *The* generalized Toeplitz matrix *(cf [6]) generated by $\mathcal{S}$ for the set $\{U_h\}_{h\in\mathcal{S}-\mathcal{S}}$ is the matrix of order M whose entries are $(\mathcal{U})_{kl} = U_{h_l-h_k}$.*

The set $\{U_h\}_{h\in\mathcal{S}-\mathcal{S}}$ is said to be (semi-)positive definite when the relevant Toeplitz matrix is (semi-)positive definite. If for a choice $\{U_h\}_{h\in\mathcal{S}-\mathcal{S}}$ there exists a positive (not necessarily atomic) density ρ on $\mathbb{U}^D$ (measure μ on $\mathbb{T}^D$) such that $\{U_h\}_{h\in\mathcal{S}-\mathcal{S}}$ are a subset of its Fourier coefficients (moments), then ρ (μ) is said to be a *representing density (measure)* for $\{U_h\}_{h\in\mathcal{S}-\mathcal{S}}$, which is said to be *extendable*. A necessary condition (cf [7]) for the existence of a representing density (measure) is that $\{U_h\}_{h\in\mathcal{S}-\mathcal{S}}$ should be (semi-)positive definite. Anyhow, for general choices of $\mathcal{S}$, this condition *is not sufficient* (counterexamples are reported in [4]). However, if $\{U_h\}_{h\in\mathcal{S}-\mathcal{S}}$ admits (possibly many) representing densities, theorem 3.4 of [4] states that between the representing densities there is always an atomic density whose finite support has cardinality $N \leqslant M$. Hence, for any finite set $\mathcal{S} \subset \mathbb{Z}^D$, the extendibility of a collection $\{U_h\}_{h\in\mathcal{S}-\mathcal{S}}$ is equivalent to the existence of a representing *atomic* density.

Given a finite set $\mathcal{S}$, a sufficient condition for a (semi-)positive definite collection of complex numbers $\{U_h\}_{h\in\mathcal{S}-\mathcal{S}}$ to yield at least one representing density, is given in theorem 5.2 of [4]. The theorem uses the concept of *Goedkoop's vector lattice*, first introduced in [8].

Definition 4. *Consider a finite-dimensional Hilbert space $\mathfrak{H}$ with a Hermitian scalar product $\langle\cdot|\cdot\rangle$ (in Dirac's notation). Consider the family of vectors $|A_h\rangle \in \mathfrak{H}$, $\boldsymbol{h} \in \mathbb{Z}^D$, having the properties:*

- *the vectors $\{|A_h\rangle\}_{h\in\mathbb{Z}^D}$ span $\mathfrak{H}$;*
- *the vectors are generated by a vector $|A_{\mathbf{0}}\rangle$, such that $\||A_{\mathbf{0}}\rangle\|^2 = \langle A_{\mathbf{0}}|A_{\mathbf{0}}\rangle = 1$, and by D commuting unitary linear operators $\mathsf{G}_1, \ldots, \mathsf{G}_D$ through the relations*

$$|A_h\rangle = \mathsf{G}_1^{h_1}\mathsf{G}_2^{h_2}\ldots\mathsf{G}_D^{h_D}|A_{\mathbf{0}}\rangle \qquad \forall\, \boldsymbol{h} = (h_1, \ldots, h_D) \in \mathbb{Z}^D.$$

A consequence of these properties is that

- $\||A_h\rangle\|^2 = \langle A_h|A_h\rangle = 1$, *for every* $\boldsymbol{h} \in \mathbb{Z}^D$,
- $\langle A_h|A_{h+k}\rangle = \langle A_{h'}|A_{h'+k}\rangle$, *for every* $\boldsymbol{h}, \boldsymbol{h}', \boldsymbol{k} \in \mathbb{Z}^D$.

The set of vectors $\mathcal{G} \equiv \{|A_h\rangle\}_{h\in\mathbb{Z}^D}$ is called Goedkoop's vector lattice. *The term 'lattice' is used to stress that $\mathbb{Z}^D$ is homomorphic to $\mathcal{G}$ after defining in $\mathcal{G}$ the sum operation as $|A_h\rangle \boxplus |A_k\rangle = |A_{h+k}\rangle$.*

Goedkoop [8] showed that when a representing atomic density exists, one can always construct this vector lattice (cf definition 6). Theorem 5.2 of [4] states that, given any finite set $\mathcal{S} \subset \mathbb{Z}^D$ and any (semi-)positive definite collection of complex numbers $\{U_h\}_{h\in\mathcal{S}-\mathcal{S}}$, the existence of Goedkoop's vector lattice such that $\langle A_h|A_k\rangle = U_{k-h}\ \forall\, \boldsymbol{h}, \boldsymbol{k} \in \mathcal{S}$ is also sufficient for the extendibility of $\{U_h\}_{h\in\mathcal{S}-\mathcal{S}}$. Goedkoop's vector lattice turns out to be a very powerful analytical tool. We will show, in fact, that this construction also allows us to solve problem 2.

2. Crystallographic formulation and splitting

2.1. Some crystallographic definitions

Presently, the definition of *crystal* is still evolving towards more general formulations [9]. The most recent definition requires a crystal to have an *essentially discrete* diffraction pattern. This means that the scattered intensity is concentrated in a set of sharp Bragg peaks. In the following, however, our analysis will deal with *ideal crystals*, defined below.

Definition 5. *A D-dimensional ideal crystal results from the periodic replica of its unit cell* C*, determined by a set of D linearly independent vectors* $\underline{\boldsymbol{a}}_1,\ \underline{\boldsymbol{a}}_2, \ldots, \underline{\boldsymbol{a}}_D \in \mathbb{R}^D$ *via*

$$\mathsf{C} \equiv \left\{ x_1\underline{\boldsymbol{a}}_1 + \cdots + x_D\underline{\boldsymbol{a}}_D \;\middle|\; 0 \leqslant x_\alpha < 1, \alpha = 1, \ldots, D \right\}.$$

The unit cell contains N atoms with specified atomic numbers $Z_1, Z_2, \ldots, Z_N$*. The D-plas* $\boldsymbol{r} \equiv (x_1, \ldots, x_D)$ *(conventionally* fractional coordinates*) evidently define* $\mathbb{U}^D$ *(cf equation (1)). The electron density in the unit cell hence may be reduced (cf [10]) to the form*

$$\rho(\boldsymbol{r}) = \sum_{j=1}^{N} Z_j \delta(\boldsymbol{r} - \boldsymbol{r}_j). \tag{6}$$

We arrive at exactly the form of equation (2) normalizing as in remark 1, with

$$n_j = Z_j/(Z_1 + \cdots + Z_N). \tag{7}$$

The diffraction diagram consists of sharp Bragg peaks of intensity $I_{\boldsymbol{h}}$ located at the lattice of points $\boldsymbol{q}_{\boldsymbol{h}}$ of reciprocal space (*reciprocal lattice*) given by

$$\boldsymbol{q}_{\boldsymbol{h}} = h_1\underline{\boldsymbol{a}}_1^* + \cdots + h_D\underline{\boldsymbol{a}}_D^* \qquad \text{with} \quad \boldsymbol{h} \equiv (h_1, \ldots, h_D) \in \mathbb{Z}^D. \tag{8}$$

Here vectors[4] $\underline{\boldsymbol{a}}_\alpha^*$s, which define the reciprocal unit cell, are determined by the relations $\underline{\boldsymbol{a}}_\alpha^* \cdot \underline{\boldsymbol{a}}_\beta = \delta_{\alpha\beta}$, $\alpha, \beta = 1, \ldots, D$, where $\delta_{\alpha\beta}$ is the Kronecker symbol.

Note. $\boldsymbol{q}_{\boldsymbol{h}}$ is conveniently replaced by $\boldsymbol{h}$. Distances are preserved by defining $\|\boldsymbol{h}\| \equiv \|\boldsymbol{q}_{\boldsymbol{h}}\| = (\boldsymbol{q}_{\boldsymbol{h}} \cdot \boldsymbol{q}_{\boldsymbol{h}})^{1/2}$. From equation (8), $\|\boldsymbol{h}\| = (\boldsymbol{h}\mathsf{M}^*\boldsymbol{h})^{1/2}$, with $\mathsf{M}^*_{\alpha\beta} = \underline{\boldsymbol{a}}_\alpha^* \cdot \underline{\boldsymbol{a}}_\beta^*$. If λ is the incident radiation wavelength, the maximum experimental range for crystal diffraction is $\|\boldsymbol{h}\| < L = 2/\lambda$ (*limiting sphere*, cf [10]).

The Fourier coefficients of equation (6) normalized as in equation (7) are

$$U_{\boldsymbol{h}} \equiv \sum_{j=1}^{N} n_j \mathrm{e}^{2\pi \mathrm{i}\boldsymbol{h}\cdot\boldsymbol{r}_j} \tag{9}$$

coincident with equation (3). The $U_{\boldsymbol{h}}$s are called *unitary structure factors* (USF). After the appropriate corrections, the Bragg peak intensity at $\boldsymbol{q}_{\boldsymbol{h}}$ (cf [10, 11]) is related to $U_{\boldsymbol{h}}$ by

$$I_{\boldsymbol{h}} = \left|U_{\boldsymbol{h}}\right|^2. \tag{10}$$

The crystal structure is known when the positions $\{\boldsymbol{r}_j\}_{j=1,\ldots,N}$ of the atoms in a unit cell are known. The *crystallographic phase problem* is equivalent to problem 1 of section 1.1, because it consists in determining the atomic positions knowing the intensity of the Bragg peaks $I_{\boldsymbol{h}}$ in a given limiting sphere $\|\boldsymbol{h}\| < L = 2/\lambda$.

2.2. Splitting of the problem

A procedure which is able to yield the solution(s) of the crystallographic phase problem (generally formulated in section 1.1, problem 1) is the so-called *algebraic approach* (cf [2, 12–16], and the recent review in [1]). The crystallographic phase problem may be split into two problems (as anticipated in section 1.1) to be solved subsequently. Let us show this point.

Splitting of the problem. The substitution of equation (9) into equation (10) yields

$$I_{\boldsymbol{h}} = \sum_{j=1}^{N} n_j^2 \quad + \sum_{1 \leqslant i \neq j \leqslant N} n_i n_j \mathrm{e}^{2\pi \mathrm{i}\boldsymbol{h}\cdot(\boldsymbol{r}_i - \boldsymbol{r}_j)}. \tag{11a}$$

[4] The crystallographic (three-dimensional) notation is $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ for $\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \underline{\boldsymbol{a}}_3$ and $\boldsymbol{a}^*, \boldsymbol{b}^*, \boldsymbol{c}^*$ for $\underline{\boldsymbol{a}}_1^*, \underline{\boldsymbol{a}}_2^*, \underline{\boldsymbol{a}}_3^*$.

Each exponential present on the right-hand side of (11a) is left invariant by the transformation

$$\boldsymbol{r}_i - \boldsymbol{r}_j \quad \longmapsto \quad \vec{\delta}_{ij} = (\boldsymbol{r}_i - \boldsymbol{r}_j) \mod 1 \tag{11b}$$

where the modulo operation is applied to each component. Thus, we can associate to each $(\boldsymbol{r}_i - \boldsymbol{r}_j)$ a vector $\vec{\delta}_{ij}$ uniquely determined by the condition that it lies inside the unit cell $\mathbb{U}^D$ (i.e. $0 \leqslant \delta_{ij,k} < 1$, for $k = 1, 2, \ldots, D$). Now label the *distinct* $\vec{\delta}_{ij}$s with index $\hat{\boldsymbol{\jmath}}$ running from 1 to $\bar{\mathcal{N}}$, and denote by $\mathcal{L}_{\hat{\boldsymbol{\jmath}}}$ the set of the distinct ordered pairs (i, j) such that $\vec{\delta}_{ij} = \vec{\delta}_{\hat{\boldsymbol{\jmath}}}$. Even though it is known that $N \leqslant \bar{\mathcal{N}} \leqslant N(N-1)/2$ and that $\bigcup_{\hat{\boldsymbol{\jmath}}=1}^{\bar{\mathcal{N}}} \mathcal{L}_{\hat{\boldsymbol{\jmath}}}$ is the set of all the pairs (i, j) with $1 \leqslant i \neq j \leqslant N$, these conditions are not yet sufficient to determine $\bar{\mathcal{N}}$ and the $\mathcal{L}_{\hat{\boldsymbol{\jmath}}}$s. Nonetheless, equation (11a) can be written as

$$I_{\boldsymbol{h}} - \sum_{j=1}^{N} n_j^2 = \sum_{\hat{\boldsymbol{\jmath}}=1}^{\bar{\mathcal{N}}} \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{h} \cdot \vec{\delta}_{\hat{\boldsymbol{\jmath}}}} \sum_{(i,j) \in \mathcal{L}_{\hat{\boldsymbol{\jmath}}}} n_i n_j. \tag{11c}$$

After putting

$$\tilde{\nu}_{\hat{\boldsymbol{\jmath}}} \equiv \sum_{(i,j) \in \mathcal{L}_{\hat{\boldsymbol{\jmath}}}} n_i n_j \left(\sum_{1 \leqslant r \neq s \leqslant N} n_r n_s \right)^{-1} \qquad \mathcal{J}_{\boldsymbol{h}} \equiv \left(I_{\boldsymbol{h}} - \sum_{j=1}^{N} n_j^2 \right) \left(\sum_{1 \leqslant r \neq s \leqslant N} n_r n_s \right)^{-1} \tag{12}$$

equation (11c) becomes

$$\mathcal{J}_{\boldsymbol{h}} = \sum_{\hat{\boldsymbol{\jmath}}=1}^{\bar{\mathcal{N}}} \tilde{\nu}_{\hat{\boldsymbol{\jmath}}} \, \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{h} \cdot \vec{\delta}_{\hat{\boldsymbol{\jmath}}}}. \tag{13}$$

Remark 6. Firstly, we suppose we know the number N and the values $\{n_j\}_{j=1,\ldots,N}$ of the normalized charges (cf equation (7)). Equation (12) implies that the quantities $\mathcal{J}_{\boldsymbol{h}}$ are fully known in terms of the intensities $I_{\boldsymbol{h}} = |U_{\boldsymbol{h}}|^2$, and vice versa. Secondly, equation (13) shows that the quantities which determine (and are determined by) the intensities $\{I_{\boldsymbol{h}}\}_{\boldsymbol{h} \in \mathbb{Z}^D}$ are the $\vec{\delta}_{\hat{\boldsymbol{\jmath}}}$s, the $\tilde{\nu}_{\hat{\boldsymbol{\jmath}}}$s and $\bar{\mathcal{N}}$.

At this point it suffices to note that equation (13) has the same structure as equations (3) or (9), because $\vec{\delta}_{\hat{\boldsymbol{\jmath}}} \in \mathbb{U}^D$ (similarly to the $\boldsymbol{r}_j$) and the $\tilde{\nu}_{\hat{\boldsymbol{\jmath}}}$s (as the n_js) are positive and their sum is normalized to 1 by construction. Therefore, $\mathcal{J}_{\boldsymbol{h}}$ are the Fourier coefficients of the atomic density

$$\rho_{\mathrm{P}}(\boldsymbol{r}) = \sum_{\hat{\boldsymbol{\jmath}}=1}^{\bar{\mathcal{N}}} \tilde{\nu}_{\hat{\boldsymbol{\jmath}}} \, \delta(\boldsymbol{r} - \vec{\delta}_{\hat{\boldsymbol{\jmath}}}). \tag{14}$$

This describes a set of $\bar{\mathcal{N}}$ scattering centres with charges $\tilde{\nu}_{\hat{\boldsymbol{\jmath}}}$ set at points $\vec{\delta}_{\hat{\boldsymbol{\jmath}}}$. Hence the problem is split into two parts.

- First, from a finite subset of $\mathcal{J}_{\boldsymbol{h}}$ we must determine the number $\bar{\mathcal{N}}$, the vectors $\{\vec{\delta}_{\hat{\boldsymbol{\jmath}}}\}_{\hat{\boldsymbol{\jmath}}=1,\ldots,\bar{\mathcal{N}}}$ and charges $\{\tilde{\nu}_{\hat{\boldsymbol{\jmath}}}\}_{\hat{\boldsymbol{\jmath}}=1,\ldots,\bar{\mathcal{N}}}$. This is exactly (symbols apart) problem 2 of section 1.1.
- Second, after determining $\bar{\mathcal{N}}$, $\{\vec{\delta}_{\hat{\boldsymbol{\jmath}}}\}_{\hat{\boldsymbol{\jmath}}=1,\ldots,\bar{\mathcal{N}}}$ and $\{\tilde{\nu}_{\hat{\boldsymbol{\jmath}}}\}_{\hat{\boldsymbol{\jmath}}=1,\ldots,\bar{\mathcal{N}}}$, we must determine the atomic positions $\{\boldsymbol{r}_j\}_{j=1,\ldots,N}$ also using the prior knowledge of N and of the charges $\{n_j\}_{j=1,\ldots,N}$.

Note. We call the density ρ_{P} in equation (14) *the Patterson map* because it is simply related by equations (11a)–(13) to the Fourier transform of the intensities $I_{\boldsymbol{h}}$. Conventionally the latter, calculated only from the observed $\{I_{\boldsymbol{h}}\}_{\|\boldsymbol{h}\| < 2/\lambda}$ (see the note below equation (8b)), is called the

Patterson map (see [10, 17–19]). In the form of equation (14), ρ_{P} is infinitely resolved and hence it is an atomic density.

The solution of the second problem is purely numerical in the sense that:

Proposition 1. All the possible different *configurations* $\{\boldsymbol{r}_j\}_{j=1,\dots,N}$ *of* N *scattering centres with charges* $\{n_j\}_{j=1,\dots,N}$ *which produce a given Patterson map (see equation (14)) turn out to be determined after performing a finite number of arithmetical operations.*

Proof of proposition 1. For the proof and a precise description of this algorithm see section 3.2 of [1]. □

3. The complex amplitude problem

This and the following sections deal with problem 2 of section 1.1, the *complex amplitude problem* (CAP). For clarity, hereinafter we will always use the symbols U_h, N, n_j, $\boldsymbol{r}_j$, as in section 1.1, instead of $\mathcal{J}_h$, $\bar{\mathcal{N}}$, $\tilde{\nu}_{\hat{\jmath}}$, $\vec{\delta}_{\hat{\jmath}}$ as in section 2.2.

Since the number of the unknown n_j and $\boldsymbol{r}_j$ quantities, leaving aside momentarily N, is $(D+1)N$, from equation (9) and from the theory of implicit functions one concludes that only $(D+1)N$ USFs (Fourier coefficients) can be functionally independent, while the remaining USFs can be expressed in terms of the former. The troublesome step is the determination of the independent USFs.

3.1. The algebraic approach: case $D = 1$

To solve the CAP, useful hints are obtained from Ott's one-dimensional solution [2]. In the one-dimensional case, we can rewrite equation (9) as

$$U_h = \sum_{j=1}^{N} n_j \xi_j^h \qquad \xi_j \equiv \exp(2\pi \mathrm{i} x_j). \tag{15}$$

The ξ_js are all distinct. They are the (*unimodular*) roots of a polynomial equation of degree N given by

$$P_N(z) = \prod_{j=1}^{N}(z - \xi_j) = \sum_{\ell=0}^{N} \alpha_\ell z^{N-\ell} = 0 \tag{16}$$

where $\alpha_0 = 1$ and the other α_ℓs are the ℓth degree fully symmetrical polynomial functions [20] of $\xi_1, \dots \xi_N$. For *any* $j \in \{1, \dots, N\}$, after putting $z = \xi_j$ in the last of equalities (16), we obtain $\xi_j^N = -\alpha_N \xi_j^0 - \alpha_{N-1}\xi_j^1 - \cdots - \alpha_1 \xi_j^{N-1}$. Substituting the result in equation (15) it follows that

$$U_{N+m} = -\sum_{k=0}^{N-1} \alpha_{N-k} U_k. \tag{17}$$

In this way, all the Fourier coefficients U_k with $k \geqslant N$ (hence the whole diffraction pattern $\{U_h\}_{h\in\mathbb{Z}}$) may be determined iteratively starting from $U_0, U_1, \dots, U_{N-1}$. Otherwise, once the coefficients $\{\alpha_1, \dots, \alpha_N\}$ are known, we may as well solve the polynomial equation (16) determining $\{\xi_j\}_{j=1,\dots,N}$ as its roots. Successively the charges $\{n_j\}_{j=1,\dots,N}$ are obtained by solving the system of N linear equations obtained from equation (15) with $h = 0, 1, \dots, N-1$.

We shall now show how to determine N and coefficients $\{\alpha_k\}_{k=1,\dots,N}$. This can be done by analysing the vectorial structure underlying the CAP.

Definition 6. *Following Goedkoop [8] (see, also, [21]), consider an N-dimensional Hilbert space $\mathfrak{H}$ with an orthonormal basis $\{|e_j\rangle\}_{j=1,\dots,N}$ and the Hermitian scalar product $\langle\cdot|\cdot\rangle$. Now construct the vectors*

$$|A_{\boldsymbol{h}}\rangle \equiv \sum_{j=1}^{N} n_j^{1/2} \mathrm{e}^{2\pi \mathrm{i} \boldsymbol{r}_j \cdot \boldsymbol{h}} |e_j\rangle \qquad \text{with} \quad \boldsymbol{h} \in \mathbb{Z}^D. \tag{18}$$

With normalization in remark 1, we have

$$\langle A_{\boldsymbol{k}} | A_{\boldsymbol{h}+\boldsymbol{k}} \rangle = \langle A_{\boldsymbol{0}} | A_{\boldsymbol{h}} \rangle = U_{\boldsymbol{h}} \qquad \forall\, \boldsymbol{h}, \boldsymbol{k} \in \mathbb{Z}^D \tag{19}$$

$$\| |A_{\boldsymbol{h}}\rangle \|^2 = \langle A_{\boldsymbol{h}} | A_{\boldsymbol{h}} \rangle = U_{\boldsymbol{0}} = 1 \qquad \forall\, \boldsymbol{h} \in \mathbb{Z}^D. \tag{20}$$

Recalling definition 4, it is evident that $\{|A_{\boldsymbol{h}}\rangle\}_{\boldsymbol{h}\in\mathbb{Z}^D}$ form a Goedkoop vector lattice satisfying equation (19). In particular, we note that, for any $\boldsymbol{h} = (h_1, \dots, h_D)$, equation (18) can be written as $|A_{\boldsymbol{h}}\rangle = \mathsf{G}_1^{h_1} \dots \mathsf{G}_D^{h_D} |A_0\rangle$, where $\{\mathsf{G}_\alpha\}_{\alpha=1,\dots,D}$ are linear operators represented in the basis $\{|e_j\rangle\}_{j=1,\dots,N}$ by matrices $\mathsf{G}_\alpha = \mathrm{diag}(\exp(2\pi \mathrm{i} \boldsymbol{a}_\alpha^* \cdot \boldsymbol{r}_1), \dots, \exp(2\pi \mathrm{i} \boldsymbol{a}_\alpha^* \cdot \boldsymbol{r}_N))$ which are evidently unitary and commuting.

Note. Basis vectors $\{|e_j\rangle\}_{j=1,\dots,N}$ are the common eigenvectors of the operators $\{\mathsf{G}_\alpha\}_{\alpha=1,\dots,D}$. The degeneration of the set of eigenvalues $\{\exp(2\pi \mathrm{i} \boldsymbol{a}_\alpha^* \cdot \boldsymbol{r}_j)\}_{\alpha=1,\dots,D;\ j=1,\dots,N}$ is completely removed because the scattering centres $\{\boldsymbol{r}_j\}_{j=1,\dots,N}$ are all distinct modulo 1. Determining the scattering centres is hence equivalent to determining the eigenvalues of $\{\mathsf{G}_\alpha\}_{\alpha=1,\dots,D}$.

Now consider the ordered index sequence $\{\boldsymbol{h}_1, \dots, \boldsymbol{h}_q\}$. Take the associated vector sequence $\{|A_{\boldsymbol{h}_1}\rangle, |A_{\boldsymbol{h}_2}\rangle, \dots, |A_{\boldsymbol{h}_q}\rangle\}$ and form the corresponding Gram matrix [22] $(\mathcal{U}_q)$, whose entries are

$$(\mathcal{U}_q)_{ij} = \langle A_{\boldsymbol{h}_i} | A_{\boldsymbol{h}_j} \rangle = U_{\boldsymbol{h}_j - \boldsymbol{h}_i}. \tag{21}$$

The matrix in equation (21) (cf also definition 3) has different denominations. According to the first expression, it is a Gram matrix, and according to the second it is a Toeplitz matrix (for $D > 1$, a generalized Toeplitz matrix). This double identification is the strongest way to characterize (cf section 1.2) the Fourier coefficients of a positive atomic density. Moreover, its determinant is a generalized Vandermonde determinant [23]. Finally, according to crystallographers, it is a Karle–Hauptman (KH) matrix after their basic work [7]. We will say that the matrix $(\mathcal{U}_q)$ is *generated* by the sequence $\{\boldsymbol{h}_1, \dots, \boldsymbol{h}_q\}$.

Definition 7. *If the KH matrix generated by a set of indices $\mathcal{B}$ is non-singular while that generated by $\mathcal{B} \cup \boldsymbol{h}_0$ is singular, $\boldsymbol{h}_0$ will be said a* KH zero *for set $\mathcal{B}$.*

Returning to the case $D = 1$, now consider sequences $\{\boldsymbol{h}_1, \boldsymbol{h}_2, \dots, \boldsymbol{h}_q\} = \{0, 1, \dots, q-1\}$. Let us override the trivial cases $q = 1$ or 2. According to a basic property of Gram matrices [22], we have $\det(\mathcal{U}_3) \neq 0$ if and only if the three generating vectors are linearly independent. Assume that this condition is met and enlarge the previous set including the vector $|A_{\boldsymbol{h}_4}\rangle = |A_3\rangle$. Denote by $(\mathcal{U}_4)$ the resulting Gram matrix. Assume that $\det(\mathcal{U}_4) \neq 0$ and iterate the procedure by adding, at each step, the next vector $|A_{\boldsymbol{h}_{q+1}}\rangle = |A_q\rangle$ with $q = 4, 5, \dots$. As q increases, a singular matrix (or a KH zero) must necessarily be found because the space $\mathfrak{H}$ has dimension N. The crucial point is that, having always chosen $\boldsymbol{h}_q = q - 1$, the order of the first singular matrix will be *exactly* $N + 1$. This is evident from the analytical expression of the determinant of $(\mathcal{U}_q)$ given by equation (A1) in the appendix. In this way, N can be determined. At this point it is possible to determine the α_j coefficients from $(\mathcal{U}_N)$, the largest non-singular matrix, by the relations (see equations (2.16), (2.43*b*) and (2.45) of [1]), as

$$\alpha_{N-(j-1)} = \sum_{i=1}^{N} (\mathcal{U}_N^{-1})_{ji} U_{N+1-i} \qquad j = 1, \dots, N. \tag{22}$$

3.2. The algebraic approach: case $D > 1$

Definition 8. *We will speak of the linear dependence or independence of a set of indices* $\{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_m\}$ *meaning the linear dependence or independence of the associated Goedkoop vectors* $\{|A_{\boldsymbol{h}_1}\rangle, \ldots, |A_{\boldsymbol{h}_m}\rangle\}$. *We will then call* base set *a set of distinct, linearly independent indices* $\mathcal{B} = \{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_N\}$ *having the largest cardinality, so that the associated Goedkoop vectors* $\{|A_{\boldsymbol{h}_1}\rangle, \ldots, |A_{\boldsymbol{h}_N}\rangle\}$ *form a base of* $\mathfrak{H}$.

An algebraic solution of the CAP is possible, *whatever D*, when an algorithm has been devised which is able to single out a base set. The dimension of the space $\mathfrak{H}$ is not known *a priori*, even if we know that it is finite. In general, knowing only the scalar products of general vectors, it would not be possible to determine its dimension with a finite number of operations. However, the properties of Goedkoop's vector lattice and of the related system of polynomial equations (see section 6) makes the answer positive in this case.

The procedures, presented here to single out a base set, simultaneously also determine a complete set. In fact, the determination of a base set needs the knowledge of several KH matrices which in turn contain several Fourier coefficients. The same coefficients are a complete set $\mathcal{C}$ in the sense of definition 2. For the proof see proposition 3 in section 6. Let us now describe exactly which Fourier coefficients are needed.

Base sets are obtained by enlargement procedures like that reported in section 3.1. Some singular KH matrices (i.e. some KH zeros) will be found while $\mathcal{B}$ is singled out. The set of these KH zeros forms the *complementary set* of $\mathcal{B}$, denoted by $\mathcal{B}^c = \{\boldsymbol{h}'_1, \ldots, \boldsymbol{h}'_M\}$, in the order in which they have been found. Take $\boldsymbol{h}'_k$ as the kth element of $\mathcal{B}^c$. Let $\mathcal{B}^-(k)$ denote the subset of $\mathcal{B}$ formed by those indices which were included in $\mathcal{B}$ *before* finding the KH zero $\boldsymbol{h}'_k$, and let $\mathcal{B}^+(k) \equiv \mathcal{B}^-(k) \cup \{\boldsymbol{h}'_k\}$. Now we can explain the first property claimed in remark 5. It is sufficient to substitute the sets denoted there by $\mathcal{S}_k$ with the sets $\mathcal{B}^+(k)$. In fact, the linear independent subsets $\mathcal{B}^-(k)$ obey $\mathcal{B}^-(k) \subset \mathcal{B}^-(k+1)$ by construction. Therefore, $\boldsymbol{h}'_k$ is the only element of $\mathcal{B}^+(k)$ which does not belong to $\mathcal{B}^+(k+1)$, so $\mathcal{B}^+(k) \setminus (\mathcal{B}^+(k) \cup \mathcal{B}^+(k+1)) = \{\boldsymbol{h}'_k\}$. On the other hand, as $\mathcal{B}^+(k) \cup \mathcal{B}^+(k+1) = \mathcal{B}^-(k)$, the set $\mathcal{B}^+(k+1) \setminus (\mathcal{B}^+(k) \cup \mathcal{B}^+(k+1))$ is not empty as it contains at least $\boldsymbol{h}'_{k+1}$. Note that to establish the linear independence of the set $\mathcal{B}^-(k)$ and the linear dependence of the set $\mathcal{B}^+(k)$ we have to verify that the KH matrix generated by the first set is non-singular and that generated by the second set is singular. Hence, for every $k = 1, \ldots, M$, we need to know all the Fourier coefficients (or USFs) which enter these matrices, that is to say all the Fourier coefficients with index in $\mathcal{B}^+(k) - \mathcal{B}^+(k)$. A complete set is hence given by

$$\mathcal{C} = \bigcup_{k=1}^{M} \big(\mathcal{B}^+(k) - \mathcal{B}^+(k)\big). \tag{23}$$

Remark 7. We remark now that:

- sets $\mathcal{B}^c$ and $\mathcal{C}$ are uniquely determined by $\mathcal{B}$;
- given an atomic density, base sets and complete sets are not unique, but the cardinality of a base set, in contrast with that of a complete set, is unique.

The radius $L_{\mathcal{C}}$ of a complete set and its cardinality depend on the selected base set $\mathcal{B}$. In the presence of a high-frequency cut-off L (cf section 1.1, questions 4 and 5) we must require $L_{\mathcal{C}} < L$. For this reason, among the possible procedures for singling out base sets, those yielding complete sets with smaller radius are the most useful ones.

We now sketch the search algorithm reported in [1] for the case $D = 2$ so as to better clarify the concepts stated above.

Definition 9. *We will denote by $\mathcal{Q}_{\boldsymbol{h}}$ the positive quadrant of $\mathbb{Z}^2$ originating at $\boldsymbol{h}_0$,*

$$\mathcal{Q}_{\boldsymbol{h}_0} \equiv \{\boldsymbol{h} = (h, k) \mid h \geqslant h_0,\ k \geqslant k_0\}.$$

This algorithm performs an *axial search* in the first quadrant $\mathcal{Q}_{\mathbf{0}}$. The limitation to $\mathcal{Q}_{\mathbf{0}}$ is legitimate because the difference set $\mathcal{B} - \mathcal{B}$ is left invariant by a translation $\mathcal{B} \longmapsto \mathcal{B} + \boldsymbol{h}$. *Axial search* means that the base set is enlarged preferably along one of the two axes of $\mathbb{Z}^2$, $\boldsymbol{a}_1^* \equiv (1, 0)$ or $\boldsymbol{a}_2^* \equiv (0, 1)$ (in [1] they were denoted by $\boldsymbol{a}^*$ and $\boldsymbol{b}^*$).

Notation. As different search directions generally yield different base sets, when we wish to specify the chosen search direction $\boldsymbol{a}_\alpha^*$ we will write $\mathcal{B}^{\boldsymbol{a}_\alpha^*}$.

Let us choose, for definiteness, the axis $\boldsymbol{a}_1^*$. As in the one-dimensional case, one successively includes in $\mathcal{B}$ indices $\boldsymbol{h}_r = (r - 1)\boldsymbol{a}_1^* = (r - 1, 0)$ with $r = 1, 2, 3, \ldots$ until we find the first KH zero at, say, $r = m_1 + 1$. At this point, index $\boldsymbol{h}_{m_1+1} = (m_1, 0)$ is included in $\mathcal{B}^c$ rather than in $\mathcal{B}$. Then we continue enlarging $\mathcal{B}$ by including indices $(r, 1)$, $r = 0, 1, 2, \ldots$, lying on the next upper half-axis parallel to $\boldsymbol{a}_1^*$. This continues until

- *either* after including an index $(m_2, 1)$ with $m_2 < m_1$, we find another singular matrix. Then, $(0, 1), \ldots, (m_2 - 1, 1)$ are included in $\mathcal{B}$ and $(m_2, 1)$ in $\mathcal{B}^c$;
- *or* we end up including index $(m_1 - 1, 1)$ without finding a KH zero. As it must necessarily be $m_2 \leqslant m_1$ (cf remark 8, proposition 2) at this point we may avoid examining index $(m_2, 1)$ (*it must be* a KH zero) and also omit including it in $\mathcal{B}^c$ as its knowledge is 'redundant'.

Then we pass to the next upper half-line and so on.

Remark 8. With this procedure (see proposition 2) every KH zero has a broader meaning. If $\boldsymbol{h}'_k$ is the kth KH zero found and $\mathcal{B}^-(k)$ is the portion of the base set isolated previously, it means not only that $\boldsymbol{h}'_k$ is linearly dependent on $\mathcal{B}^-(k)$ but *all of the positive quadrant* $\mathcal{Q}_{\boldsymbol{h}'_k}$ (see definition 9) is linearly dependent on $\mathcal{B}^-(k)$. Hence the whole quadrant $\mathcal{Q}_{\boldsymbol{h}'_k}$ may be excluded from a further search. A weaker version (property 1) of this property will be used to construct more general base sets (cf section 5). A consequence is that $\mathcal{B}$ is delimited by the positive half-axes $\boldsymbol{a}_1^*$ and $\boldsymbol{a}_2^*$ and by a staircase connecting them. In fact, the lengths of the rows of $\mathcal{B}$ form a decreasing sequence, $m_1 \geqslant m_2 \geqslant \cdots \geqslant m_p$ (and likewise for the columns).

The inclusion procedure terminates once we arrive at a certain $(p + 1)$th upper half-axis such that its *first* index $(0, p + 1)$ turns out to be a KH zero for the set $\mathcal{B}$ just singled out. At this point $\mathcal{B}$ is a base set, $(0, p + 1)$ is added to the complementary set $\mathcal{B}^c$.

For the reason explained in remark 8, $\mathcal{B}$ will be called the *(first quadrant) ladder base set* in the sense defined below.

Definition 10. *A* ladder set *in the first quadrant of $\mathbb{Z}^2$ is defined as a union of p decreasing rows $\bigcup_{k=1}^{p}\{(h - 1, k - 1) \mid k = 1, \ldots, p, k = 1, \ldots, m_k\}$ with $m_1 \geqslant \cdots \geqslant m_p$. Equivalently one may speak of the union of $q = m_1$ decreasing columns. A similar definition can be given for the other three quadrants. A* centred *ladder set in $\mathbb{Z}^2$ may then be defined as the union of ladder sets in the four quadrants. The first quadrant ladder sets will be indicated simply as ladder sets. We will use a more detailed notation for ladder sets in $\mathbb{Z}^2$. When we need to specify the number of elements contained in the rows (ordered bottom-up) of $\mathcal{B}$, we shall write $\mathcal{B}^{\boldsymbol{a}_1^*}[m_1 \ldots m_p]$. Alternatively, if we wish to specify the elements present in the columns (ordered left-to-right), we shall write $\mathcal{B}^{\boldsymbol{a}_1^*}\{\mu_1 \ldots \mu_q\}$ where $q = m_1$, $\mu_1 = p$ and $\mu_j = \max\{i | m_i \geqslant j\}$, $j = 1, \ldots, q$.*

Proposition 2. *The structure of the base set $\mathcal{B}$ is closely related to the geometry of the support of ρ. Consider the lengths of the rows of $\mathcal{B}^{a_1^*}[m_1 \dots m_p]$. m_1 is equal to the number of* distinct *projections $\boldsymbol{a}_1^* \cdot \boldsymbol{r}_j$ of the scattering centres $\{\boldsymbol{r}_j\}_{j=1,\dots,N}$. Clearly, if the projections are all distinct, then $m_1 = N$, $p = 1$ and the base set is linear. If some projections coincide, m_2 is the number of the* distinct *$\boldsymbol{a}_1^*$-projections common to two or more scattering centres, m_3 is the number of* distinct *$\boldsymbol{a}_1^*$-projections common to three or more scattering centres, and so on. Conversely, $\mu_1 \dots$ denote the multiplicities of the $\boldsymbol{a}_1^*$-projections, in decreasing order.*

Proof. The proof is given in [1], section 1.2.2 and appendix C. □

Proposition 2 has one important consequence. Unless many of the projections along the chosen search direction coincide, the resulting base set (and hence the associated complete set) will be very elongated along the search direction. Of course, one may try another of the D possible search directions, but the resulting set will probably be very long. A search procedure which is equally simple and effective while yielding more compact base sets is hence preferable. The next two sections are devoted to the construction of more compact base sets.

Note. If for some $1 \leqslant \alpha \leqslant D$ the projections $\{\boldsymbol{r}_j \cdot \boldsymbol{a}_\alpha^*\}_{j=1,\dots,N}$ are all *distinct*, $m_1 = N$ and the base set $\mathcal{B}$ is linear along $\boldsymbol{a}_\alpha^*$, i.e. $\mathcal{B} = \{k\boldsymbol{a}_\alpha^* \mid k = 0, \dots, N-1\}$. Let us consider other directions $\hat{\boldsymbol{\sigma}} = (n_1, \dots, n_D) \in \mathbb{Z}^2$, $\hat{\boldsymbol{\sigma}} \neq \boldsymbol{a}_\alpha^*$ for all α. We can define the orthogonal *modular* projections $\Pi_j(\hat{\boldsymbol{\sigma}}) \equiv \boldsymbol{r}_j \cdot \hat{\boldsymbol{\sigma}} \bmod 1$. The modulo operation is introduced because $|\hat{\boldsymbol{\sigma}}|^2 = \hat{\boldsymbol{\sigma}} \cdot \hat{\boldsymbol{\sigma}} > 1$. Let us ask: is it always possible to choose a direction $\hat{\boldsymbol{\sigma}}$ in $\mathbb{Z}^D$ such that the orthogonal *modular* projections $\boldsymbol{r}_j \cdot \hat{\boldsymbol{\sigma}} \bmod 1$ are all distinct and hence (cf [1], appendix B) is a linear set $\mathcal{B} = \{k\hat{\boldsymbol{\sigma}} \mid k = 0, \dots, N-1\}$ always a base set? The answer is negative. The following example makes this point clear in the case $D = 2$. Consider, in fact, an atomic density consisting of four scattering centres $(0, 0)$, $(0, \frac{1}{2})$, $(\frac{1}{2}, 0)$ and $(\frac{1}{2}, \frac{1}{2})$. It is easy to check that the modular projections of these positions along any $\hat{\boldsymbol{\sigma}} \in \mathbb{Z}^2$ always yield *two* different values. Thus, the linearly independent sets lying on any line of $\mathbb{Z}^2$ always contain only two elements and therefore cannot be base sets.

4. A property of the ladder sets

As already mentioned in section 3, the KH matrix generated by the set of indices $\mathcal{B}$ involves all the USFs relevant to indices in $\mathcal{B} - \mathcal{B}$. Thus, for singling out small complete sets, it is convenient to consider $\mathcal{B}$ sets as compact as possible and made up of indices close to the origin.

Definition 11. *Consider a generic (first-quadrant)* ladder set $\mathcal{B}[m_1 \dots m_p]$ *(see definition 10). Then define the extremal corner points $\boldsymbol{h}_0'$ and $\boldsymbol{h}_0''$ as, respectively, the index at the outermost right of the top row and the index at the top of the rightmost column of $\mathcal{B}[m_1 \dots m_p]$.*

One has the important property (recall definition 9):

Property 1. *If $\mathcal{B} = \mathcal{B}[m_1 \dots m_p]$ is a linearly dependent ladder set, i.e. $\mathcal{B}$ contains a KH zero, then all the vectors with an index in $\mathcal{Q}_{\boldsymbol{h}_0'}$ (or in $\mathcal{Q}_{\boldsymbol{h}_0''}$) are linear combination of the vectors with an index in $\mathcal{Q}_{\boldsymbol{0}} \setminus \mathcal{Q}_{\boldsymbol{h}_0'}$ (or in $\mathcal{Q}_{\boldsymbol{0}} \setminus \mathcal{Q}_{\boldsymbol{h}_0''}$).*

Proof. Property 1 clearly holds true for $|A_{\boldsymbol{h}_0'}\rangle$. Now translate $\mathcal{B}[m_1 \dots m_p]$ by $(1, 0)$. The translated set shares with set $\mathcal{Q}_{\boldsymbol{h}_0'}$ only the indices $\boldsymbol{h}_0'$ and $(\boldsymbol{h}_0' + \boldsymbol{a}_1^*)$, while its remaining indices lie in $\mathcal{Q}_{\boldsymbol{0}} \setminus \mathcal{Q}_{\boldsymbol{h}_0'}$. The translated set is also linearly dependent. Thus, $|A_{\boldsymbol{h}_0'+(1,0)}\rangle$ can be written as a linear combination of $|A_{\boldsymbol{h}_0'}\rangle$ and of the vectors associated with the remaining indices of

the translated set. As noted above, these indices belong to $\mathcal{Q}_{\mathbf{0}} \setminus \mathcal{Q}_{\boldsymbol{h}'_0}$, and since $|A_{\boldsymbol{h}'_0}\rangle$ is a linear combination of vectors with indices in $\mathcal{Q}_{\mathbf{0}} \setminus \mathcal{Q}_{\boldsymbol{h}'_0}$, one also concludes that $|A_{\boldsymbol{h}'_0+(1,0)}\rangle$ is a linear combination of vectors with indices in $\mathcal{Q}_{\mathbf{0}} \setminus \mathcal{Q}_{\boldsymbol{h}'_0}$. In the same way, by taking $m = 2, 3, \ldots$, one shows recursively that each $|A_{\boldsymbol{h}'_0+(m,0)}\rangle$ is a linear combination of vectors of $\mathcal{Q}_{\mathbf{0}} \setminus \mathcal{Q}_{\boldsymbol{h}'_0}$. By doing so, all the vectors with index of the lowest row of $\mathcal{Q}_{\boldsymbol{h}'_0}$ are proven to be linear combinations of vectors with index in $\mathcal{Q}_{\mathbf{0}} \setminus \mathcal{Q}_{\boldsymbol{h}'_0}$. Actually they depend only on the elements belonging to $\mathcal{Q}_{\mathbf{0}} \setminus \mathcal{Q}_{\boldsymbol{h}'_0}$ and to the rows going from zero to the value of the row where $\boldsymbol{h}'_0$ is located. Now translate $\mathcal{B}$ by (0, 1). The resulting set shares with $\mathcal{Q}_{\boldsymbol{h}'_0}$ the elements $\boldsymbol{h}'_0$, $\boldsymbol{h}'_0+(0, 1)$ and possibly some more $\boldsymbol{h}'_0 + (m, 0)$ (with $m > 0$) resulting from the translation by (0, 1) of $\mathcal{B}$. The vectors associated with $(\boldsymbol{h}'_0 + (m, 0))$ with $m \geqslant 0$ already being determined, it follows that $|A_{\boldsymbol{h}'_0+(0,1)}\rangle$ is a linear combination of vectors associated with indices of $\mathcal{Q}_{\mathbf{0}} \setminus \mathcal{Q}_{\boldsymbol{h}'_0}$. Again, by considering the further successive translations by $(m, 0)$ with $m = 1, 2, \ldots$, property 1 is proven for all the vectors with index in the second row of $\mathcal{Q}_{\boldsymbol{h}'_0}$ and, by iteration, for all the vectors associated with $\mathcal{Q}_{\boldsymbol{h}'_0}$. The proof for $\mathcal{Q}_{\boldsymbol{h}''_0}$ is identical. □

In conclusion, property 1 amounts to saying that finding a KH zero of a ladder set is equivalent to stating a condition of *known linear dependence* for the *countable infinity* of vectors associated with the positive quadrant with origin on $\boldsymbol{h}'_0$ (or $\boldsymbol{h}''_0$).

5. Construction of ladder base sets in $\mathbb{Z}^2$

Property 1 makes the construction of ladder base sets straightforward once an inclusion procedure has been devised which at each step ensures that the enlarged set is ladder-shaped. Each step of the searching procedure set consists in enlarging a linearly independent set $\mathcal{B}$ by including a further element and in evaluating the determinant of the associated KH matrix. If the determinant is not zero, the previous procedure is iterated starting from the set $\mathcal{B}$ just enlarged. In contrast, if the determinant is zero, the last element is stored separately in $\mathcal{B}^c$, a quadrant is excluded from further search, and then we can iterate the step with an element not yet considered. The search is complete when the set cannot be further enlarged. Therefore, it is important to find a condition ensuring that the considered set cannot be further enlarged.

First-quadrant square search

As a translated base set is still a base set we may restrict the search to the first quadrant of $\mathbb{Z}^2$.

Consider the following algorithm. Start from $\mathbf{0}$ and include further elements in such a way that the resulting set is as close as possible to a *square*. More explicitly, consider elements in the following order:

$$\begin{array}{l} (0, 0); \\ (1, 0),\ (0, 1),\ (1, 1); \\ (2, 0),\ (0, 2),\ (2, 1),\ (1, 2),\ (2, 2); \\ \vdots \end{array} \tag{24}$$

Remark 9. This search order guarantees that, at each step, the set is a (first-quadrant) ladder set (cf definition 10) and the last added index is *always* one of its two *extremal corner points* $\boldsymbol{h}'_0$ or $\boldsymbol{h}''_0$ (cf definition 11).

At each step, the set is ladder-shaped and one evaluates the relevant KH determinant. Let $\boldsymbol{h}_0^{(1)}$ be the first KH zero found. Owing to remark 9, either $\boldsymbol{h}_0^{(1)} = \boldsymbol{h}'_0$ or $\boldsymbol{h}_0^{(1)} = \boldsymbol{h}''_0$. Assume, for definiteness, that the first case occurs. We now use property 1. Then we may exclude all

indices lying in $\mathcal{Q}_{\boldsymbol{h}_0^{(1)}}$ from our search. Hence we carry on the inclusion procedure remaining in the region $\mathcal{Q}_{\mathbf{0}} \setminus \mathcal{Q}_{\boldsymbol{h}_0^{(1)}}$. Let $\boldsymbol{h}_0^{(2)}$ denote the next KH zero. Once again either $\boldsymbol{h}_0^{(2)} = \boldsymbol{h}_0'$ or $\boldsymbol{h}_0^{(2)} = \boldsymbol{h}_0''$. In both cases, the indices of $\mathcal{Q}_{\boldsymbol{h}_0^{(2)}}$ may be excluded from further search. More precisely, owing to property 1 in the appendix, we can restrict a further search to the indices of the region $\mathcal{Q}_{\mathbf{0}} \setminus (\mathcal{Q}_{\boldsymbol{h}_0^{(1)}} \cup \mathcal{Q}_{\boldsymbol{h}_0^{(2)}})$. At this point, the upper border of the region containing linearly independent indices presents one finite step. The finiteness of N ensures that, after a finite number of iterations, the lower boundary of the excluded region will be a staircase connecting the vertical to the horizontal axis. In this way, no further indices of $\mathcal{Q}_{\mathbf{0}}$ can be added to the $\mathcal{B}$ set. Thus, set $\mathcal{B}$ will be a ladder base set and its cardinality will yield the *a priori* unknown N value. The complementary set $\mathcal{B}^c$ consists of all the KH zeros found during the aforementioned search and the complete set $\mathcal{C}$ is obtained as in equation (23).

This search procedure was reported, as a numerical example, at the end of section 1.2.2 of [1] suggesting that it would *always* yield a base set. This claim is now proven. It is worth noticing that a procedure obtained from the one just described by exchanging the role of the two axes in equation (24) will generally lead to a different base set.

Even though the inclusion procedures reported above do not choose at each step the element closest to the origin, nonetheless the elements are still chosen among the ones closer to the origin. Thus, it can confidently be expected that the radius of the resulting complete set should be close to the minimal one. This is a noticeable improvement with respect to the procedure reported in section 3.2.

Centred-square search

We now report another searching procedure which acts more symmetrically than the one just described. It consists in looking for a base set which is as close as possible to a square *centred* at the origin of reciprocal space. The inclusion proceeds as follows:

$(0, 0)$;
$(1, 0), (0, 1), (-1, 0), (0, -1)$; $(1, 1), (-1, 1), (-1, -1), (1, -1)$;
$(2, 0), (0, 2), (-2, 0), (0, -2)$; $(2, 1), (-1, 2), (-2, -1), (1, -2)$; $(2, 2), \ldots$.

The resulting set is a centred ladder set (cf definition 10). We now need to generalize definitions 9 and 11.

Definition 12. *The region $\mathcal{Q}_{\boldsymbol{h}^*}$ associated with $\boldsymbol{h}^* \equiv (h^*, k^*)$ is defined as follows:*

- *if $h^*k^* \neq 0$, $\mathcal{Q}_{\boldsymbol{h}^*}$ is the quadrant*

  $$\mathcal{Q}_{\boldsymbol{h}^*} \equiv \left\{ \boldsymbol{h} = (h, k) \;\middle|\; |h| \geqslant |h^*|,\ |k| \geqslant |k^*|,\ \operatorname{sign}(hh^*) = 1,\ \operatorname{sign}(kk^*) = 1 \right\}$$

  internal to the quadrant of $\mathbb{Z}^2$ where $\boldsymbol{h}^*$ lies;
- *otherwise if $\boldsymbol{h}^*$ lies on one of the axes, say $\boldsymbol{h}^* = (h^*, 0)$, $\mathcal{Q}_{\boldsymbol{h}^*}$ is a half-plane*

  $$\mathcal{Q}_{\boldsymbol{h}^*} \equiv \left\{ \boldsymbol{h} = (h, k) \;\middle|\; |h| \geqslant |h^*|,\ \operatorname{sign}(hh^*) = 1,\ |k| \geqslant 0 \right\}.$$

Definition 13. *Consider a generic centred ladder set $\mathcal{B}$ (cf definition 10). For each of the $2^D = 4$ quadrants $\mathsf{Q}_1 \equiv \{(h \geqslant 0, k \geqslant 0)\}$, $\mathsf{Q}_2 \equiv (h \leqslant 0, k \geqslant 0)$, $\mathsf{Q}_3 \equiv (h \leqslant 0, k \leqslant 0)$, $\mathsf{Q}_4 \equiv (h \geqslant 0, k \leqslant 0)$, we can define $D = 2$* extremal corner points *$\boldsymbol{h}'_{\mathsf{Q}_i}$ and $\boldsymbol{h}''_{\mathsf{Q}_i}$. $\boldsymbol{h}'_{\mathsf{Q}_i}$ will be the index $(h_{i;1}, h_{i;2})$ having*

$$|h_{i;2}| = \max \left\{ |\boldsymbol{h} \cdot \boldsymbol{a}_2^*| \mid \boldsymbol{h} \in \mathcal{B} \cap \mathsf{Q}_i \right\}$$
$$|h_{i;1}| = \max \left\{ |\boldsymbol{h} \cdot \boldsymbol{a}_1^*| \;\middle|\; \boldsymbol{h} \in \mathcal{B} \cap \mathsf{Q}_i;\ |\boldsymbol{h} \cdot \boldsymbol{a}_2^*| = h_{i;2} \right\}.$$

An analogous definition holds for $\boldsymbol{h}''_{\mathsf{Q}_i}$ exchanging the maximization order.

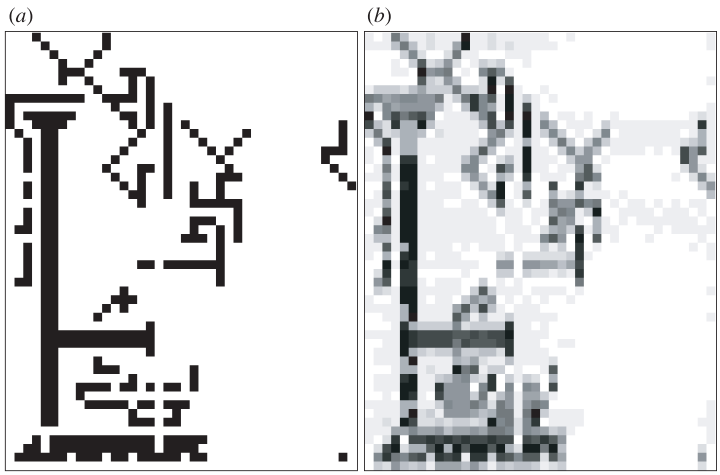

Figure 3. (*a*) A 40 × 50 pixel image. A 256-level greyscale is used (black = 255, white = 0). By attributing the grey-level value to the centre of each pixel, and with some simple rescaling, the image can be encoded as an atomic density ρ (cf definition 1) and its Fourier coefficients can be evaluated (cf equation (3)). In this case, as only 344 points have a non-zero grey level, ρ is formed by $N = 344$ scattering centres. (*b*) A Fourier reconstruction of the image in (*a*) using the Fourier coefficients inside the cut-off (ellipse) in figure 2. Note the loss of resolution. In contrast, image (*a*) can be *exactly* reconstructed from the complete set shown in figure 2.

Using these two definitions, property 1 can be straightforwardly generalized. Now note that every last added index is one of the eight extremal corner points. Having made all the due generalizations, all the considerations made for the *first-quadrant square search* apply to this procedure. This procedure will determine a base set restricted to a region containing the origin. Due to the greater symmetry of the search, the resulting complete set will generally have a smaller radius with respect to any of the procedures reported before.

An example of centred-square search is given in figure 2. The density ρ is illustrated in figure 3(*a*) as a 40 × 50-pixel greyscale image (black = 255, white = 0). The image is encoded as an atomic density (cf definition 1), with some simple rescaling, by attributing the grey-level value to the centre of each pixel and appropriately rescaling coordinates and weights. The Fourier coefficients are evaluated by (3). The complete set is considerably smaller than that obtained from the same density by the algorithm of section 3.2, which is shown in figure 1. Figure 2 also shows an ellipse corresponding to the smallest high-resolution cut-off that contains the complete set $\mathcal{C}$. Figure 3(*b*) shows the image obtained by calculating the inverse Fourier transform with such a cut-off. The resolution is evidently rather poor. In contrast, we stress that an *exact* reconstruction of the image of figure 3(*a*) can be performed as explained in section 6 using only the Fourier coefficients with index in $\mathcal{C}$.

Spherical search procedure

Both the *square* and the *centred-square* search procedures do not ensure that the radius of the resulting complete set is really minimal. Anyhow, they yield complete sets of radius close to the minimal one. Their simplicity makes them well suited to problems where the high-resolution cut-off is large enough to allow some freedom. For the sake of proof we describe now a *spherical* construction procedure which might well yield the minimal radius but loses much in terms of simplicity. It might be useful as a last resort after one of the former algorithms

has failed to isolate a complete set for a given cut-off L. The spherical procedure runs as for the former except that the sequence of indices is ordered by increasing radius. For instance, with the Euclidean metric ($\mathsf{M}^* = \mathbf{1}$), it may proceed as

$[(0,0)]$
$[(0,1),(1,0),(0,-1),(-1,0)],\ [(1,1),(-1,1),(-1,-1),(1,-1)]$
$[(2,0),\dots],\ [(2,1),\dots],\ [(2,2),\dots]$
$[(3,0),\dots],\ [(3,1),\dots],\ [(3,2),\dots],\ [(4,0),\dots],\ [(3,3),\dots],\ \dots.$

The square brackets contain indices related by a $\pi/2$ rotation. Note that $(4,0)$ comes before $(3,3)$ because $3^2+3^2 > 4^2$. This implies that the last added index is not always coincident with an extremal corner point (cf definition 13). Let us call $\mathcal{B}$ the portion of a base set isolated at a certain point of the procedure and let $\boldsymbol{h}^\triangle$ be the last added index, assumed to be different from the extremal corner points of its quadrant. For simplicity we will suppose it to be in the first quadrant ($\boldsymbol{h}^\triangle \in \mathsf{Q}_1$) and we will use the simpler notation $\boldsymbol{h}'_0$, $\boldsymbol{h}''_0$ for the extremal corner points (as in definition 11) instead of the more complex notation of definition 13. Suppose now that the inclusion of $\boldsymbol{h}^\triangle$ makes the set linearly dependent. We cannot straightforwardly exclude $\boldsymbol{h}^\triangle$ and the related quadrant $\mathcal{Q}_{\boldsymbol{h}^\triangle}$. Now, the linear dependence is a collective property and one may always exclude $\boldsymbol{h}'_0$ or $\boldsymbol{h}''_0$ (and the related quadrant) *instead of* the last added index $\boldsymbol{h}^\triangle$ when they do not coincide. The trouble at this point is that the exclusion of $\boldsymbol{h}'_0$ or $\boldsymbol{h}''_0$ instead of $\boldsymbol{h}^\triangle$ *does not ensure* that the resulting set is linearly independent. In fact, it may be that $\{\boldsymbol{h}^\triangle\} \cup \mathcal{B}'$ (with $\mathcal{B}'$ properly contained in $\mathcal{B}$) is linearly dependent. In this case the index (and related quadrant) to be excluded would be an extremal corner point *of the subset* $\mathcal{B}'$.

Anyhow, we now show how to isolate $\mathcal{B}'$ by a finite set of operations. First, we choose an extremal corner point of $\mathcal{B}$, say $\boldsymbol{h}'_0$ for definiteness. Start from $\boldsymbol{h}'_0$ and perform the following test. If $\mathcal{B} \cup \{\boldsymbol{h}^\triangle\}$ is linearly dependent (as supposed) and $(\mathcal{B} \cup \{\boldsymbol{h}^\triangle\}) \setminus \{\boldsymbol{h}'_0\}$ is linearly independent then $\boldsymbol{h}^\triangle$ is included in $\mathcal{B}$, the quadrant $\mathcal{Q}_{\boldsymbol{h}'_0}$ is excluded and the algorithm can restart. If not, we must begin deleting indices from $\mathcal{B}$. We delete first $\boldsymbol{h}'_0$ and we proceed deleting 'backwards', or right to left and top to bottom when a row is empty. Before deleting each new point we perform the same test as before (with $\mathcal{B}$ substituted by $\mathcal{B}'$ which is what remains of $\mathcal{B}$ after the deletions). By construction, for this reduced set $\mathcal{B}'$ the point under test is always an extremal corner point, just like $\boldsymbol{h}'_0$. As soon as we find a point $\boldsymbol{h}^*$ for which the test is successful, we can exclude the quadrant $\mathcal{Q}_{\boldsymbol{h}^*}$. The quadrant $\mathcal{Q}_{\boldsymbol{h}^*}$ can be excluded by using the subset $\mathcal{B}'$ instead of $\mathcal{B}$ for enforcing property 1. In the worst case we will arrive at the point $\boldsymbol{h}^\triangle$. In such a case the test *must* be successful and this subprocess will always reach an end point. Excluding $\boldsymbol{h}^\triangle$, in fact, we are sure to obtain a linearly independent set.

Successively, we must re-examine all the formerly deleted points. So we proceed back from $\mathcal{Q}_{\boldsymbol{h}^*}$ (left to right and bottom-up) retracing our steps (except the newly excluded quadrant(s)) up to $\boldsymbol{h}'_0$. For each re-included point we test the KH determinant and when we find a KH zero we may exclude a new quadrant. After exhausting this second subprocess, we may go back to the main cycle and restart the 'spherical' inclusion.

Comment. Even though the search procedure ensures that each newly added index is closest to the origin, the conclusion that the final base set has the smallest radius is not rigorous. The final shape of the base set may depend on the order in which indices with equal radius are considered. Anyhow, the number of different ways of including these indices is finite. The base set with the smallest radius will be found by considering all of these ways.

6. Reconstruction procedure

Clearly, any density $\rho : \mathbb{U}^D \longmapsto \mathbb{C}$ and its IDP $\{U_{\boldsymbol{h}}\}_{\boldsymbol{h}\in\mathbb{Z}^D}$ are related by a duality relation, in the sense that the knowledge of ρ determines the IDP via equation (3) and, vice versa, the knowledge of the IDP determines ρ via equation (4). For a positive atomic density, as reported below *question 6*, the duality can be restricted to ρ and any complete set $\mathcal{C}$. In fact, it holds that:

Proposition 3. *A complete set $\mathcal{C}$ as defined in section 3.2, that is the minimal set of indices needed to be known in order to isolate a ladder base set, is also a complete set as in definition 2: the knowledge of $U_{\boldsymbol{h}}$ with $\boldsymbol{h} \in \mathcal{C}$ is sufficient to determine all the remaining ones.*

Proof. The proof of a weaker statement is reported in [1], section 2.4. In that case, simplicity was gained by allowing a bit of redundancy. The difference is the definition of the *complementary set* (and, consequently, that of the associated *complete* set associated with a given *base* set). In the less precise form of [1], the complementary set (let us denote this variant by $\mathcal{B}^c_{[1]}$) is defined as $\mathcal{B}^c_{[1]} \equiv \{\boldsymbol{h} + \boldsymbol{a}^*_\alpha \mid \boldsymbol{h} \in \mathcal{B},\ \alpha = 1, \ldots, D\} \setminus \mathcal{B}$. The complete set was taken as $\mathcal{C}_{[1]} = (\mathcal{B} - \mathcal{B}) \cup (\mathcal{B}^c_{[1]} - \mathcal{B}) \cup (\mathcal{B} - \mathcal{B}^c_{[1]})$. The latter set in the union is irrelevant as $U_{\boldsymbol{h}} = U_{-\boldsymbol{h}}$. In this paper, instead, $\mathcal{B}^c$ consists only of the KH zero found during the search of the base set, as the search is based on the use of property 1. So, $\mathcal{B}^c \subset \mathcal{B}^c_{[1]}$.

We first sketch the simpler proof given in [1] using $\mathcal{B}^c_{[1]}$. If $\mathcal{B} = \{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_N\}$ then $\{|A_{\boldsymbol{h}_1}\rangle, \ldots, |A_{\boldsymbol{h}_N}\rangle\}$ is a basis of $\mathfrak{H}$ and $(\mathcal{U}_\mathcal{B})$, the relevant non-singular KH-matrix, is its Gram matrix. Recall that any translated base set $\mathcal{B} + \boldsymbol{h}$ is still a base set as the matrix $(\mathcal{U}_\mathcal{B})$ is left invariant. So, for any $\boldsymbol{h} \in \mathbb{Z}^D$, we can write (cf also section 3.1)

$$|A_{\boldsymbol{h}}\rangle = \sum_{p,q=1}^{N} |A_{\boldsymbol{h}_p}\rangle \left(\mathcal{U}_\mathcal{B}^{-1}\right)_{pq} \langle A_{\boldsymbol{h}_q}|A_{\boldsymbol{h}}\rangle. \tag{25}$$

Left-multiplying by $\langle A_{\boldsymbol{0}}|$ and using equation (19), we obtain

$$U_{\boldsymbol{h}} = \sum_{p,q=1}^{N} U_{\boldsymbol{h}_p} \left(\mathcal{U}_\mathcal{B}^{-1}\right)_{pq} U_{\boldsymbol{h}-\boldsymbol{h}_q}. \tag{26}$$

Now choose $\boldsymbol{h} = \boldsymbol{k} + \boldsymbol{a}^*_\beta$, with $\boldsymbol{k} \in \mathcal{B}^c_{[1]}$ and $1 \leqslant \beta \leqslant D$. Rewrite equation (26) using the translated base set $\mathcal{B} + \boldsymbol{a}^*_\beta$, and the right-hand side of equation (26) contains only elements of $\mathcal{C}_{[1]}$. After exhausting all the $\boldsymbol{k} \in \mathcal{B}^c_{[1]}$ and $1 \leqslant \beta \leqslant D$, we recursively determine the $U_{\boldsymbol{h}}$ with $\boldsymbol{h} = \boldsymbol{k} + 2\boldsymbol{a}^*_\beta$, $\boldsymbol{h} = \boldsymbol{k} + 3\boldsymbol{a}^*_\beta$, ...

For the proof with the smaller sets $\mathcal{B}^c$ and $\mathcal{C}$, it is sufficient to show how one obtains $\mathcal{B}^c_{[1]}$ from $\mathcal{C}$. Substituting in equation (26) successively every $\boldsymbol{h} \in \mathcal{C}_{[1]} \setminus \mathcal{C}$, choosing only one index from every pair $(\boldsymbol{h}, -\boldsymbol{h})$, we obtain a linear system with as many equations as unknowns. In fact, the right-hand side contains only elements of $\mathcal{C}_{[1]}$. The system is clearly determined because equation (26) is identically satisfied only when $\boldsymbol{h} \in \mathcal{B}$. □

Note. The proof just given has its mathematical value. On practical grounds, however, it is more convenient to refer to $\mathcal{B}^c_{[1]}$ and $\mathcal{C}_{[1]}$. In fact, the difference between $\mathcal{B}^c$ and $\mathcal{B}^c_{[1]}$ (and consequently between $\mathcal{C}$ and $\mathcal{C}_{[1]}$) is typically a very small number of indices lying on the border of $\mathcal{B}$. The few indices added very rarely would be outside a high-resolution cut-off (cf figure 2) when $\mathcal{C}$ is inside.

We have now seen how to determine recursively all the $U_{\boldsymbol{h}}$s with $\boldsymbol{h} \notin \mathcal{C}$. In this way, as already mentioned, the density can be obtained using equation (4). However, as implied by

the noted duality, ρ can be obtained analytically from the knowledge of $\mathcal{C}$. In fact, as shown in section 2.5 of [1], the scattering centres are determined by the unimodular roots of a system of polynomial equations, one equation for each index of $\mathcal{B}^c$. In this respect, having found that $\mathcal{B}^c \subset \mathcal{B}^c_{[1]}$, the number of polynomial equations to be solved is smaller than that reported in [1].

We now show how to combine these two methods to convert the system of polynomial equations in two variables in a set of polynomial equations in a single variable. In this way, the task of determining the coordinates $\{\boldsymbol{r}_j\}_{j=1,\dots,N}$ is considerably simplified. An important role is played by the base sets of [1], also described in section 3.2. These sets are characterized by a growth direction along one of the coordinate axes $\boldsymbol{a}_1^*$ or $\boldsymbol{a}_2^*$.

Algorithm 1

First step. Let $\mathcal{B}$ denote a known base set and let $\mathcal{C}$ denote the relevant complete set, also known. Thus, we can evaluate any $U_{\boldsymbol{h}}$ with index $\boldsymbol{h} \notin \mathcal{C}$ by the recursive procedure just mentioned. In particular, we shall evaluate all the $U_{\boldsymbol{h}}$s required to single out the base set $\mathcal{B}^{a_1^*}[m_1 \dots m_p]$ by the procedure reported prior to remark 8 in section 3.2. Adopting the column notation (cf definition 10) $\mathcal{B}^{a_1^*}[m_1 \dots m_p]$ reads $\mathcal{B}^{a_1^*}[\mu_1 \dots \mu_q]$ with $p = \mu_q$ and $q = m_1$.

Second step. Since m_1 is equal to the number of the *different* projections $x_{\hat{\jmath}}$ along $\boldsymbol{a}_1^* = (1, 0)$ of all the $\boldsymbol{r}_j$s and the $\mu_{\hat{\jmath}}$s are the multiplicities of the $x_{\hat{\jmath}}$s for $\hat{\boldsymbol{\jmath}} = 1, \dots, m_1$, we can reorder the scattering centres as

$$\{\boldsymbol{r}_j\}_{j=1,\dots,N} = \left\{\boldsymbol{r}_{\hat{\jmath},r} = (x_{\hat{\jmath}}, y_{\hat{\jmath},r})\right\}_{[\hat{\jmath}=1,\dots,m_1;\ r=1,\dots,\mu_{\hat{\jmath}}]}.$$

The $x_{\hat{\jmath}}$s can be determined as follows. Firstly, we observe that the set of indices $(0, 0), \dots, (0, m_1)$, lying in the first row, is linearly dependent. Then we use equation (22) with $N = m_1$ to evaluate coefficients $\alpha_0, \alpha_1, \dots, \alpha_{m_1}$ which, in turn, determine the m_1th-degree polynomial equation (16). After solving the latter, the roots yields the m_1 distinct values $\xi_{\hat{\jmath}} \equiv \exp(2\pi \mathrm{i} x_{\hat{\jmath}})$, $\hat{\boldsymbol{\jmath}} = 1, \dots, m_1$ and the $x_{\hat{\jmath}}$s are determined.

Third step. We can write (see section 2.5 of [1]):

$$U_{\boldsymbol{h}} \equiv U_{(h,k)} = \sum_{\hat{\jmath}=1}^{m_1} \xi_{\hat{\jmath}}^h B_{\hat{\jmath},k} \tag{27}$$

after putting

$$B_{\hat{\jmath},k} = \sum_{r=1}^{\mu_{\hat{\jmath}}} n_{\hat{\jmath},r} \mathrm{e}^{2\pi \mathrm{i} y_{\hat{\jmath},r} k}. \tag{28}$$

By letting h range over $0, 1, \dots, (m_1 - 1)$ and by keeping k fixed, equations (27) form a system of linear equations in the $B_{\hat{\jmath},k}$s for $\hat{\boldsymbol{\jmath}} = 1, \dots, m_1$. In fact, all the required $U_{\boldsymbol{h}}$ values can be evaluated by the aforementioned recursive procedure, while the $\xi_{\hat{\jmath}}$s have been determined at the previous step. Besides, the $\xi_{\hat{\jmath}}$s being all different, the determinant is different from zero and the $B_{\hat{\jmath},k}$s are uniquely determined.

Fourth step. Now take equation (28) with $k \neq 0$ and multiply $B_{\hat{\jmath},k}$ by

$$\zeta_{\hat{\jmath}} = \left(B_{\hat{\jmath},0}\right)^{-1} = \left(\sum_{r=1}^{\mu_{\hat{\jmath}}} n_{\hat{\jmath},r}\right)^{-1} > 0.$$

Put $\nu'_{\hat{\jmath},r} \equiv n_{\hat{\jmath},r}\zeta_{\hat{\jmath}}$. Evidently, $\nu'_{\hat{\jmath},r} > 0$ and $\sum_{r=1}^{\mu_{\hat{\jmath}}} \nu'_{\hat{\jmath},r} = 1$. Equation (28) becomes

$$B_{\hat{\jmath},k}\zeta_{\hat{\jmath}} = \sum_{r=1}^{\mu_{\hat{\jmath}}} \nu'_{\hat{\jmath},r} \mathrm{e}^{2\pi \mathrm{i} y_{\hat{\jmath},r} k}. \tag{29}$$

For each $\hat{\jmath}$, the structure on the right-hand side is that of the Fourier coefficient with index k relevant to a one-dimensional atomic density with $\mu_{\hat{\jmath}}$ scattering centres. Thus, the $y_{\hat{\jmath},r}$s can be determined algebraically, as in step 1 above, by solving a polynomial equation of degree $\mu_{\hat{\jmath}}$. After determining the $y_{\hat{\jmath},r}$s for $r = 1, 2, \ldots, \mu_{\hat{\jmath}}$, equations (29) can be solved in the $\nu'_{\hat{\jmath},1}, \ldots, \nu'_{\hat{\jmath},\mu_{\hat{\jmath}}}$. In this way, we determine all the weight $\nu'_{\hat{\jmath},r}$ associated with $y_{\hat{\jmath},r}$.

In conclusion, the determination of the N scattering centres and weights is equivalent to solving $p+1$ single-variable polynomial equations of degrees $m_1, \mu_1, \ldots, \mu_p$ and subsequently p systems of linear equations.

7. Generalization and conclusions

The results worked out in sections 4–6 will now be generalized to the case $D \geqslant 3$.

Firstly, we need to generalize the definition of ladder set. To this end we need some formal definitions. Let $\mathcal{B}$ denote a finite and simply connected subset of $\mathbb{Z}^D$. Recall (see the beginning of section 1.1) the definition of *first neighbourhood* of an index $\boldsymbol{h}$.

Definition 14. *We call the* inner frontier $\overline{\mathcal{F}}_{\mathcal{B}}$ *of set* $\mathcal{B}$ *the set of points* $\boldsymbol{h} \in \mathcal{B}$ *such that* $\mathcal{N}_{\boldsymbol{h}}$ *contains both points of* $\mathcal{B}$ *and points outside* $\mathcal{B}$. *Similarly,* $\mathcal{F}_{\mathcal{B}}$, *the* outer frontier *of* $\mathcal{B}$, *is defined as the set of points* $\boldsymbol{h} \notin \mathcal{B}$ *with the same property. Then define a* corner point *as a point* $\hat{\boldsymbol{c}} \in \overline{\mathcal{F}}_{\mathcal{B}}$ *such that set* $(\mathcal{N}_{\hat{\boldsymbol{c}}} \cap \mathcal{B}) - \hat{\boldsymbol{c}}$ *contains exactly* D *unit vectors which span* $\mathbb{R}^D$. *In other words, a corner point is one of the convex 'corners' of* $\mathcal{B}$. *The* indentation points *of* $\mathcal{B}$ *are defined complementarily as the concave 'corners' of* $\mathcal{B}$. *Formally, a point of* $\mathcal{F}_{\mathcal{B}}$ *is called an* indentation point *and is denoted by* $\hat{\boldsymbol{\imath}}$ *if it has only* D *nearest neighbours* $\boldsymbol{h}_{\hat{\imath},\alpha}$, $\alpha = 1, \ldots, D$, *which belong to* $\mathcal{F}_{\mathcal{B}}$ *and such that* $\boldsymbol{h}_{\hat{\imath},\alpha} - \hat{\boldsymbol{\imath}}$, *for* $j = 1, \ldots, D$, *point along the coordinate axes of* $\mathbb{Z}^D$. *We generalize definition 9 by introducing the* positive hyperquadrant *of indices associated with* $\hat{\boldsymbol{\imath}}$ *and denoted by* $\mathcal{Q}_{\hat{\imath}}$, *as*

$$\mathcal{Q}_{\hat{\imath}} \equiv \left\{ \boldsymbol{h} \;\middle|\; \boldsymbol{h} - \hat{\boldsymbol{\imath}} = m_1\boldsymbol{a}_1^* + \cdots + m_D\boldsymbol{a}_D^*, m_\alpha = 0, 1, \ldots, \alpha = 1, \ldots, D \right\}.$$

Finally, the set $\mathcal{B}$ *is a* ladder set *if for each of its indentation points, denoted by* $\hat{\boldsymbol{\imath}}_j$ *with* $j = 1, \ldots, M$, *one has* $\mathcal{B} \cap \mathcal{Q}_{\hat{\imath}_j} = \emptyset$ *for* $j = 1, \ldots, M$.

We now report a procedure for enlarging a set $\mathcal{B}$ of linearly independent indices such that the ladder condition is ensured at each step and the enlarged set is as close as possible to a *hypercube*. Assume we have found a linearly independent *hypercubic* set $\mathcal{B}_{M-1}$ of edge $(M-1)$, i.e. $\mathcal{B}_{M-1} = \{\boldsymbol{h} | 0 \leqslant h_j \leqslant (M-1),\ j = 1, \ldots, D\}$. The assumption is legitimate because for $M = 0$ one has $\mathcal{B}_0 = \{\boldsymbol{0}\}$ and this set is clearly linearly independent. In order to construct the set $\mathcal{B}_M$ we need to add to $\mathcal{B}_{M-1}$ the square faces $(M, M, \ldots, j, M, \ldots, k, M, \ldots, M)$ with $1 \leqslant j < k \leqslant M$. Thus, we must include the required indices, one at a time, in such a way that the enlarged set is a ladder-shaped one at each step. After ordering the faces, this can be done by adopting for each face the inclusion procedure considered in the case $D = 2$. For instance,

in the case $D = 3$, one chooses the order $(1, 2) \Rightarrow 1$, $(2, 3) \Rightarrow 2$, $(3, 1) \Rightarrow 3$ and, assuming $M = 3$, one enlarges in the following order:

$$[(3, 0, 3), (0, 3, 3)],\ [(3, 3, 0), (3, 0, 3)],\ [(0, 3, 3), (3, 3, 0)],$$
$$[(3, 1, 3), (1, 3, 3)],\ [(3, 3, 1), (3, 1, 3)],\ [(1, 3, 3), (3, 3, 1)],$$
$$[(3, 2, 3), (2, 3, 3)],\ [(3, 3, 2), (3, 2, 3)],\ [(2, 3, 3), (3, 3, 2)],\ [(3, 3, 3)].$$

For greater clarity, the elements relevant to the same face are inside square brackets. Actually, elements already included must no longer be considered and the effective inclusion procedure is

$$(3, 0, 3), (0, 3, 3), (3, 3, 0), (3, 1, 3), (1, 3, 3), (3, 3, 1), (3, 2, 3), (2, 3, 3), (3, 3, 2), (3, 3, 3).$$

At each step, in order to include the index and enlarge the linearly independent set of indices, the resulting KH determinant must be different from zero. When this is equal to zero, the index is a KH zero and it contributes to $\mathcal{B}^c$. By construction, this element is found when a particular face is considered. This has a ladder border and the index, by construction, lies either at the right of the top line or at the right of the bottom line. We observe now that this index, denoted by $\hat{\boldsymbol{\imath}}$, is an indentation point for the set of linearly independent indices. In fact, it determines a quadrant external to the face on the plane containing the face and, considering the directions orthogonal to the plane, one obtains a hyperquadrant external to the set $\mathcal{B}$. Besides, we can now apply property 1 in order to express all the vectors relevant to the indices of the quadrant contained in the considered face and originating at $\hat{\boldsymbol{\imath}}$. Then, by moving along each direction orthogonal to the face by a step at the time externally to $\mathcal{B}$, we express all the relevant vectors as linear combinations of the vectors relevant to the indices of $\mathcal{Q}_{\mathbf{0}} \setminus \mathcal{Q}_{\hat{\imath}}$. At this point, the search of the base set proceeds as explained above by adding indices lying in $\mathcal{Q}_{\mathbf{0}} \setminus \mathcal{Q}_{\hat{\imath}}$. We can now apply the same considerations of section 5, since the proof of property 1 is easily generalized to the case $D > 2$, and in this way we may single out a base set, close to a *hypercube*. Similarly, we may generalize the *centred square* algorithm (definitions 12 and 13 are easily extended to $D > 2$). Also the more complex *spherical* algorithm reported in section 5 is easily extended to $D > 2$.

The generalization of the results reported in section 6 is now immediate, and we can conclude that the results of sections 4–6 apply to any D.

Comment. In concluding the paper, however, we would like to comment on the applicability of the former procedures. So far, it has been assumed that the Fourier coefficients (or their moduli) are known exactly. In practice, they are only known approximately owing to unavoidable experimental uncertainties. Thus, the reported algorithms, in order to be used practically, need to be supplemented with a technique which is able to correct the Fourier coefficients. For crystallographic applications to structure solution, we have already found a satisfactory procedure for the one-dimensional case [24] and we are working to extend it to the three-dimensional case by the reported results. However, the numerical difficulties increase quite sharply with N. This is due to the characteristic ill-conditionedness of Toeplitz matrices, even though the cure for this problem is registering a continuous progress since Toeplitz matrices and generalized Vandermonde determinants (see, e.g., [23, 25–28] and references therein) form an active field of research, because they have applications in as many fields as the Fourier series.

On theoretical grounds, however, the interest of the former results can hardly be denied. Returning to crystallography, crystal structures are routinely solved by other methods (such as direct methods, maximum-entropy and shake-and-bake, cf [11, 29, 30]). They are all based on constrained minimization of a statistical indicator in a chosen configuration space. However,

the search for more powerful and reliable techniques, where the structure of constraints plays a decisive role, is very active. This fact motivated our research for a deeper understanding of the problem. Hence the interest of an *analytical* method of solution goes beyond the present possibility of application.

Acknowledgments

We are grateful to Professor K F Fischer for useful correspondence. Financial support from MURST through 60% and 40% funds is acknowledged.

Appendix

Let $\boldsymbol{h}^{(1)}, \ldots, \boldsymbol{h}^{(p)}$ denote the first p KH zeros during the search procedure of a ladder base set. We now show that:

Property A1. *All the elements of* $\bigcup_{j=1}^{p} \mathcal{Q}_{\boldsymbol{h}^{(j)}}$ *are linear combinations of elements lying in* $\mathcal{Q}_{\boldsymbol{0}} \setminus (\bigcup_{j=1}^{p} \mathcal{Q}_{\boldsymbol{h}^{(j)}})$.

Before proving this point, we need two further properties. The first states that:

Property A2. *Suppose that between the* $\{\boldsymbol{r}_j\}_{j=1,\ldots,N}$ *we have exactly* N_1 distinct *projections* $x_{j,1} = \boldsymbol{r}_j \cdot \boldsymbol{a}_1^* = \boldsymbol{r}_j \cdot (1, 0)$. *Reorder* $\{\boldsymbol{r}_j\}_{j=1,\ldots,N} = \{(x_{\hat{\jmath}}, y_{\hat{\jmath},r})\}_{\hat{\jmath}=1,\ldots,N_1;\, r=1,\ldots,\mu_{\hat{\jmath}}}$, *where* $y_{\hat{\jmath},r}$, $r = 1, \ldots, \mu_{\hat{\jmath}}$, *are the distinct* $(0, 1)$ *projections of points sharing the same* $x_{\hat{\jmath}}$. *Then the line subset of vectors* $|A_{(h,k)}\rangle$ *(defined in equation (18)) with* k *fixed and* $h = 0, 1, \ldots,$ *contains exactly* N_1 *linearly independent vectors.*

Proof. The generic element of the associated KH matrix of order q reads

$$\begin{aligned}(\mathcal{U}_q)_{h_1+1,h+1} &= \langle A_{(h_1,k)}|A_{(h,k)}\rangle = U_{(h,k)-(h_1,k)} \\ &= \sum_{j=1}^{N} n_j \mathrm{e}^{2\pi \mathrm{i} x_{j,1}(h-h_1)} = \sum_{j=1}^{N} n_j \xi_j^{h-h_1}\end{aligned}$$

where $\xi_j \equiv \mathrm{e}^{2\pi \mathrm{i} x_{j,1}}$. It does not depend on k. The determinant of $(\mathcal{U}_q)$ is a generalized Vandermonde determinant, given by (see equation (B.12) of [1])

$$\det(\mathcal{U}_q) = \sum_{1 \leqslant j_1 < j_2 < \cdots < j_q \leqslant N} n_{j_1} \ldots n_{j_q} \prod_{1 \leqslant r < s \leqslant q} |\xi_{j_s} - \xi_{j_r}|^2. \tag{A1}$$

This formula shows that the determinant is equal to zero when $q > N_1$. Thus, for $h \geqslant N_1$, $|A_{(h,k)}\rangle$ is a linear combination of $\{|A_{(h,k)}\rangle\}_{h=0,\ldots,N_1-1}$. By the same reasoning, after denoting by N_2 the number of the *distinct* $x_{j,2}$s, one concludes that, for any fixed h, only the vectors $\{|A_{(h,k)}\rangle\}_{k=0,\ldots,N_2-1}$ are linearly independent, while $|A_{(h,k)}\rangle$, when $k \geqslant N_2$, is a linear combination of $\{|A_{(h,k)}\rangle\}_{k=0,\ldots,N_2-1}$ The generalization to the case $D > 2$ is trivial. □

Property A3. *The search of a base set can always be restricted to a finite hyperparallelepiped* $\mathcal{R}$ *of* $\mathbb{Z}^D$, *with* $\mathcal{R} \equiv \{(h_1, \ldots, h_D) |\, 0 \leqslant h_1 < N_1, \ldots, 0 \leqslant h_D < N_D\}$ *and* $N_1, \ldots, N_D \leqslant N$ *are the number of distinct projections along the* D *coordinate axes.*

Proof. This is an immediate consequence of property A2. □

Proof of property A1. The proof will proceed by induction. The property is true for $p = 1$. Thus, we need to prove that the property holds true for $(p + 1)$ if it is true for p. The search

procedure is supposed to ensure that the last KH zero is located either at the far right of the top row or at the top of the line at the far right. Refer, for definiteness, to the first case and name the last KH zero by $\boldsymbol{h}^{(p+1)}$. Put $\mathcal{Q}_{\mathrm{T}} \equiv \mathcal{Q}_{\boldsymbol{h}^{(p+1)}} \cap \mathcal{Q}_p$, where $\mathcal{Q}_p = \bigcup_{j=1}^{p} \mathcal{Q}_{\boldsymbol{h}^{(j)}}$ denotes (inductive hypothesis) the set of indices whose associated vectors are linear combinations of vectors associated with $\mathcal{Q}_{\mathbf{0}} \backslash \mathcal{Q}_p$. Also put $\mathcal{S}_1 \equiv \mathcal{Q}_p \backslash \mathcal{Q}_{\mathrm{T}}, \mathcal{S}_2 \equiv \mathcal{Q}_{\boldsymbol{h}^{(p+1)}} \backslash \mathcal{Q}_{\mathrm{T}}$ and $\boldsymbol{\Lambda} \equiv \mathcal{Q}_{\mathbf{0}} \backslash (\mathcal{Q}_p \cup \mathcal{Q}_{\boldsymbol{h}^{(p+1)}})$. Adopting a simpler notation, we shall denote by $\vec{t}_j$, $\vec{s}_{(1),j}$, $\vec{s}_{(2),j}$ and $\vec{\lambda}_j$ the vectors associated with the jth point of $\mathcal{Q}_{\mathrm{T}}$, $\mathcal{S}_1$, $\mathcal{S}_2$ and $\boldsymbol{\Lambda}$, respectively. The vectors of $\mathcal{Q}_p$ either belong to $\mathcal{Q}_{\mathrm{T}}$ or to $\mathcal{S}_1$. They are linear combinations of vectors belonging to $\boldsymbol{\Lambda}$ and to $\mathcal{S}_2$, property 1 being true at value p. Thus, one has

$$\vec{s}_{(1),j} = \alpha_{j,k}\vec{s}_{(2),k} + \beta_{j,k}\vec{\lambda}_k \tag{A2}$$

$$\vec{t}_j = \gamma_{j,k}\vec{s}_{(2),k} + \epsilon_{j,k}\vec{\lambda}_k \tag{A3}$$

where (α), (β), (γ) and (ϵ) are appropriate matrices and the sums over repeated indices are understood. For the vectors of $\mathcal{Q}_{\boldsymbol{h}^{(p+1)}}$ property 1 ensures the validity of a similar set of relations which read

$$\vec{s}_{(2),j} = \alpha'_{j,k}\vec{s}_{(1),k} + \beta'_{j,k}\vec{\lambda}_k \tag{A4}$$

$$\vec{t}_j = \gamma'_{j,k}\vec{s}_{(1),k} + \epsilon'_{j,k}\vec{\lambda}_k \tag{A5}$$

where (α'), (β'), (γ') and (ϵ') are the matrices appropriate to this case. Moreover, property A3 ensures that all the indices present in equations (A2)–(A5) are a finite set. Substituting equation (A4) in the right-hand side of equations (A2) and (A3) one finds, respectively,

$$(\delta_{i,j} - \alpha_{i,k}\alpha'_{k,j})\vec{s}_{(1),j} = (\alpha_{i,k'}\beta'_{k',j} + \beta_{i,j})\vec{\lambda}_j \tag{A6}$$

$$(\gamma_{i,k}\alpha'_{k,j} - \gamma'_{i,j})\vec{s}_{(1),j} = (\epsilon'_{i,j} - \gamma_{i,k'}\beta'_{k',j})\vec{\lambda}_j. \tag{A7}$$

Equation (A7) follows by comparing (A3) and (A5). Since the $\vec{s}_{(1),j}$s are fully known owing to equations (A2), the system of linear equations (A6) and (A7) must be solvable with respect to the $\vec{s}_{(1),j}$s, while the remaining equations represent linear constraints for the $\vec{\lambda}_j$s. In this way, one concludes that the $\vec{s}_{(1),j}$s can be expressed in terms of vectors belonging to the L-shaped region $\boldsymbol{\Lambda}$ and property A1 is proven. □

References

[1] Cervellino A and Ciccariello S 1996 *Riv. Nuovo Cimento* **19** 1
[2] Ott H 1927 *Z. Kristall.* **66** 136
[3] Gabardo J-P 1999 *J. Math. Anal. Applic.* **239** 349
[4] Gabardo J-P 1998 *Trans. Am. Math. Soc.* **350** 4473
[5] Mattila P 1995 *Geometry of Sets and Measures in Euclidean Spaces. Fractals and Rectifiability* (Cambridge: Cambridge University Press) p 9
[6] Iohvidov I S 1982 *Hankel and Toeplitz Matrices and Forms* (Boston, MA: Birkhäuser) p 96
[7] Karle J and Hauptman H A 1950 *Acta Crystallogr.* **3** 181
[8] Goedkoop J B 1950 *Acta Crystallogr.* **3** 374
[9] Senechal M 1995 *Quasicrystals and Geometry* (Cambridge: Cambridge University Press) p 27
[10] Giacovazzo C, Monaco H L, Viterbo D, Scordari F, Gilli G, Zanotti G and Catti M 1992 *Fundamentals of Crystallography* ed C Giacovazzo (Oxford: IUCr-Oxford University Press) ch 1, 2, 5
[11] Giacovazzo C 1980 *Direct Methods in Crystallography* (London: Academic)
[12] Avrami M 1938 *Phys. Rev.* **50** 300
[13] Fischer K F and Pilz K 1997 *Acta Crystallogr.* A **53** 475
[14] Silva A M and Navaza J 1981 *Acta Crystallogr.* A **37** 658
[15] Navaza A and Navaza J 1992 *Acta Crystallogr.* A **48** 695
[16] Navaza J and Silva A M 1979 *Acta Crystallogr.* A **35** 266

[17] Patterson A L 1939 *Nature* **143** 939
[18] Patterson A L 1939 *Phys. Rev.* **55** 682
[19] Patterson A L 1944 *Phys. Rev.* **65** 195
[20] Caccioppoli R 1955 *Lezioni di Analisi Matematica* vol 1 (Naples: Treves)
[21] Kitaigorodskii A I 1961 *The Theory of Crystal Structure Analysis* (Transl. by Harker D and Harker K) (New York: Consultant Bureau)
[22] Gantmacher F R 1966 *Théorie des Matrices* vol 1 (Paris: Dunod)
[23] Gohberg I and Olshevsky V 1994 *Numer. Math.* **67** 71
[24] Cervellino A and Ciccariello S 1999 *Z. Kristall.* **214** 739
[25] Gohberg I and Olshevsky V 1997 *J. Complexity* **13** 208
[26] Ng M K 1997 *Linear Alg. Applic.* **259** 307
[27] Huang Y and McColl W F 1999 *J. Parallel Distrib. Comp.* **56** 99
[28] Serra S 1999 *Math. Comp.* **68** 793
[29] Bricogne G 1984 *Acta Crystallogr.* A **40** 410
[30] Miller R, Gallo S M, Khalak H G and Weeks C M 1994 *J. Appl. Crystallogr.* **27** 613