Diffraction from visible lattice points and kth power free integers

Michael Baake[1], Robert V. Moody[2] and Peter A. B. Pleasants[3]

1) Institut für Theoretische Physik, Universität Tübingen,
Auf der Morgenstelle 14, D-72076 Tübingen, Germany
2) Department of Mathematical Sciences, University of Alberta,
Edmonton, Alberta T6G 2G1, Canada
3) Department of Mathematics and Computing Science,
University of the South Pacific, Suva, Fiji

Dedicated to Ludwig Danzer on the occasion of his 70th birthday

Abstract

We prove that the set of visible points of any lattice of dimension $n \geq 2$ has pure point diffraction spectrum, and we determine the diffraction spectrum explicitly. This settles previous speculation on the exact nature of the diffraction in this situation. Using similar methods we show the same result for the 1-dimensional set of kth-power-free integers with $k \geq 2$. Of special interest is the fact that neither of these sets is a Delone set – each has holes of unbounded inradius. The paper contains a careful formulation of the mathematical ideas underlying the study of diffraction from infinite point sets.

Introduction

It has long been known that the diffraction spectrum of a crystal consists of pure Bragg peaks only, being a pure point measure supported on the lattice dual to the lattice of periods. Until about 15 years ago it was tacitly believed that crystals were the only discrete point sets with this property. Now, however, we know that many quasicrystals have pure point diffraction spectra (though in a significantly different sense, since the locations of the peaks are no longer discrete). These quasicrystals are all Meyer sets, that is, sets S that are both uniformly discrete and relatively dense and whose difference sets $\Delta = S - S$ also have these properties. (For a discussion of Meyer sets, see [11]). There has consequently been a feeling that a perfectly diffracting discrete point set, if not precisely a Meyer set, must be a closely related to a Meyer set. Indeed, the Meyer condition cannot possibly be strictly necessary since, as we shall see later, adding or removing a set of density zero does not alter the diffraction spectrum of a discrete point set, and one can clearly destroy the relative denseness of any uniformly discrete point set by removing a set of density zero.

In this paper we give some simple examples of perfectly diffractive discrete point sets that deviate much further from the Meyer properties than this; in fact we consider sets which, for arbitrarily large D, have a lattice of holes of inner diameter at least D. Such a set cannot differ from a Meyer set (or from a model set, which is a particular case of a Meyer set) only by a set of density zero. The sets comprising our examples are well known in number theory: they are the sets of visible (or primitive) points of a lattice in any dimension $n \geq 2$, see [1, 9], and the 1-dimensional sets consisting of the kth-power-free integers for $k \geq 2$, see [9, §6.6].

As well as giving a rigorous derivation of the diffraction spectra that depends on explicitly calculating the autocorrelation of the point set, we precede it by a shorter derivation of the pure point parts of the spectra only, via the Fourier transform of the point set. This follows along lines similar to previous attempts and uses a general result of A. Hof. It provides a quick, elegant way of calculating the discrete parts of the spectra which gives the correct results but does not have a full mathematical justification at present.

The rigorous approach is a response to the history of the problem for the visible lattice, described in [2]. In particular, there is a clear disagreement between earlier results in [14] and in [10] regarding the nature of the diffraction. This was partially resolved in favour of [10] by a calculation of the point part of the diffraction spectrum in terms of Dirichlet series and its comparison with a real optical experiment [2]. The older numerical calculations in [14], using the fast Fourier transform, suffer from an insufficient resolution and are misleading. In this article, we give the definitive answer to the nature of the diffraction spectrum by making the previous formal calculation of the pure point part [10, 2] rigorous and by proving that there is no continuous part to the diffraction spectrum of visible lattice points in dimension $n \geq 2$.

The paper is organized as follows. We start by describing the results from number theory and related areas we shall need, then describe some basic properties of the set of visible points. Next we discuss the background material of Fourier transforms and autocorrelations needed for diffraction spectra. This is first framed in the language of tempered distributions and then connected to measure theoretical considerations which are a basic part of the theory of diffraction. We then give the short intuitively motivated method of computing the pure point part of the diffraction spectrum, both for the visible points and the kth-power-free integers. The following two sections are devoted to an explicit calculation of the autocorrelation of the set of visible points, followed by a rigorous derivation of the diffraction spectrum. The final section treats the set of kth-power-free numbers in the same way.

Strictly speaking, neither the measure theoretical picture nor the intuitive approach to computing the diffraction of the visible points and the kth-power-free integers are necessary for the logic of the paper, and they could be omitted by the reader interested only in the bare bones of the results.

Tools from number theory

Here we set out some results we need from number theory and related areas.

Notation

We use the notation $(l, m) = \gcd(l, m)$ to denote the greatest common divisor of two integers l and m. For integers d and x, $d \mid x$ means that d is a divisor of x. Summation conditions like "$d \mid x$" are to be interpreted as meaning that d runs through positive divisors of x only (even when x is negative). The divisor function $\tau(m)$ (defined for $m \in \mathbb{N}$) counts the number of positive divisors of m, so $\tau(m) = \sum_{d|m} 1$.

Also, we will frequently use the O-notation for error estimates. For example, we say that a function $f(r)$ is $O(1/r)$ if there is a constant c such that $|f(r)|$ is bounded by c/r.

Lattices

A *lattice* in $\mathbb{R}^n$ is a set Γ of the form

$$\Gamma \;=\; \mathbb{Z}\boldsymbol{b}_1 \oplus \cdots \oplus \mathbb{Z}\boldsymbol{b}_n\,, \tag{1}$$

where $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ is a set of n linearly independent vectors called a *basis* of Γ. We define $\mathrm{vol}(\Gamma)$ to be the volume of the fundamental region $\{t_1\boldsymbol{b}_1 + \cdots + t_n\boldsymbol{b}_n \mid 0 \le t_1, \ldots, t_n < 1\}$, which turns out to be independent of the basis chosen. It can be calculated as $\mathrm{vol}(\Gamma) = |\det(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)|$. Every lattice is uniformly discrete and relatively dense in $\mathbb{R}^n$ and is a subgroup of $\mathbb{R}^n$ under vector addition.

Proposition 1 *Let Γ be a lattice and $\boldsymbol{a}$ an arbitrary vector (not necessarily in Γ). Let $N(R)$ be the number of points $\boldsymbol{x}$ in $\Gamma + \boldsymbol{a}$ with $|\boldsymbol{x}| < R$. Then there are constants c_1 and c_2, depending only on Γ, such that, for all $R > 0$,*

$$|N(R)\mathrm{vol}(\Gamma) - v_n R^n| \;\leq\; c_1 R^{n-1} + c_2\,, \tag{2}$$

where v_n is the volume of the unit ball in $\mathbb{R}^n$.

Proof: The translates of the fundamental region of Γ by vectors in $\Gamma + \boldsymbol{a}$ tile $\mathbb{R}^n$. Let $\mathcal{V}$ be the total volume of those translates that meet the ball $\{|\boldsymbol{x}| < R\}$ and v the total volume of those translates that lie entirely inside this ball. Then the volume $v_n R^n$ of the ball and the number $N(R)\mathrm{vol}(\Gamma)$ are both bounded above by $\mathcal{V}$ and below by v. So their difference is at most $\mathcal{V} - v$, which is the total volume of the translates of the fundamental region that meet the boundary of the ball. If D is the diameter[1] of the fundamental region then this volume is at most $v_n(R + D)^n - v_n(R - D)^n \leq 2^{n+1} v_n D R^{n-1}$, when $R \geq D$, and at most $v_n(R + D)^n < 2^n v_n D^n$, when $R < D$. This gives the result with $c_1 = 2^{n+1} v_n D$ and $c_2 = 2^n v_n D^n$. □

If Γ is a lattice and r is a nonzero real number then $r\Gamma$ is also a lattice (a basis of $r\Gamma$ is $\{r\boldsymbol{b}_1, \ldots, r\boldsymbol{b}_n\}$). When r is an integer $r\Gamma \subseteq \Gamma$ and $r\Gamma$ is a sublattice (and a subgroup) of Γ, of index r^n. For a nonzero integer m and two points $\boldsymbol{a}$ and $\boldsymbol{b}$ in Γ we write $\boldsymbol{a} \equiv \boldsymbol{b} \pmod{m\Gamma}$ to mean that $\boldsymbol{a} - \boldsymbol{b} \in m\Gamma$. With this notation we have the following Chinese Remainder Theorem for a lattice Γ.

Proposition 2 *Let Γ be a lattice and $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r \in \Gamma$. If $m_1, m_2, \ldots, m_r \in \mathbb{N}$ are chosen such that $(m_j, m_k) = 1$, for $1 \leq j < k \leq r$, then there is a point $\boldsymbol{a} \in \Gamma$ such that the solutions of the simultaneous congruences*

$$\boldsymbol{x} \equiv \boldsymbol{a}_1 \pmod{m_1\Gamma}, \ldots, \boldsymbol{x} \equiv \boldsymbol{a}_r \pmod{m_r\Gamma} \tag{3}$$

are precisely the points $\boldsymbol{x} \in \Gamma$ with

$$\boldsymbol{x} \equiv \boldsymbol{a} \pmod{(m_1 \cdot m_2 \cdot \ldots \cdot m_r)\Gamma}. \tag{4}$$

Proof: This follows by applying the Chinese Remainder Theorem for integers [9, Thm. 2.7.1] to each coordinate with respect to a basis of Γ. □

[1] The diameter of a bounded set $S \subset \mathbb{R}^n$ is the supremum of all distances between points of S.

We define the *content* of a nonzero lattice point $\boldsymbol{x}$ in a lattice Γ by

$$\operatorname{cont}(\boldsymbol{x}) \ := \ \max\{l \mid \boldsymbol{x} \in l\Gamma\}\,. \tag{5}$$

If $\boldsymbol{x}$ is expressed in terms of a basis, $\boldsymbol{x} = \sum x_j \boldsymbol{b}_j$, then $\operatorname{cont}(\boldsymbol{x}) = \gcd(x_1, \ldots, x_n)$ (which is therefore independent of the particular basis chosen). For consistency and convenience we define $\operatorname{cont}(\boldsymbol{0}) = \infty$. For $m \in \mathbb{N}$ and $\boldsymbol{x} \in \Gamma$, we have

$$\operatorname{cont}(m\boldsymbol{x}) \ = \ m \cdot \operatorname{cont}(\boldsymbol{x}). \tag{6}$$

It is clear from (5) that for $\boldsymbol{x} \in \Gamma \setminus \{\boldsymbol{0}\}$

$$\operatorname{cont}(\boldsymbol{x}) \leq \frac{|\boldsymbol{x}|}{L(\Gamma)}, \tag{7}$$

where $|\boldsymbol{x}|$ is the Euclidean length of $\boldsymbol{x}$ and $L(\Gamma)$ is the length of the shortest nonzero vector in Γ.

The set $V = V(\Gamma)$ of *visible points* of a lattice Γ (also known as the *primitive* points of Γ) is

$$V \ := \ \{\boldsymbol{x} \in \Gamma \mid \operatorname{cont}(\boldsymbol{x}) = 1\}\,. \tag{8}$$

In terms of a lattice basis, V consists of all points whose coordinates have no common divisor. These are precisely the lattice points that are visible from the origin, in the sense that the line segment joining them to the origin contains no other lattice point. When $n = 1$, V consists of two points, equidistant from 0, but otherwise V is infinite and, indeed, contains more than half the lattice points, as we shall see. For $m \in \mathbb{N}$ we have

$$mV \ = \ \{\boldsymbol{x} \in \Gamma \mid \operatorname{cont}(\boldsymbol{x}) = m\}\,. \tag{9}$$

Density

Let S be a uniformly discrete set of points in $\mathbb{R}^n$ (i.e. there is a $c > 0$ so that $\boldsymbol{x}, \boldsymbol{y} \in S$ with $\boldsymbol{x} \neq \boldsymbol{y}$ implies $|\boldsymbol{x} - \boldsymbol{y}| \geq c$). We say S has *natural density* $\mathcal{D} = \operatorname{dens}(S)$ if

$$\mathcal{D}_R \ = \ \mathcal{D}_R(S) \ := \ \frac{|\{\boldsymbol{x} \in S \mid |\boldsymbol{x}| < R\}|}{v_n R^n} \quad \longrightarrow \quad \mathcal{D} \tag{10}$$

as $R \to \infty$. The expression $|\mathcal{D} - \mathcal{D}_R|$ (a function of R) is called the *error term* for the natural density. For example, by Proposition 1, $\mathcal{D} = \operatorname{dens}(\Gamma)$ exists and is equal to $1/\operatorname{vol}(\Gamma)$ (with error term $O(1/R)$) for every lattice Γ.

If $\operatorname{dens}(S)$ exists and if T is an orthogonal transformation then $\operatorname{dens}(T(S)) = \operatorname{dens}(S)$, but it is a failing of the natural density that it is not always true that $\operatorname{dens}(T(S)) = \operatorname{dens}(S)/|\det(T)|$ when T is a general affine transformation.

When we use the word "density" from now on we shall mean natural density.

Power-free numbers

For an integer exponent $k \geq 1$ the set $F = F_k$ of *kth-power-free integers* is

$$F \ := \ \{n \in \mathbb{Z} \mid n \text{ is not divisible by } d^k \text{ for any integer } d > 1\}\,. \tag{11}$$

An equivalent characterization of the numbers $n \in F$ is that in their prime power factorization $n = \pm p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ every exponent a_j is less than k. In close analogy with V, the case $k = 1$ is trivial, F_1 consisting of just the two numbers ± 1, but for $k \geq 2$ F_k is infinite and contains more than half the integers.

Inclusion-exclusion

The Möbius function $\mu(m)$ is defined for $m \in \mathbb{N}$ by

$$\mu(d) \ := \ \begin{cases} 1, & \text{when } m = 1, \\ (-1)^r, & \text{when } m \text{ is a product of } r \text{ distinct primes,} \\ 0, & \text{when } m \text{ is divisible by the square of a prime.} \end{cases} \tag{12}$$

It is multiplicative in the sense that $\mu(lm) = \mu(l)\mu(m)$ when $(l, m) = 1$ (and clearly $\mu(lm) = 0$ when $(l, m) \neq 1$).

A number of inversion formulæ and variants of the inclusion-exclusion principle can be expressed in terms of this function. The result we need in this paper is

$$\sum_{d|m} \mu(d) \ = \ \begin{cases} 1, & \text{if } m = 1, \\ 0, & \text{if } m > 1, \end{cases} \tag{13}$$

see [9, Thm. 6.3.1]. For a point $\boldsymbol{x}$ in a lattice Γ this shows that

$$\chi_V(\boldsymbol{x}) \ := \ \sum_{d|\mathrm{cont}(\boldsymbol{x})} \mu(d) \tag{14}$$

is the characteristic function of the visible points V of Γ (except that it is undefined when $\boldsymbol{x} = \mathbf{0}$).

Similarly, since $d^k \mid x$ if and only if d divides the largest integer whose kth-power divides x,

$$\chi_F(x) \ := \ \sum_{d^k|x} \mu(d) \tag{15}$$

is the characteristic function of the kth-power-free integers (except for being undefined when $x = 0$).

Dirichlet series

The well-known Riemann zeta-function (see, for example, [1, ch. 12] or [9, §9.2]) is defined for complex numbers s by

$$\zeta(s) \;:=\; \sum_{m=1}^{\infty} \frac{1}{m^s} \;=\; \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}, \tag{16}$$

where the sum is called a *Dirichlet series* and the product (in which p runs through all positive prime numbers) is called an *Euler product.* The sum and product are absolutely convergent for s in the half-plane $\mathrm{Re}(s) > 1$, but $\zeta(s)$ can be meromorphically continued to the whole of $\mathbb{C}$, for example $\zeta(0) = -1/2$. The only singularity of $\zeta(s)$ is a simple pole at $s = 1$ with residue 1. Using the Euler product, we see that the Dirichlet series of $1/\zeta(s)$ is given by

$$\frac{1}{\zeta(s)} \;=\; \prod_p \left(1 - \frac{1}{p^s}\right) \;=\; \sum_{m=1}^{\infty} \frac{\mu(m)}{m^s}. \tag{17}$$

This function has infinitely many poles (all in the half-plane $\mathrm{Re}(s) < 1$). Its value at 1 is 0 and its value at 0 is -2.

Another Dirichlet series we shall encounter is

$$\xi(s) := \sum_{m=1}^{\infty} \frac{\mu(m)\tau(m)}{m^s} = \prod_p \left(1 - \frac{2}{p^s}\right). \tag{18}$$

Again, both the Dirichlet series and Euler product are absolutely convergent in the half-plane $\mathrm{Re}(s) > 1$. It can be deduced from the Euler product that $\xi(1) = 0$.

Visible points of a lattice

Here we summarize some elementary and well-known properties of the set of visible points V of a lattice Γ, together with complete proofs.

Since Γ is a free Abelian group of rank n, its automorphism group, $\mathrm{Aut}(\Gamma)$, is isomorphic to the matrix group $GL(n,\mathbb{Z})$. Explicit isomorphisms can be found by taking coordinates with respect to any lattice basis.

Proposition 3 *The orbits of the action of $GL(n,\mathbb{Z})$ on Γ are the sets mV, $m \in \mathbb{N}_0$. In particular, $GL(n,\mathbb{Z})$ acts transitively on V.*

PROOF: Since the elements of $GL(n, \mathbb{Z})$ cannot decrease content and are invertible, they preserve content. Hence each set mV is invariant under $GL(n, \mathbb{Z})$. The transitivity of $GL(n, \mathbb{Z})$ on V can be seen from the facts that every visible point belongs to some basis of Γ [5, §3,Thm. 5] and that any two bases of Γ are related by a transformation in $GL(n, \mathbb{Z})$. For $m \in \mathbb{N}$ the transitivity of $GL(n, \mathbb{Z})$ on mV follows from its transitivity on V. Transitivity on the singleton orbit $0 \cdot V = \{\mathbf{0}\}$ is trivial. $\square$

Proposition 4 *V is uniformly discrete, but has holes of arbitrary size. Moreover, for any $r > 0$ there is a set of holes in V of inradius at least r whose centres have positive density.*

PROOF: The uniform discreteness is trivial, as V is a subset of a lattice. Now let $C = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s\}$ be any finite configuration of points in Γ. Choose s integer moduli $m_1, \ldots, m_s$ that are > 1 and coprime in pairs (for example, they could be the first s primes). By Proposition 2 there is a point $\boldsymbol{a} \in V$ with

$$\boldsymbol{a} \equiv -\boldsymbol{a}_1 \pmod{m_1\Gamma}, \ldots, \boldsymbol{a} \equiv -\boldsymbol{a}_s \pmod{m_s\Gamma}. \tag{19}$$

Now for any $\boldsymbol{x} \equiv \boldsymbol{a} \pmod{m_1 m_2 \ldots m_s \Gamma}$ the configuration $C + \boldsymbol{x} = \{\boldsymbol{a}_1 + \boldsymbol{x}, \ldots, \boldsymbol{a}_s + \boldsymbol{x}\}$ is congruent, in the geometric sense, to C but no point in $C + \boldsymbol{x}$ is visible, since $\boldsymbol{a}_j + \boldsymbol{x} \in m_j\Gamma$. The points $\boldsymbol{x}$ have density $\mathrm{dens}(\Gamma)/(m_1 m_2 \ldots m_s) > 0$. $\square$

For $n = 2$, and with the density of holes not mentioned, this is Thm. 5.29 of [1]. We note that the hole with s points nearest the origin provided by this argument can be expected to be at a distance of the order s^s and that the density guaranteed for holes of inradius r is of the order $r^{-n^2 r^n}$, so large holes, while having positive density, are probably extremely sparse.

This proposition shows that V, though uniformly discrete, is not relatively dense, and hence not a Delone set. Consequently it is not a Meyer set either. (Recall that Λ is a Meyer set if and only if both Λ and $\Lambda - \Lambda$ are Delone sets.) Also, V cannot be transformed into such a set by adding a set of zero density. However, we do have:

Proposition 5 *If $n \geq 2$ then $V - V = \Gamma$.*

PROOF: Clearly, $V - V \subseteq \Gamma$. Now let $\boldsymbol{x} = \sum_{j=1}^{n} x_j \boldsymbol{b}_j \in \Gamma$, where $n \geq 2$. Then, $(x_1, x_2, \ldots, x_n) = (x_1 + 1, 1, x_3, \ldots, x_n) - (1, 1 - x_2, 0, \ldots, 0)$, and $\boldsymbol{x}$ is the difference of two visible points. $\square$

The existence of arbitrarily large holes also implies that the set of visible points cannot have a uniform density (not, at least, when its density is positive in some sense, as is the case for $n \geq 2$). It does have a natural density, however:

Proposition 6 *The visible points V of a lattice $\Gamma \in \mathbb{R}^n$ have a natural density given by*

$$\mathrm{dens}(V) \ = \ \frac{\mathrm{dens}(\Gamma)}{\zeta(n)}, \tag{20}$$

with error term $O(1/R)$ when $n \neq 2$ and $O(\log R/R)$ when $n = 2$.

This is a standard example of the use of Möbius inversion given (at least for the case $\Gamma = \mathbb{Z}^2$) in most introductory number theory books (for example, [9, Thm. 6.6.3] and [1, Thm. 3.9]). In these particular references the averages are taken over triangles and squares, respectively, instead of balls. Indeed the density is essentially independent of the shape of the region averaged over. We say more about this later.

PROOF: The proposition is trivially true for $n = 1$, when the pole of $\zeta(s)$ at 1 gives a density of 0, so we assume from now on that $n \geq 2$.

The density of V is the limit as $R \to \infty$ of

$$\frac{1}{v_n R^n} \sum_{\substack{|\boldsymbol{x}|<R \\ \boldsymbol{x}\in V}} 1\,, \tag{21}$$

which by (14) is

$$\frac{1}{v_n R^n} \sum_{\substack{\boldsymbol{x}\in\Gamma\backslash\{\boldsymbol{0}\} \\ |\boldsymbol{x}|<R}} \ \sum_{m|\mathrm{cont}(\boldsymbol{x})} \mu(m) = \frac{1}{v_n R^n} \sum_{1\leq m<cR} \mu(m) \sum_{\substack{\boldsymbol{x}\in m\Gamma\backslash\{\boldsymbol{0}\} \\ |\boldsymbol{x}|<R}} 1\,, \tag{22}$$

where $c = 1/L(\Gamma)$ (using (7)). The inner sum is equal to the number of nonzero points $\boldsymbol{y} \in \Gamma$ with $|\boldsymbol{y}| < R/m$, which by Proposition 1 is

$$\frac{v_n}{\mathrm{vol}(\Gamma)} \left(\frac{R}{m}\right)^n + O\left(\left(\frac{R}{m}\right)^{n-1}\right) + O(1). \tag{23}$$

Substituting this in (22) gives

$$\mathrm{dens}(\Gamma) \sum_{1\leq m<cR} \frac{\mu(m)}{m^n} + O\Big(\frac{1}{R} \sum_{1\leq m<cR} \frac{1}{m^{n-1}}\Big) + O\left(\frac{1}{R^{n-1}}\right) \tag{24}$$

which tends to $\mathrm{dens}(\Gamma)/\zeta(n)$ as $R \to \infty$ when $n \geq 2$ by (17). The total error is $O(1/R^{n-1})$, from the last term and the tail of the sum in the main term, and $O(1/R)$ (or $O(\log R/R)$ when $n = 2$ and the series diverges logarithmically), from the middle term. □

Calculations of densities by Möbius inversion form the core of this paper. This is the first of many.

Diffraction spectra

In this section we assemble the facts we need about distributions, Fourier transforms and diffraction spectra. The mathematics underlying diffraction is quite subtle and needs to be spelled out carefully. Although the discussion in this section does not contain much that is new, it is nonetheless difficult to extract it all from any convenient source. We shall use a formulation based upon tempered distributions. For an essentially parallel approach which starts with measures, we refer to [7].

Autocorrelations

In dealing with diffraction, and therefore Fourier transforms, it is appropriate to use tempered distributions, whose test space $\mathcal{S}$ consists of the Schwartz functions (also known as "rapidly decreasing functions"). We refer to ([15, 13]) for this and the details on the standard topology which is used to describe convergence in $\mathcal{S}$. A *tempered distribution* T is a continuous (in the sense of this topology) linear functional on $\mathcal{S}$. The space of tempered distributions is denoted by $\mathcal{S}'$ and is equipped with the weak*-topology. A simple example is $\delta_{\boldsymbol{x}}$, *Dirac's delta-distribution*[2] at the point $\boldsymbol{x}$, defined by $(\delta_{\boldsymbol{x}}, \psi) := \psi(\boldsymbol{x})$ for all $\psi \in \mathcal{S}$.

If S is a uniformly discrete subset of $\mathbb{R}^n$, we call

$$\omega_S \;:=\; \sum_{\boldsymbol{x}\in S} \delta_{\boldsymbol{x}} \tag{25}$$

its *Dirac comb*. It is also a tempered distribution, and in fact the sum (25) is convergent (with any ordering) in the weak*-topology.

To describe the diffraction from a Dirac comb $\omega = \omega_S$ we need its *natural autocorrelation distribution* (also called its generalized Patterson function) defined by

$$\gamma_\omega \;:=\; \lim_{R\to\infty} \frac{1}{v_n R^n} \sum_{\boldsymbol{x},\boldsymbol{y}\in S_R} \delta_{\boldsymbol{x}-\boldsymbol{y}}\,, \tag{26}$$

where $S_R = S \cap B_R(\mathbf{0})$ and $B_R(\mathbf{0})$ is the ball of radius R centre $\mathbf{0}$. We shall simply call γ_ω the "autocorrelation" of S from now on. The existence of this limit in the weak*-topology is a prerequisite for the diffraction spectrum to be well defined. (The word "natural" refers to the use of the expanding ball $B_R(0)$ in the averaging process. Replacing it by an expanding region of some other shape might lead to a different limit.) It is clear from the definition that enlarging or diminishing S by a set of

[2]Nowadays it is usually called Dirac's point measure for reasons that will become clear shortly.

density 0 does not change its autocorrelation γ_ω. In particular, adding or removing any finite number of points does not change γ_ω.

We say that S has *finite local complexity* if the closure of $\Delta = S - S$ in $\mathbb{R}^n$ is discrete (as is clearly the case for all sets whose diffraction spectra we seek in this paper). Equivalently, we can say that Δ is a closed discrete set. Such an S itself is then uniformly discrete because $0 \in \Delta$ is isolated. The existence of the limit (26) in the weak*-topology is now controlled by the following result.

Lemma 1 *Let S be a set of finite local complexity and $\omega = \omega_S$ its Dirac comb. Then S has a natural autocorrelation γ_ω if and only if the coefficients $w(\boldsymbol{a})$*

$$w(\boldsymbol{a}) \;:=\; \lim_{R\to\infty} \frac{1}{v_n R^n} \sum_{\substack{|\boldsymbol{x}|,|\boldsymbol{x}-\boldsymbol{a}|<R \\ \boldsymbol{x},\boldsymbol{x}-\boldsymbol{a}\in S}} 1 \tag{27}$$

exist, i.e. the right hand side is convergent for all $\boldsymbol{a} \in \Delta$. In this case, $w(\boldsymbol{a}) \geq 0$ for all $\boldsymbol{a} \in \Delta$ and γ_ω is the tempered distribution of positive type[3] *given by the weak*-convergent sum*

$$\gamma_\omega \;=\; \sum_{\boldsymbol{a}\in\Delta} w(\boldsymbol{a})\delta_{\boldsymbol{a}}\,. \tag{28}$$

PROOF: The existence of γ_ω clearly implies the existence of $w(\boldsymbol{a})$ for all $\boldsymbol{a} \in \Delta$ because Δ is discrete and closed by assumption and the Schwartz space $\mathcal{S}$ contains all C^∞-functions of compact support, whence we can focus on any $\boldsymbol{a}$ we want.

Conversely, assume that the $w(\boldsymbol{a})$ exist (which are then ≥ 0 by definition). Since Δ is closed and discrete, its intersection with any compact subset of $\mathbb{R}^n$ contains at most finitely many points. Consequently, the $w(\boldsymbol{a})$ are locally summable and the right hand side of (28) defines a distribution over the space $\mathcal{D}$ of all C^∞-functions of compact support. We have to show that it is actually also a tempered distribution. This follows from the translation boundedness[4] of ω which is then inherited by γ_ω, see [7, Prop. 2.2]. Finally, γ_ω can also be written as a certain volume-normalized convolution (see below) which implies that it is a distribution of positive type. $\square$

Fourier transforms

The Fourier transform $\hat{T}$ of a tempered distribution $T \in \mathcal{S}'$ is defined by $(\hat{T},\psi) := (T,\hat{\psi})$, where we use the definition

$$\hat{\psi}(\boldsymbol{y}) \;:=\; \int_{\mathbb{R}^n} e^{-2\pi i \boldsymbol{y}\cdot\boldsymbol{x}}\,\psi(\boldsymbol{x})d\boldsymbol{x} \tag{29}$$

[3]See next section for a definition.

[4]We will explain this below in the context of measures.

for the Fourier transform of functions $\psi \in \mathcal{S}$. The Fourier transform maps the space $\mathcal{S}$ onto itself and is continuous on $\mathcal{S}$ in the standard topology for Schwartz functions [13, Thm. 7.7], hence it maps $\mathcal{S}'$ onto itself and is continuous on $\mathcal{S}'$ in the weak*-topology [13, Thm. 7.15].

As special cases, we mention $\hat{\delta}_{\mathbf{0}} = 1$ and the well-known *Poisson summation formula* for lattice Dirac combs

$$\hat{\omega}_{\Gamma} \; = \; \mathrm{dens}(\Gamma) \cdot \omega_{\Gamma^*} \tag{30}$$

where Γ^* is the dual or reciprocal lattice defined by

$$\Gamma^* \; := \; \{\boldsymbol{y} \mid \boldsymbol{y} \cdot \boldsymbol{x} \in \mathbb{Z}\,, \; \text{ for all } \boldsymbol{x} \in \Gamma\}\,. \tag{31}$$

(Eq. (30) can easily be derived from Poisson's summation formula for Schwartz functions [15].)

For a set S of finite local complexity, its *diffraction pattern* (also called its *diffraction distribution* or *diffraction spectrum*) is the Fourier transform $\hat{\gamma}_\omega$ of its autocorrelation γ_ω (when the latter exists). In view of the remarks above, $\hat{\gamma}_\omega$ is a tempered distribution. The autocorrelation of a lattice Γ, for example, is supported on Γ itself (since $\Delta = \Gamma - \Gamma = \Gamma$) and each peak has equal amplitude $\mathrm{dens}(\Gamma)$. So the autocorrelation is $\gamma_\omega = \mathrm{dens}(\Gamma)\omega_\Gamma$ and the corresponding diffraction spectrum is $\hat{\gamma}_\omega = \mathrm{dens}(\Gamma)^2\omega_{\Gamma^*}$, a constant multiple of the Fourier transform of ω_Γ itself. However, as we will see, the Fourier transform of a general point set (even of finite local complexity) does not describe its diffraction in such a simple way.

Distributions and Measures

The diffraction pattern $\hat{\gamma}_\omega$ is a tempered distribution. However, it is also a positive measure. This remarkable fact is indispensible for the general theory of diffraction, making available to it a vast array of concepts and tools. For example, it makes it immediately evident that the diffraction pattern may be viewed as having a pure point part and a continuous part, a fact of considerable physical significance.

The next two sections sketch out this distribution theory – measure theory connection in the context of diffraction. Although this background material is not explicitly needed to understand the proofs, it is essential to see the natural connection to measure theory. Also, it is this picture that has a generalization to the diffraction theory of translation bounded measures on locally compact Abelian groups which is an appropriate setting for more general questions.

A distribution T is *positive* if for all positive test functions ψ, $(T, \psi) \geq 0$. A distribution T is of *positive type* or is *positive definite* if for all test functions ψ, $(T, \psi * \psi^*) \geq 0$, where $\psi^*(x) := \overline{\psi(-x)}$. These two concepts are related by the Bochner-Schwarz theorem (see [15, Thm. VII.XVIII] or [12, Thm. IX.10]) which asserts that a tempered distribution is positive if and only if it is the Fourier transform of a tempered distribution of positive type.

Let $\mathcal{C}$ denote the space of complex-valued continuous functions of compact support on $\mathbb{R}^n$ and let $\|\cdot\|$ denote the supremum norm on $\mathcal{C}$. A (complex) *measure* ν on $\mathbb{R}^n$ is defined as a linear functional on $\mathcal{C}$ such that for every compact set $K \subset \mathbb{R}^n$ there is a constant a_K such that

$$|\nu(\phi)| \;\leq\; a_K \|\phi\| \tag{32}$$

for all $\phi \in \mathcal{C}$ with support in K. Such measures are in one-to-one correspondence with regular Borel measures due to the Riesz-Markov representation theorem [3, Ch.8, Sec. 69], and we thus identify these two pictures. We deal only with regular Borel measures in this paper. If for each $\phi \in \mathcal{C}$ inequality (32) holds uniformly for all translates of ϕ, we say that ν is *translation bounded*. This turns out to be a very useful concept because the Fourier transform of a tempered measure is a tempered distribution, but it need not be a measure again. Let us explain how to control this subtle point in the context of diffraction theory.

Distributions and measures act on different spaces of functions, and they are equipped with different topologies. Nonetheless, there is an important connection between them which comes through the fact that the space $\mathcal{D}$ of infinitely differentiable functions of compact support is dense both in $\mathcal{C}$ and in $\mathcal{S}$.

If a measure ν defines a tempered distribution T_ν by

$$(T_\nu, \psi) \;=\; \nu(\psi) \;=\; \int \psi \, d\nu \tag{33}$$

for all $\psi \in \mathcal{S}$, the measure ν is called *tempered.* A sufficient condition for a measure to be tempered is that it is slowly increasing in the sense that $\int (1 + |\boldsymbol{x}|^{-k}) |\nu|(d\boldsymbol{x}) < \infty$ for some $k \in \mathbb{N}$, see [15, Thm. VII.VII] or [13, Ex. 7.12 b]. Here, $|\nu|$ is the unique absolute value of ν, i.e. the smallest positive measure ρ such that $|\nu(\phi)| \leq \rho(|\phi|)$ for all $\phi \in \mathcal{C}$. It is also called the variational measure of ν. As a partial converse, any positive tempered distribution is a positive tempered measure. Thus, under the assumption of positiveness, tempered measures and tempered distributions can be viewed as the same thing, and the Bochner-Schwartz theorem can be restated as follows: ν is a positive tempered measure if and only if it is the Fourier transform of a tempered distribution of positive type.

Now, let ν be a translation bounded measure. Clearly ν is tempered. Our previous definition of the autocorrelation of a Dirac comb has natural extension to the autocorrelation γ_ν of ν by means of a volume-normalized convolution of ν with ν^*, where ν^* is defined by $\nu^*(\phi) := \overline{\nu(\phi^*)}$, see [7] for details. This construction of γ_ν, if it exists, guarantees that it is a positive definite tempered measure, so combining the previous arguments we see that the corresponding diffraction pattern $\hat{\gamma}_\nu$ is a tempered positive measure.

This is important because it is this object that describes what one actually sees on the screen in a diffraction experiment. In this article, we will only meet the simple case that ν is a Dirac comb ω_S. However, already the convolution with a function of compact support (a "profile" of the scatterer) shows why the more general setting is useful. Let us summarize this by the following result, which is a combination of [7, Prop. 2.2] and [7, Prop. 3.3].

Proposition 7 *Let ν be a translation bounded measure. If its natural autocorrelation, γ_ν, exists, it is a translation bounded (hence tempered) positive definite measure. Furthermore, $\hat{\gamma}_\nu$ is a positive measure and also translation bounded.*

Decomposition of measures

A *pure point* of a measure ν is a point $\boldsymbol{x} \in \mathbb{R}^n$ with $\nu(\{\boldsymbol{x}\}) \neq 0$. Since ν is a regular Borel measure, it has at most countably many pure points. Consequently, the sum of $|\nu(\{\boldsymbol{x}\})|$ over the pure points in any compact set K is convergent, and the pure points alone serve to define a measure ν_{pp} called the *pure point part* of ν. Thus ν has a unique decomposition as

$$\nu \; = \; \nu_{pp} + \nu_c \, , \tag{34}$$

where $\nu_c := \nu - \nu_{pp}$ is the so-called *continuous part*[5] of ν, and is characterized by having no pure points. A *pure point measure* is a measure whose continuous part is 0.

When a tempered measure ν is decomposed in this way the components ν_{pp} and ν_c are, of course, measures, but not necessarily tempered. For a translation bounded measure ν, however, the components are both translation bounded and hence tempered. If ν is a *positive* tempered measure, the decomposition is automatically into tempered components.

[5] The word "continuous" here does not refer to being continuous as a function, since a line measure in the plane, for example, is continuous. It refers to the intermediate value property that if there are sets $A \subset C$ with $\nu_c(A) < b < \nu_c(C)$ then there is a set B with $A \subset B \subset C$ and $\nu_c(B) = b$. Some authors use the word diffuse instead of continuous and the words atomic or purely discrete instead of pure point.

Returning now to diffraction patterns, we meet the special situation that $\hat{\gamma}_\omega$ is a *positive* measure. Consequently, it decomposes into a pure point part and a continuous part, both of which are positive measures. The pure point part is called the *Bragg spectrum* (of the point set S that created it). We say that the diffraction pattern is *pure point* or that S has a *pure point diffraction spectrum* if its diffraction pattern is a pure point measure, i.e. if it is equal to its Bragg spectrum.

Let us finally mention that, relative to Lebesgue measure, the continuous part ν_c can be further decomposed into an *absolutely continuous* and a *singular continuous* part, $\nu_c = \nu_{ac} + \nu_{sc}$, see [3, 12] for details. In our examples, we show that ν_c vanishes, so it means that there is neither an absolutely continuous nor a singular continuous component present.

Pure point distributions

Our principal concern is with showing that the visible points and kth-power-free points have a pure point spectrum. In this section we consider a special class of point measures which we will use in the sequel.

Consider an arbitrary point measure. It can be expressed in the form

$$\nu \;=\; \sum_{\boldsymbol{x}\in S} w(\boldsymbol{x})\delta_{\boldsymbol{x}}, \tag{35}$$

where the point set S is countable, but not necessarily uniformly discrete, and the coefficients or *weights* $w(\boldsymbol{x}) \in \mathbb{C}$ are not necessarily equal to one another. Note that the weights may be complex numbers. Let us assume in addition that the measure is translation bounded. This can be expressed as the condition that for every compact set $K \subset \mathbb{R}^n$ the sum

$$\sum_{\boldsymbol{x}\in S\cap(K+\boldsymbol{a})} |w(\boldsymbol{x})| \tag{36}$$

is convergent and bounded uniformly in $\boldsymbol{a}$. We denote the space of all translation bounded point measures by $\mathcal{T}$. All these measures are tempered, and we identify ν and T_ν in each case.

In the weak*-topology of $\mathcal{S}'$, $\mathcal{T}$ is not a closed subset of the tempered distributions $\mathcal{S}'$. For example, the sequence of pure point distributions

$$\{j^{-n}\omega_{\mathbb{Z}^n/j}\} \tag{37}$$

tends to the constant function 1 as $j \to \infty$. (For a test function ψ the numbers $j^{-n}\omega_{\mathbb{Z}^n/j}(\psi)$ are approximating sums to the integral[6] of ψ.) A similar argument shows

[6] Seen as a sequence of measures, (37) weak*-converges to Lebesgue measure.

that every bounded continuous function is a limit of pure point distributions. Unfortunately, as this example shows, the limit of such a sequence cannot, in general, be deduced by looking at the pointwise limit of the sequence.

However, taking pointwise limits in $\mathcal{T}$ is jusifiable under certain circumstances, and fortunately these apply in the cases of interest to us here. We introduce a "locally defined" norm on $\mathcal{T}$ by

$$\|\nu\|_{\mathrm{loc}} := \sup_K \int_K d|\nu| \;=\; \sup_K \sum_{\boldsymbol{x}\in S\cap K} |w(\boldsymbol{x})|\,, \tag{38}$$

where the supremum is taken over all compact sets K of diameter < 1. This norm defines a topology on $\mathcal{T}$ stronger than the weak*-topology and it provides a simultaneous "M-test" for pointwise and weak*-convergence of infinite sums of translation bounded point measures:

Lemma 2 *If $\nu_j \in \mathcal{T}$, $j \in \mathbb{N}$, and $\sum_{j=1}^{\infty} \|\nu_j\|_{\mathrm{loc}}$ is convergent then $\sum_{j=1}^{\infty} \nu_j$ is pointwise and weak*-convergent to the same sum $\nu \in \mathcal{T}$.*

PROOF: Choose $\psi \in \mathcal{S}$ and let $\varepsilon > 0$. On any compact set K of diameter less than 1 we have $|(\nu_j,\psi)| \le \|\nu_j\|_{\mathrm{loc}}\|\psi_K\|$, where ψ_K is the restriction of ψ to K. Covering $\mathbb{R}^n$ with unit balls and using the fact that ψ is rapidly decreasing, we see that we can find a positive constant C depending only on ψ such that

$$|(\nu_j,\psi)| \le C\|\nu_j\|_{\mathrm{loc}} + \frac{\varepsilon}{2^j}\,. \tag{39}$$

Since $\sum \|\nu_j\|_{\mathrm{loc}}$ is convergent, for all $N, M \in \mathbb{N}$ with $N \ge n_0(\varepsilon)$

$$|\sum_{j=N}^{N+M} (\nu_j,\psi)| \le C \sum_{j=N}^{N+M} \|\nu_j\|_{\mathrm{loc}}\|\psi\| + \varepsilon < 2\varepsilon\,. \tag{40}$$

Thus $\{(\nu_j,\psi)\}_j$ is convergent and we can define a linear functional ν on $\mathcal{S}$ by

$$(\nu,\psi) \;=\; \sum_{j=1}^{\infty}\sum_{\boldsymbol{x}\in S_j} w_j(\boldsymbol{x})\psi(\boldsymbol{x}) \;=\; \sum_{\boldsymbol{x}\in S} w(\boldsymbol{x})\psi(\boldsymbol{x}) \tag{41}$$

where $S := \bigcup S_j$, and $w(\boldsymbol{x}) = \sum_{j=1}^{\infty} w_j(\boldsymbol{x})$.

For any compact set K and any $\boldsymbol{a} \in \mathbb{R}^n$ we have

$$\sum_{\boldsymbol{x}\in S\cap(K+\boldsymbol{a})} |w(\boldsymbol{x})| \;=\; \sum_{j=1}^{\infty} \sum_{\boldsymbol{x}\in S_j\cap(K+\boldsymbol{a})} |w_j(\boldsymbol{x})| \;\le\; \sum_{j=1}^{\infty} C_K\,\|\nu_j\|_{\mathrm{loc}} \tag{42}$$

for some constant C_K depending only on K. From this it is easy to see that ν is a tempered measure (i.e. ν is continuous on $\mathcal{S}$) and finally that $\nu \in \mathcal{T}$. $\square$

$$\begin{tikzcd}
\omega \arrow[r, "*"] \arrow[d, "\mathrm{FT}"'] & \gamma_\omega \arrow[d, "\mathrm{FT}"] \\
\hat{\omega} \arrow[r, "|\cdot|^2"] & \hat{\gamma}_\omega
\end{tikzcd}$$

Figure 1: Wiener diagram

An intuitive derivation of the point spectra

This section describes a short, intuitive way of calculating the pure point part of the diffraction spectrum of the visible points V of a lattice Γ and also of the kth-power-free numbers $F = F_k$. Its purpose is to give a taste of our later number theoretic methods in a simpler setting and also to contrast this intuitively clear calculation with the more circuitous route we take later in rigorously establishing the complete diffraction spectra. It seems almost miraculous when the longer method eventually reduces to the same simple result. The intuitive method depends on the fact that since the autocorrelation γ_ω is a volume-normalized convolution of $\omega = \omega_V$ with itself, its Fourier transform, the diffraction pattern $\hat{\gamma}_\omega$, should be a normalized square of $\hat{\omega}$.

In this context, the coefficients of $\hat{\omega}$ are usually called *amplitudes*, even if they only exist formally, while those of $\hat{\gamma}_\omega$ are called *intensities* which relates to the fact that they are real and non-negative. The appropriate operation now is to determine the intensities as the absolute squares of the corresponding amplitudes of peaks, as indicated in the commutative Wiener diagram in Figure 1 where the vertical arrows represent the Fourier transform, the upper horizontal arrow the volume-normalized convolution, and the lower horizontal arrow taking the absolute squares of amplitudes.

This observation dates back at least 100 years in optics and is the standard procedure in diffraction theory [4]. In the context of diffraction from infinite arrangements, it has been made rigorous (at least for the pure point part of the spectrum) by Hof [7, 8]. The intuitive approach to the visible points runs through the Wiener diagram along the “low road”, via the difficult to interpret $\hat{\omega}$.

Visible points

We start from

$$\omega_V = \sum_{m=1}^{\infty} \mu(m) \omega_{m\Gamma \setminus \{\mathbf{0}\}}, \tag{43}$$

which follows from (14) since the sum on the right is supported on $\Gamma \setminus \{\mathbf{0}\}$ and the amplitude of the peak at $\boldsymbol{x} \in \Gamma \setminus \{\mathbf{0}\}$ is

$$\sum_{m|\mathrm{cont}(\boldsymbol{x})} \mu(m) \;=\; \chi_V(\boldsymbol{x}). \tag{44}$$

Although the sum on the left of (44) is finite for each $\boldsymbol{x}$, these pointwise sums are not uniformly convergent, so the sum on the right of (43) does not pass the M-test of Proposition 2 as a sum of pure point distributions. Nevertheless, it is clear that the sum does converge to ω_V in the weak*-topology (though not in the norm topology). Taking the Fourier transform of (43) term-by-term we obtain

$$\hat{\omega}_V = \sum_{m=1}^{\infty} \mu(m) \left(\frac{1}{m^n} \omega_{\Gamma^*/m} - 1 \right). \tag{45}$$

Since ω_V is a tempered distribution it has a Fourier transform $\hat{\omega}_V$ which is also a tempered distribution, and since the Fourier transform is a continuous operator on the space of tempered distributions with the weak*-topology [12] the sum on the right of (45) converges to $\hat{\omega}_V$ in the weak*-topology (though again it does not pass the M-test). This time the sums of the amplitudes at individual points of $\mathbb{Q}\Gamma^*$ are infinite but uniformly convergent, when $n \geq 2$.

It is difficult to interpret (45) as a distribution and we suspect that it is not a measure. Nevertheless, we can identify the pure points and their amplitudes formally (even though they may not be real peaks).

Define the *denominator* of a non-zero point $\boldsymbol{x} \in \mathbb{Q}\Gamma^*$ by

$$\mathrm{den}(\boldsymbol{x}) \;:=\; \gcd\{m \in \mathbb{Z} \mid m\boldsymbol{x} \in \Gamma^*\} \tag{46}$$

(it is the smallest positive integer such that $m\boldsymbol{x} \in \Gamma^*$). Then the sum of the amplitudes in (45) at a point $\boldsymbol{x} \in \mathbb{Q}\Gamma^* \setminus \{0\}$ with denominator d is

$$\begin{aligned}
\sum_{l=1}^{\infty} \frac{\mu(ld)}{(ld)^n} \;=\; \frac{\mu(d)}{d^n} \sum_{\substack{l=1 \\ (l,d)=1}}^{\infty} \frac{\mu(l)}{l^n} \;&=\; \frac{\mu(d)}{d^n} \prod_{p \nmid d} \left(1 - \frac{1}{p^n}\right) \\
&=\; \frac{\mu(d)}{\zeta(n) d^n} \prod_{p|d} \left(1 - \frac{1}{p^n}\right)^{-1} \\
&=\; \frac{\mu(d)}{\zeta(n)} \prod_{p|d} \frac{1}{p^n - 1}.
\end{aligned} \tag{47}$$

When $n \geq 2$ the pointwise sum of the pure point parts is thus nonzero at all points of $\mathbb{Q}\Gamma^*$ with squarefree denominator but is not absolutely locally summable. It is, however, locally square summable. The squaring operation gives, for the diffraction spectrum $\hat{\gamma}_\omega$ of V, the distribution with a pure point at each point of $\mathbb{Q}\Gamma^*$ with squarefree denominator, the peak at such a point with denominator d having intensity

$$\frac{1}{\zeta^2(n)} \prod_{p|d} \frac{1}{(p^n - 1)^2} \,. \tag{48}$$

These coefficients are uniformly absolutely locally summable and in fact are the correct intensities of the diffraction pattern. They were first derived, based on similar arguments, in [2]. However, one may well ask why this works. How can we justify the squaring operation as the appropriate mechanism to obtain the amplitudes? The intuition is supported as follows.

Hof's results

Two results of Hof [7, 8] are

Proposition 8 *Let ν be a translation bounded measure on $\mathbb{R}^n$ such that $\hat{\nu}$ is also a translation bounded measure. Then*

$$\hat{\nu}(\{\boldsymbol{x}\}) \;=\; \lim_{R\to\infty} \frac{1}{v_n R^n} \int_{B_R(\boldsymbol{0})+\boldsymbol{a}} e^{-2\pi i \boldsymbol{x}\cdot\boldsymbol{y}} \, \nu(d\boldsymbol{y}) \tag{49}$$

for every $\boldsymbol{x} \in \mathbb{R}^n$ and the limit exists uniformly in $\boldsymbol{a}$.

Proposition 9 *Let ν be a translation bounded measure with unique autocorrelation γ and suppose that for all $\boldsymbol{x} \in \mathbb{R}^n$,*

$$m_{\boldsymbol{x}} \;:=\; \lim_{R\to\infty} \frac{1}{v_n R^n} \int_{B_R(\boldsymbol{0})+\boldsymbol{a}} e^{-2\pi i \boldsymbol{x}\cdot\boldsymbol{y}} \, \nu(d\boldsymbol{y}) \,, \tag{50}$$

exists uniformly in $\boldsymbol{a}$. Then, for all $\boldsymbol{x}$, we have

$$\hat{\gamma}(\{\boldsymbol{x}\}) \;=\; |m_{\boldsymbol{x}}|^2 \,. \tag{51}$$

Taken together, these show that if the Fourier transform $\hat{\nu}$ of a translation bounded tempered distribution ν is also translation bounded (and hence can be decomposed as a pure point part and a continuous part) then the pure points of the diffraction spectrum $\hat{\gamma}_\nu$ are the same as the pure points of the Fourier transform $\hat{\nu}$, but their intensities are the absolute squares of the amplitudes of $\hat{\nu}$. Proposition 9 alone says that this continues to be true even when $\hat{\nu}$ is not translation bounded, at least as long as the formal expressions for the pure point amplitudes of $\hat{\nu}$ (which may now not represent peaks of $\hat{\nu}$ in the accepted sense) converge uniformly with respect to translation of physical space. We now give the analogous intuitive derivation for the kth-power-free integers before we come back to the status of the method.

kth-power-free integers

The parallel intuitive calculation for the pure point part of the diffraction spectrum of the set $F = F_k$ of kth-power-free integers in $\mathbb{R}$ goes like this. We have

$$\omega_F \;=\; \sum_{m=1}^{\infty} \mu(m)\omega_{m^k\mathbb{Z}\setminus\{0\}}, \tag{52}$$

hence

$$\hat{\omega}_F \;=\; \sum_{m=1}^{\infty} \mu(m)\left(\frac{1}{m^k}\omega_{\mathbb{Z}/m^k} - 1\right). \tag{53}$$

Let d_k ($\geq d$) be the smallest positive integer such that d_k^k is divisible by d. Then the formal sum of the amplitudes of the peaks of (53) at a point $x \in \mathbb{Q}$ with denominator d is

$$\begin{aligned}
\sum_{l=1}^{\infty} \frac{\mu(ld_k)}{(ld_k)^k} = \frac{\mu(d_k)}{d_k^k} \sum_{\substack{l=1\\(l,d)=1}}^{\infty} \frac{\mu(l)}{l^k} \;&=\; \frac{\mu(d_k)}{\zeta(k)d_k^k} \prod_{p|d}\left(1 - \frac{1}{p^k}\right)^{-1} \\
&=\; \frac{\mu(d^*)}{\zeta(k)} \prod_{p|d} \frac{1}{p^k - 1},
\end{aligned} \tag{54}$$

since, when d_k is squarefree, we have $d_k = d^* = \prod_{p|d} p$ (the squarefree kernel of d) and $d_k^k = \prod_{p|d} p^k$. Consequently, the pure point part of the diffraction spectrum of F has peaks at the points of $\mathbb{Q}$ with $(k+1)$th-power-free denominator, and its intensity at a such a point, with denominator d, is

$$\frac{1}{\zeta^2(k)} \prod_{p|d} \frac{1}{(p^k - 1)^2}. \tag{55}$$

This is a pure point distribution and also agrees with the result of Theorem 5.

Status of the method

In what follows we completely avoid Hof's formula (50) (which in any case could only determine the pure point part of the spectrum without any assurance that there is no continuous part). What prevents us from making our derivation of the pure point part of the spectrum rigorous by citing Proposition 9 is that for our examples the limit in (50) is not uniform in $\boldsymbol{a}$. This is most easily seen by noting that our point sets have arbitrarily large holes, so that no matter how large R is there will be an $\boldsymbol{a}$ such that $B_R(\mathbf{0}) + \boldsymbol{a}$ lies entirely within a hole. However, large holes are very sparse, so the limit is, in a sense, "nearly uniform". For irrational $\boldsymbol{x}$ the limit (50) can be shown to be 0 (as it should be) but again is almost certainly not uniform in $\boldsymbol{a}$.

The fact that our examples are outside the domain where Hof's result applies underscores the fact that we are in new territory not only for calculating the continuous part of the diffraction spectrum but even for calculating the diffraction peaks.

Autocorrelation of the visible points

The remainder of this paper consists of a rigorous derivation of the results of the previous section via the "high road" of the autocorrelation. So we begin by using an elaboration of the kind of Möbius inversion argument used in proving Proposition 6 to calculate the autocorrelation of the visible points V of a lattice Γ.

Theorem 1 *For $n \geq 3$, the natural autocorrelation γ of the set V of visible points of a lattice Γ exists and is supported on Γ, the weight of a point $\boldsymbol{a} \in \Gamma$ in the autocorrelation of V being given by*

$$w(\boldsymbol{a}) \;=\; \mathrm{dens}(\Gamma)\xi(n) \prod_{p|\mathrm{cont}(\boldsymbol{a})} \left(1 + \frac{1}{p^n - 2}\right), \tag{56}$$

with error term (as defined just after (10)) equal to $O(1/R)$, where the implied constant depends on $\boldsymbol{a}$ as well as on Γ.

Proof: By Prop. 5, $V - V = \Gamma$, so the autocorrelation of V (if it exists) is supported on Γ. The amplitude of a point $\boldsymbol{a} \in \Gamma$ in the autocorrelation of V is the limit as $R \to \infty$ of

$$\frac{1}{v_n R^n} \sum_{\substack{|\boldsymbol{x}|,|\boldsymbol{x}-\boldsymbol{a}|<R \\ \boldsymbol{x},\boldsymbol{x}-\boldsymbol{a}\in V}} 1 \tag{57}$$

and by Lemma 1 the existence of this limit for each $\boldsymbol{a} \in \Gamma$ is sufficient to ensure the existence of the autocorrelation.

It is convenient to drop the condition $|\boldsymbol{x}-\boldsymbol{a}|<R$ in (57), which then becomes

$$\frac{1}{v_n R^n} \sum_{\substack{|\boldsymbol{x}|<R \\ \boldsymbol{x},\boldsymbol{x}-\boldsymbol{a}\in V}} 1. \tag{58}$$

The difference between these sums is $O(1/R)$, due to the extra lattice points $\boldsymbol{x}$ within a constant distance $|\boldsymbol{a}|$ of the boundary of $B_R(\mathbf{0})$ that are included in the latter. We note that the latter sum has a natural geometric interpretation as the proportion of integer points in a large ball that are visible both from the origin and from the viewpoint $\boldsymbol{a}$. By (14) and (7), (58) can be expressed as

$$\frac{1}{v_n R^n} \sum_{\substack{\boldsymbol{x}\in\Gamma\setminus\{\mathbf{0},\boldsymbol{a}\} \\ |\boldsymbol{x}|<R}} \sum_{l|\mathrm{cont}(\boldsymbol{x})} \mu(l) \sum_{m|\mathrm{cont}(\boldsymbol{x}-\boldsymbol{a})} \mu(m) = \frac{1}{v_n R^n} \sum_{1\le l<S} \sum_{1\le m<S} \mu(l)\mu(m) \sum_{\substack{\boldsymbol{x}\in\Gamma\setminus\{\mathbf{0},\boldsymbol{a}\} \\ \boldsymbol{x}\equiv\mathbf{0} \pmod{l\Gamma} \\ \boldsymbol{x}\equiv\boldsymbol{a} \pmod{m\Gamma} \\ |\boldsymbol{x}|<R}} 1, \tag{59}$$

where $S=(R+|\boldsymbol{a}|)/L(\Gamma)$. Collecting together terms with the same value of $d=(l,m)$, noting that all solutions $\boldsymbol{x}$ of the congruences belong to $d\Gamma$ and that the congruences have no solution at all unless $\boldsymbol{a}\in d\Gamma$, and putting $l'=l/d$, $m'=m/d$, $\boldsymbol{x}'=\boldsymbol{x}/d$, $\boldsymbol{a}'=\boldsymbol{a}/d$, we obtain

$$\frac{1}{v_n R^n} \sum_{d|\mathrm{cont}(\boldsymbol{a})} \sum_{1\le l'<S/d} \sum_{\substack{1\le m'<S/d \\ (l',m')=1}} \mu(l'd)\mu(m'd) \sum_{\substack{\boldsymbol{x}'\in\Gamma\setminus\{\mathbf{0},\boldsymbol{a}'\} \\ \boldsymbol{x}'\equiv\mathbf{0} \pmod{l'\Gamma} \\ \boldsymbol{x}'\equiv\boldsymbol{a}' \pmod{m'\Gamma} \\ |\boldsymbol{x}'|<R/d}} 1. \tag{60}$$

Since l' and m' are bound variables of summation and $\boldsymbol{x}'$ and $\boldsymbol{a}'$ will not be referred to again, we can drop the dashes: from now on l and m are the new l' and m' but $\boldsymbol{a}$ is the original $\boldsymbol{a}$.

By Propositions 2 and 1, and the fact that the number of points $\boldsymbol{x}\in lm\Gamma+\boldsymbol{b}$ with $|\boldsymbol{x}|<R$ is the same as the number of points $\boldsymbol{y}\in\Gamma+(\boldsymbol{b}/lm)$ with $|\boldsymbol{y}|<R/lmd$, the inner sum is

$$\mathrm{dens}(\Gamma)v_n\left(\frac{R}{dlm}\right)^n+O\left(\frac{R}{dlm}\right)^{n-1}+O(1) \tag{61}$$

These three terms give a main term and two error terms in (60).

The first error term is majorized by

$$O\left(\frac{1}{R}\sum_{1\le l<S}\frac{1}{l^{(n-1)}}\sum_{1\le m<S}\frac{1}{m^{(n-1)}}\right)=\begin{cases} O(1/R), & \text{if } n\ge 3,\\ O((\log S)^2/R), & \text{if } n=2,\end{cases} \tag{62}$$

since the sums are convergent when $n-1 \geq 2$ and increase logarithmically when $n-1=1$.

The second error term is majorized by $O(S^2/R^n) = O(1/R^{n-2})$. So when $n \geq 3$ both error terms are $O(1/R)$ (since $S = O(R)$ and tends to 0 as $R \to \infty$).

The main term is (with $\mathcal{D} = \mathrm{dens}(\Gamma)$)

$$\mathcal{D} \sum_{d|\mathrm{cont}(\boldsymbol{a})} \sum_{1 \leq l < S/d} \sum_{\substack{1 \leq m < S/d \\ (l,m)=1}} \frac{\mu(ld)\mu(md)}{(dlm)^n} = \mathcal{D} \sum_{d|\mathrm{cont}(\boldsymbol{a})} \sum_{\substack{1 \leq l < S/d \\ (l,d)=1}} \sum_{\substack{1 \leq m < S/d \\ (m,d)=1 \\ (l,m)=1}} \frac{\mu^2(d)\mu(l)\mu(m)}{(dlm)^n} \tag{63}$$

since $\mu(ld)$ is $\mu(l)\mu(d)$ when $(l,d)=1$ and 0 otherwise, and similarly for $\mu(md)$,

$$= \mathcal{D} \sum_{d|\mathrm{cont}(\boldsymbol{a})} \frac{\mu^2(d)}{d^n} \sum_{\substack{1 \leq l < S/d \\ (l,d)=1}} \sum_{\substack{1 \leq m < S/d \\ (m,d)=1}} \frac{\mu(lm)}{(lm)^n} \tag{64}$$

since $\mu(l)\mu(m)$ is $\mu(lm)$ when $(l,m)=1$ and 0 otherwise

$$\to \mathcal{D} \sum_{d|\mathrm{cont}(\boldsymbol{a})} \frac{\mu^2(d)}{d^n} \sum_{\substack{r=1 \\ (r,d)=1}}^{\infty} \frac{\mu(r)\tau(r)}{r^n} \quad \text{as } R \to \infty \tag{65}$$

since the double sum is absolutely convergent. The difference between this limit and the partial sum (64) is $O(1/R^{n-1})$, so falls within the error term estimate $O(1/R)$.

Using (18), the expression for the limit (65) can be rearranged as

$$\begin{aligned} \mathcal{D} \sum_{\substack{d|\mathrm{cont}(\boldsymbol{a}) \\ d \text{ squarefree}}} \frac{1}{d^n} \prod_{p \nmid d} \left(1 - \frac{2}{p^n}\right) &= \mathcal{D}\xi(n) \sum_{\substack{d|\mathrm{cont}(\boldsymbol{a}) \\ d \text{ squarefree}}} \frac{1}{d^n} \prod_{p|d} \left(1 - \frac{2}{p^n}\right)^{-1} \\ &= \mathcal{D}\xi(n) \prod_{p|\mathrm{cont}(\boldsymbol{a})} \left(1 + \frac{1}{p^n}\left(1 - \frac{2}{p^n}\right)^{-1}\right) \\ &= \mathcal{D}\xi(n) \prod_{p|\mathrm{cont}(\boldsymbol{a})} \left(1 + \frac{1}{p^n - 2}\right). \end{aligned} \tag{66}$$

This completes the proof. □

To establish the existence of the autocorrelation function for the visible points when $n = 2$, we need to reduce the second error term in the above proof. This we do by modifying the argument slightly, first using the characteristic function (14) to replace the constraint on $\boldsymbol{x} - \boldsymbol{a}$ only, then discarding large values of m from the sum before using (14) again to replace the constraint on $\boldsymbol{x}$.

Theorem 2 *Theorem 1 holds for $n = 2$ with the error term increased to $O(1/\sqrt{R})$.*

PROOF: Using (14) to replace the constraint on $\boldsymbol{x} - \boldsymbol{a}$, (58) becomes

$$\frac{1}{\pi R^2} \sum_{\substack{\boldsymbol{x} \in V \\ \boldsymbol{x} \neq \boldsymbol{a} \\ |\boldsymbol{x}| < R}} \sum_{m | \mathrm{cont}(\boldsymbol{x} - \boldsymbol{a})} \mu(m) = \frac{1}{\pi R^2} \sum_{1 \leq m < S} \mu(m) \sum_{\substack{\boldsymbol{x} \in V \\ \boldsymbol{x} \neq \boldsymbol{a} \\ \boldsymbol{x} \equiv \boldsymbol{a} \pmod{m\Gamma} \\ |\boldsymbol{x}| < R}} 1, \tag{67}$$

where S is as before. The inner sum is trivially $O(R^2/m^2)$ (since $m < S = O(R)$) so the contribution to (58) from the terms with $m \geq \sqrt{R}$ is

$$O\Big(\sum_{m \geq \sqrt{R}} \frac{1}{m^2} \Big) = O(1/\sqrt{R}). \tag{68}$$

For the terms in (67) with $m < \sqrt{R}$ we use the characteristic function (14) to replace the constraint on $\boldsymbol{x}$, obtaining

$$\frac{1}{\pi R^2} \sum_{1 \leq m < \sqrt{R}} \mu(m) \sum_{\substack{\boldsymbol{x} \in \Gamma \setminus \{\boldsymbol{0}, \boldsymbol{a}\} \\ \boldsymbol{x} \equiv \boldsymbol{a} \pmod{m\Gamma} \\ |\boldsymbol{x}| < R}} \sum_{l | \mathrm{cont}(\boldsymbol{x})} \mu(l) = \frac{1}{\pi R^2} \sum_{1 \leq l < S} \sum_{1 \leq m < \sqrt{R}} \mu(l)\mu(m) \sum_{\substack{\boldsymbol{x} \in \Gamma \setminus \{\boldsymbol{0}, \boldsymbol{a}\} \\ \boldsymbol{x} \equiv \boldsymbol{0} \pmod{l\Gamma} \\ \boldsymbol{x} \equiv \boldsymbol{a} \pmod{m\Gamma} \\ |\boldsymbol{x}| < R}} 1. \tag{69}$$

This is now identical to (59), except that the second S is replaced by $\sqrt{R}$, and as before contributes to (58) a main term that tends to

$$\xi(2) \prod_{p | \mathrm{cont}(\boldsymbol{a})} \left(1 + \frac{1}{p^2 - 2} \right) \tag{70}$$

as $R \to \infty$ and two error terms $O((\log R)^2/R)$ and $O(1/\sqrt{R})$. The difference between the main term and its limit is also $O(1/\sqrt{R})$. So the total error term is majorized by $O(1/\sqrt{R})$. $\square$

Note that this modified argument does not reduce the error term when $n \geq 3$, which contains a term $O(1/R)$ arising from the boundary of the ball when $l = m = 1$. For $n = 2$, a similar but more complicated argument gives an improved error term $O(R^{-3/4}(\log R)^c)$ for some constant c, but we suppress further details here.

Remark Theorems 1 and 2 show that the amplitude of $\boldsymbol{a}$ in the autocorrelation of V depends only on the content of $\boldsymbol{a}$ and the density of Γ. Using Proposition 3, we see that the group $SL^{\pm}(n,\mathbb{Z})$ preserves both the content and the density of Γ. In fact our results are $SL^{\pm}(n,\mathbb{Z})$-invariant, but to see this we have to be a bit careful. Our definition of the autocorrelation uses the concept of natural density, averaging over balls, and these are not invariant under this group if $n > 1$. We need to show that we can replace averaging by balls by averaging over an arbitrary expanding ellipsoid.

We may argue as follows. Consider our ambient space as affine n-space $\mathbb{E}^n$ in which we have a chosen origin and a chosen measure of volume. Then $\mathbb{E}^n$ is already a group and our latttice Γ is a certain subgroup.

Now we make $\mathbb{E}^n$ into $\mathbb{R}^n$ by some choice of basis which has unit volume and put on the standard metric. Then using balls in this metric we get our natural autocorrelation, and our lattice diffracts as expected with all the right amplitudes. This is our theorem. Now we go back and do it again with a different coordinatization. The answer is the same. But seen from the point of view of the first coordinatization we have now averaged over ellipsoids. The relationship between the two bases is some element of $SL^{\pm}(n,\mathbb{R})$ which can be made arbitrary. In fact, the choice of region can be generalized even further, see Appendix A.

Diffraction spectrum of the visible points

The final step in obtaining the diffraction spectrum of the visible points V is to take the Fourier transform of the autocorrelation of V.

Theorem 3 *The diffraction spectrum of the set of visible points of an n-dimensional lattice Γ exists and is a pure point measure whose underlying point set is the set of points in $\mathbb{Q}\Gamma^*$ with squarefree denominator and whose intensity at a point with such a denominator q is*

$$\frac{1}{\zeta^2(n)} \prod_{p|q} \frac{1}{(p^n-1)^2}\,. \tag{71}$$

This measure can also be represented as

$$\mathrm{dens}(\Gamma)^2 \xi(n) \sum_{\substack{d=1 \\ d \text{ squarefree}}}^{\infty} \Big(\prod_{p|d} \frac{1}{p^{2n}-2p^n} \Big) \omega_{\Gamma^*/d}\,, \tag{72}$$

a weak-convergent sum of Dirac combs.*

PROOF: Let γ be the autocorrelation of V. The right hand side of (66) can be expressed in the form (again, with $\mathcal{D} = \mathrm{dens}(\Gamma)$)

$$w(\boldsymbol{a}) \;=\; \mathcal{D}\xi(n) \sum_{\substack{d=1 \\ d \text{ squarefree} \\ \boldsymbol{a}\in d\Gamma}}^{\infty} \frac{1}{d^n} \prod_{p|d} \left(1 - \frac{2}{p^n}\right)^{-1} . \tag{73}$$

So by Theorems 1 and 2 and Lemma 1

$$\gamma = \mathcal{D}\xi(n) \sum_{\substack{d=1 \\ d \text{ squarefree}}}^{\infty} \frac{1}{d^n} \prod_{p|d} \left(1 - \frac{2}{p^n}\right)^{-1} \omega_{d\Gamma} . \tag{74}$$

Since $\|\omega_{d\Gamma}\|_{\mathrm{loc}} = O(1)$ and the coefficient of $\omega_{d\Gamma}$ is $O(1/d^n)$ when $n \geq 2$, this sum of tempered distributions is convergent in the weak*-topology by Lemma 2 (and this is easy to see by other means in this case, where the resulting sum is uniformly discrete). Its term-by-term Fourier transform is

$$\mathcal{D}^2\xi(n) \sum_{\substack{d=1 \\ d \text{ squarefree}}}^{\infty} \frac{1}{d^{2n}} \prod_{p|d} \left(1 - \frac{2}{p^n}\right)^{-1} \omega_{\Gamma^*/d}, \tag{75}$$

which weak*-converges to the diffraction spectrum of V, since the Fourier transform operator is weak*-continuous. Since $\|\omega_{\Gamma^*/d}\|_{\mathrm{loc}} = O(d^n)$ and the coefficient of $\omega_{\Gamma^*/d}$ is $O(1/d^{2n})$, Lemma 2 tells us that the weak*-sum is a translation bounded pure point measure equal to the pointwise sum of its terms.

This establishes the series form (72) for the diffraction spectrum, The explicit values of the intensities can now be evaluated quite simply. Let $\boldsymbol{p}$ be a point in $\mathbb{Q}\Gamma^*$ with denominator q. If q is not square-free, there is no contribution from (72) at all. So let q be square-free. The terms in (72) that contribute to the amplitude at $\boldsymbol{p}$ are those with $d = mq$ (m squarefree in $\mathbb{N}$ and prime to q), so the amplitude at $\boldsymbol{p}$ is

$$\mathcal{D}^2\xi(n) \prod_{p|q} \frac{1}{p^{2n} - 2p^n} \sum_{\substack{m=1 \\ m \text{ squarefree} \\ (m,q)=1}}^{\infty} \prod_{p|m} \frac{1}{p^{2n} - 2p^n}. \tag{76}$$

This simplifies to

$$\mathcal{D}^2\xi(n) \prod_{p|q} \frac{1}{p^{2n} - 2p^n} \prod_{p\nmid q} \left(1 + \frac{1}{p^{2n} - 2p^n}\right)$$

$$= \mathcal{D}^2\xi(n)\prod_{p|q}\frac{1}{p^{2n}}\Big(1-\frac{2}{p^n}\Big)^{-1}\prod_{p\nmid q}\Big(1-\frac{1}{p^n}\Big)^2\Big(1-\frac{2}{p^n}\Big)^{-1}$$
$$= \frac{\mathcal{D}^2}{\zeta^2(n)}\prod_{p|q}\frac{1}{p^{2n}}\Big(1-\frac{1}{p^n}\Big)^{-2} \tag{77}$$

(using the Euler products in (17) and (18)) which agrees with (71). □

An explicit example of the diffraction (for $\Gamma = \mathbb{Z}^2$) is shown in [2], and compared with an optical experiment.

kth-power-free numbers

In this section we derive the diffraction spectrum of the 1-dimensional set F consisting of the kth-power-free numbers in $\mathbb{Z}$. Again, this is a set with arbitrarily long gaps but it nevertheless has a pure point diffraction spectrum. The proof of the second assertion closely parallels the corresponding proof for the visible points V in the previous two sections, with the parameter k for F playing the rôle of the dimension n of V in the formalism. There are some differences of detail, however, particularly with the error terms.

Proposition 10 *F is uniformly discrete, but has gaps of arbitrary length. Moreover, for any $L > 0$ the set of gaps of length at least L has positive density.*

Proof: The uniform discreteness is trivial since $F \subset \mathbb{Z}$. Given a positive integer r choose r integer moduli $m_1, \ldots, m_r$ that are > 1 and coprime in pairs (for example, the first r primes). By the Chinese Remainder Theorem there is an integer N with

$$N \equiv -j + 1 \pmod{m_j^k} \quad \text{for } j = 1, \ldots, r. \tag{78}$$

Now for $x \equiv N \pmod{m_1^k m_2^k \ldots m_r^k}$ we have $m_1^k \mid x$, $m_2^k \mid (x+1)$,..., $m_r^k \mid (x+r-1)$, so none of the numbers $x, x+1, \ldots, x+r-1$ is kth-power-free. The numbers x have density $(m_1 m_2 \ldots m_r)^{-k}$. □

This argument gives a distance of the order L^{kL} between gaps of length L, so again long gaps can be expected to be extremely sparse.

Proposition 11 *For $k \geq 2$, the kth-power-free integers F have a natural density given by*

$$\mathrm{dens}(F) = \frac{1}{\zeta(k)}, \tag{79}$$

with error term $O(R^{-1+(1/k)})$.

At least for the squarefree numbers, this is again a standard example of Möbius inversion, see [9, Thm. 6.6.1].

PROOF: The natural density of F is the limit as $R \to \infty$ of

$$\frac{1}{2R} \sum_{\substack{|x|<R \\ x \in F}} 1, \tag{80}$$

which by (15) is

$$\begin{aligned} \frac{1}{2R} \sum_{\substack{|x|<R \\ x \neq 0}} \sum_{m^k | x} \mu(m) &= \frac{1}{2R} \sum_{1 \leq m < R^{1/k}} \mu(m) \sum_{\substack{|x|<R \\ x \equiv 0 \pmod{m^k} \\ x \neq 0}} 1 \\ &= \sum_{1 \leq m < R} \frac{\mu(m)}{m^k} + O\left(\frac{R^{1/k}}{R}\right), \end{aligned} \tag{81}$$

since the inner sum on the right of (81) is $2R/m^k + O(1)$. This last expression tends to $1/\zeta(k)$ as $R \to \infty$ when $n \geq 2$, with the errors due to the tail of the sum and to the explicit error term both $O(R^{-1+(1/k)})$. □

Theorem 4 *For $k \geq 2$, the natural autocorrelation of the set F of kth-power-free integers exists and is supported on $\mathbb{Z}$, the amplitude of an integer a in the autocorrelation of V being given by*

$$w(a) = \xi(k) \prod_{p^k | a} \left(1 + \frac{1}{p^k - 2}\right), \tag{82}$$

with error term $O(R^{-(1-(1/k))^2})$, where the implied constant depends on a.

PROOF: Clearly the autocorrelation is supported on $\mathbb{Z}$. The amplitude of an integer a in the autocorrelation of F is the limit as $R \to \infty$ of

$$\frac{1}{2R} \sum_{\substack{|x|,|x-a|<R \\ x, x-a \ k-\text{free}}} 1 \tag{83}$$

and, as in the proof of Theorem 1, to show the existence of the autocorrelation it is enough to show that this limit exists for every a. Again as in the proof of Theorem 1, we can drop the condition $|x-a| < R$ from the sum to obtain

$$\frac{1}{2R} \sum_{\substack{|x|<R \\ x, x-a \ k-\text{free}}} 1 \tag{84}$$

with error $O(1)$.

This can be evaluated by an argument exactly parallel to that for the visible points, except for using the characteristic function (15) in place of (14). We use the argument in the modified form we used in the proof of Theorem 2 since, in the context of kth-power-free numbers, it gives an improved error term in all cases owing to the fact that, for a 1-dimensional set, errors due to the boundary of a large region are trivial. (Using the unmodified form would give error term $O(R^{-1+(2/k)})$ instead of the $O(R^{-1+(2/k)-(1/k)^2})$ we obtain here. As with the visible points, this would not be small enough to establish the existence of the limit for the squarefree numbers $k = 2$ – the paradigm case.)

Using the characteristic function (15) to replace the constraint on $x - a$, (84) becomes

$$\frac{1}{2R} \sum_{\substack{|x|<R \\ x\ k-\text{free} \\ x \neq a}} \sum_{m^k|(x-a)} \mu(m) = \frac{1}{2R} \sum_{1 \leq m < S^{1/k}} \mu(m) \sum_{\substack{|x|<R \\ x\ k-\text{free} \\ x \neq a \\ x \equiv a \pmod{m^k}}} 1, \tag{85}$$

where $S = R + |a|$. The inner sum is trivially $O(R/m^k)$ (since $m^k < S$) so the contribution to (84) from the terms with $m \geq T$, where T is a parameter tending to infinity with R to be chosen later, is

$$O\Big(\sum_{m \geq T} \frac{1}{m^k}\Big) = O(T^{1-k}). \tag{86}$$

For the terms in (85) with $m < T$ we use the characteristic function (15) to replace the constraint on x, obtaining

$$\frac{1}{2R} \sum_{1 \leq m < T} \mu(m) \sum_{\substack{|x|<R \\ x \neq 0,a \\ x \equiv a \pmod{m^k}}} \sum_{l^k|x} \mu(l) = \frac{1}{2R} \sum_{1 \leq l < R^{1/k}} \sum_{1 \leq m < T} \mu(l)\mu(m) \sum_{\substack{|x|<R \\ x \neq 0,a \\ x \equiv 0 \pmod{l^k} \\ x \equiv a \pmod{m^k}}} 1. \tag{87}$$

Collecting together terms with the same value of $d = (l, m)$, noting that all solutions x of the congruences are divisible by d^k and that the congruences have no solution unless $d^k \mid a$, and putting $l' = l/d$, $m' = m/d$, $x' = x/d^k$, $a' = a/d^k$, we obtain

$$\frac{1}{2R} \sum_{d^k|a} \sum_{1 \leq l' < R^{1/k}/d} \sum_{\substack{1 \leq m' < T/d \\ (l',m')=1}} \mu(l'd)\mu(m'd) \sum_{\substack{|x'|<R/d^k \\ x' \neq 0,a' \\ x' \equiv 0 \pmod{l'^k} \\ x' \equiv a' \pmod{m'^k}}} 1. \tag{88}$$

Again, we can drop the dashes from now on.

By the Chinese Remainder Theorem the inner sum is

$$\frac{2R}{(dlm)^k} + O(1), \tag{89}$$

giving a main term and another error term in (84).

The error term is majorized by $O(R^{-1+(1/k)}T)$.

The main term is

$$\sum_{d^k|a} \sum_{1\le l<R^{1/k}/d} \sum_{\substack{1\le m<T/d \\ (l,m)=1}} \frac{\mu(ld)\mu(md)}{(dlm)^k}$$

$$= \sum_{d^k|a} \sum_{\substack{1\le l<R^{1/k}/d \\ (l,d)=1}} \sum_{\substack{1\le m<T/d \\ (m,d)=1 \\ (l,m)=1}} \frac{\mu^2(d)\mu(l)\mu(m)}{(dlm)^k} \tag{90}$$

since $\mu(ld)$ is $\mu(l)\mu(d)$ when $(l,d)=1$ and 0 otherwise, and similarly for $\mu(md)$,

$$= \sum_{d^k|a} \frac{\mu^2(d)}{d^k} \sum_{\substack{1\le l<R^{1/k}/d \\ (l,d)=1}} \sum_{\substack{1\le m<T/d \\ (m,d)=1}} \frac{\mu(lm)}{(lm)^k} \tag{91}$$

since $\mu(l)\mu(m)$ is $\mu(lm)$ when $(l,m)=1$ and 0 otherwise

$$\to \sum_{d^k|a} \frac{\mu^2(d)}{d^k} \sum_{\substack{r=1 \\ (r,d)=1}}^{\infty} \frac{\mu(r)\tau(r)}{r^k} \quad \text{as } R\to\infty \tag{92}$$

since the double sum is absolutely convergent. The difference between this limit and the partial sum (91) is $O(1/R^{-(k-1)/k}) + O(T^{-(k-1)})$ and each of these terms is subsumed by one of the error estimates we already have.

This last expression is

$$\begin{aligned} \sum_{\substack{d^k|a \\ d \text{ squarefree}}} \frac{1}{d^k} \prod_{p\nmid d}\left(1-\frac{2}{p^k}\right) &= \xi(k) \sum_{\substack{d^k|a \\ d \text{ squarefree}}} \frac{1}{d^k} \prod_{p|d}\left(1-\frac{2}{p^k}\right)^{-1} \\ &= \xi(k) \prod_{p^k|a}\left(1+\frac{1}{p^k}\left(1-\frac{2}{p^k}\right)^{-1}\right) \\ &= \xi(k) \prod_{p^k|a}\left(1+\frac{1}{p^k-2}\right). \end{aligned} \tag{93}$$

Finally, we choose $T = R^{(1/k)-(1/k)^2}$, making both error terms $O(R^{-1+(2/k)-(1/k)^2})$. $\square$

We can now fill a small gap in our analogy by the following result.

Corollary 1 *For $k \geq 2$, we have $F - F = \mathbb{Z}$.*

PROOF: Since $F - F \subset \mathbb{Z}$, γ_ω is clearly supported on $\mathbb{Z}$. From the previous Theorem, we get $w(a) > 0$ for all $a \in \mathbb{Z}$, hence also $\mathbb{Z} \subset F - F$. $\square$

Let us finally describe the diffraction of kth-power-free integers.

Theorem 5 *The diffraction spectrum of the set F of kth-power-free integers exists and is a pure point measure. Its underlying point set is the set of numbers $a/q \in \mathbb{Q}$ with q being $(k+1)$-power-free. The diffraction intensity at a point with such a denominator q is*

$$\frac{1}{\zeta^2(k)} \prod_{p|q} \frac{1}{(p^k-1)^2}\,. \tag{94}$$

This measure can also be represented as

$$\xi(k) \sum_{\substack{d=1 \\ d \text{ squarefree}}}^{\infty} \Big(\prod_{p|d} \frac{1}{p^{2k} - 2p^k}\Big) \omega_{\mathbb{Z}/d^k}\,, \tag{95}$$

a weak-convergent sum of Dirac combs.*

PROOF: As in the proof of Theorem 3, it follows from (93) that

$$\gamma = \xi(k) \sum_{\substack{d=1 \\ d \text{ squarefree}}}^{\infty} \frac{1}{d^k} \prod_{p|d} \left(1 - \frac{2}{p^k}\right)^{-1} \omega_{d^k\mathbb{Z}}, \tag{96}$$

where γ is the autocorrelation of F and this sum of tempered distributions is convergent in the $\|\cdot\|_{\text{loc}}$-topology for $k \geq 2$. Its term-by-term Fourier transform is

$$\xi(k) \sum_{\substack{d=1 \\ d \text{ squarefree}}}^{\infty} \frac{1}{d^{2k}} \prod_{p|d} \left(1 - \frac{2}{p^k}\right)^{-1} \omega_{\mathbb{Z}/d^k}, \tag{97}$$

also convergent in the $\|\cdot\|_{\text{loc}}$-topology, and by Lemma 2 this sum weak*-converges to a translation bounded pure point measure equal to the pointwise sum of its terms. Since the Fourier transform operator is weak*-continuous this pure point distribution is the diffraction spectrum of F.

This establishes (95). To obtain (94) we note that the terms in (95) that contribute to the intensity at a/q (where q has to be $(k+1)$-free) are those with $d = mq^*$, where q^* is the square-free kernel of q and $m \in \mathbb{N}$ is square-free and prime to q. Thus the amplitude at a/q is

$$\xi(k) \prod_{p|q} \frac{1}{p^{2k} - 2p^k} \sum_{\substack{m=1 \\ m \text{ squarefree} \\ (m,q)=1}}^{\infty} \prod_{p|m} \frac{1}{p^{2k} - 2p^k}. \tag{98}$$

This is identical to (76) without the term $\mathcal{D}^2$ and with n replaced by k, so reduces to (94). □

Summary

We have demonstrated that the set of visible lattice points in dimensions $n \geq 2$ and the set of kth-power-free integers with $k \geq 2$ both possess well-defined autocorrelations and pure point diffraction spectra. We have shown how to replace the intuitive but incomplete derivation of the spectra by a rigorous number theoretical argument based on explicit computation of the autocorrelation followed by Fourier inversion and repeated application of the Poisson summation formula.

Acknowledgements

It is our pleasure to thank M. Schlottmann for a number of clarifying discussions and for carefully reading the manuscript.

Appendix A: Densities of discrete sets

Appendix B: Higher order correlations

References

[1] T. M. Apostol, *Introduction to Analytic Number Theory*, Springer, New York (1976).

[2] M. Baake, U. Grimm and D. H. Warrington, "Some remarks on the visible points of a lattice", *J. Phys.* **A27** (1994) 2669–74 and 5041 (Erratum).

[3] S. K. Berberian, *Measure and Integration*, Chelsea, New York (1970).

[4] J. M. Cowley, *Diffraction Physics*, 3rd ed., North Holland, Amsterdam (1995).

[5] P. M. Gruber and C. G. Lekkerkerker, *Geometry of Numbers*, 2nd ed., North Holland, Amsterdam (1987).

[6] E. Hewitt and K. A. Ross, *Abstract Harmonic Analysis I*, 2nd ed., Springer, New York (1979).

[7] A. Hof, "On diffraction by aperiodic structures", *Commun. Math. Phys.* **169** (1995) 25–43.

[8] A. Hof, "Diffraction by aperiodic structures", in: *The Mathematics of Long-Range Aperiodic Order*, ed. R. V. Moody, NATO ASI Series C 489, Kluwer, Dordrecht (1997), pp. 239–68.

[9] L. K. Hua, *Introduction to Number Theory*, Springer, Berlin (1982).

[10] R. Mosseri, "Visible points in a lattice", *J. Phys.* **A25** (1992) L25–9.

[11] R. V. Moody, "Meyer sets and their duals", in: *The Mathematics of Long-Range Aperiodic Order*, ed. R. V. Moody, NATO ASI Series C 489, Kluwer, Dordrecht (1997), pp. 403–41.

[12] M. Reed and B. Simon, *Methods of Modern Mathematical Physics. I: Functional Analysis*, 2nd ed., Academic Press, San Diego (1980).

[13] W. Rudin, *Functional Analysis*, 2nd ed., McGraw-Hill, New York (1991).

[14] M. R. Schroeder, *Number Theory in Science and Communication*, 3rd ed., Springer, Berlin (1997).

[15] L. Schwartz, *Théorie des Distributions*, 2nd ed., Hermann, Paris (1966).