Entropy and Diffraction of the *k*-Free Points in *n*-Dimensional Lattices

Peter A. B. Pleasants · Christian Huck

Abstract We consider the *k*th-power-free points in *n*-dimensional lattices and explicitly calculate their entropies and diffraction spectra. This is of particular interest since these sets have holes of unbounded inradius.

1 Introduction

In [4] the diffraction properties of the visible points and the *k*th-power-free numbers were studied and it was shown that these sets have positive, pure-point, translation-bounded diffraction spectra with countable, dense support. The interest of this lay in the fact that these sets fail to be Delone sets: they are uniformly discrete (subsets of lattices, in fact) but not relatively dense. The lack of relative denseness means that these sets have arbitrarily large "holes" and hence are not repetitive in the sense of [14]. It is of interest to ask for more precise information about the irregularity of these sets, and Lenz (private communication) has asked what their entropy is.

There are two kinds of entropy commonly associated with arrays of symbols (of which subsets of lattices are a particular case): *patch-counting entropy* which is defined simply by counting patches and depends only on the adjacency relation between sites, not on any metric of the ambient space; and *measure entropy* which is defined in terms of the frequency of occurrence of patches in space. The patch-counting entropy is an upper bound for the measure entropy, whatever measure is

P. A. B. Pleasants
Department of Mathematics, University of Queensland, Brisbane, QLD 4072, Australia

C. Huck (✉)
Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, Bielefeld, Germany
e-mail: huck@math.uni-bielefeld.de

used. We show that the sets considered here have measure entropy zero (relative to a canonically constructed measure) but positive patch-counting entropy, contrasting with regular model sets [18], for which both entropies are zero [5]. In [4], a model set construction for the visible points and the kth-power-free numbers was described, with the internal spaces adelic, instead of Euclidean as in more usual cut-and-project sets. In this construction, the boundaries of the windows have positive measure, however, so they are not regular model sets.

In Sect. 2, we define patch-counting and measure entropies, while in Sect. 3 we define the set of k-free points, whose entropies we investigate, and show that they possess patch frequencies which can be explicitly calculated in terms of infinite products. This is just a mild generalization to the case of lattices other than $\mathbb{Z}$ of the results of Mirsky [16] on kth-power-free integers. To keep the route to our main results as clear as possible we have been content with weak error terms in Sect. 3, but for the record we show in Sect. 8 how error terms like those in [15] carry over to the general case. Section 4 gives some examples of patch frequencies and Sect. 5 completes the calculation of the entropies, with the aid of a key lemma (for the measure entropy case) that gives a small upper bound for the frequencies of the great majority of patches. In Sect. 6, we give a short discussion of the variational principle, which relates the two kinds of entropy. In Sect. 7, we demonstrate how the results in [4] on the diffraction spectra of the kth-power-free integers and visible lattice points carry over to the general case.

For the special case of square-free numbers (resp., kth-power-free numbers), some of our results were found independently by employing alternative methods from the theory of dynamical systems by Cellarosi and Sinai [7], Cellarosi and Vinogradov [8] and by Sarnak [22]. Furthermore, these references also contain results on the ergodic properties of the underlying invariant measures that go beyond what we cover here.

In the course of the paper, we need to call on a number of standard results in number theory, which for convenience we have collected in an appendix.

Peter A. B. Pleasants gave me (CH) an early draft of this paper already in 2006. After his untimely death in 2008, Michael Baake asked me to finish the manuscript. At that time, it already contained the entire calculation of the entropies (Sects. 1–5). Moreover, Peter had planned two further sections, one on improved error terms and one on a model set construction including the sets in question together with an upper bound for the topological entropies that is intrinsic to the corresponding window. While the former is now included (Sect. 8), the latter is still work in progress. Instead, the text now has two additional sections, one on a variational principle (Sect. 6) and one on the diffraction of the sets studied here (Sect. 7).

2 Definitions of Entropy

Let X be a subset of a lattice Λ in $\mathbb{R}^n$. Given a radius $\rho > 0$ and a point $t \in \Lambda$, the ρ-patch of X at t is

$$(X - t) \cap B_\rho(\mathbf{0}),$$

the translation to the origin of the part of X within a distance ρ of t. We denote by $\mathcal{A}(\rho)$ the set of all ρ-patches of X and by $N(\rho) = |\mathcal{A}(\rho)|$ the number of distinct ρ-patches of X. Then the *patch-counting entropy* of X is

$$h_{\mathrm{pc}}(X) := \lim_{\rho \to \infty} \frac{\log_2 N(\rho)}{\rho^n v_n}, \tag{1}$$

where v_n is the volume of an n-dimensional ball of radius 1, i.e. $v_n = \pi^{n/2}/\Gamma(1 + \frac{n}{2})$ (so that the denominator is the volume of the open ball $B_\rho(\mathbf{0})$). It can be shown by a subadditivity argument that this limit exists for every $X \subset \Lambda$. In [5, Theorem 1 and Remark 2] Baake, Lenz and Richard show that, for the dynamical system of coloured Delone sets of finite local complexity, the patch-counting entropy coincides with the topological entropy; see Sect. 6 for more on the natural dynamical system associated with a subset X of Λ and the k-free points in particular.

To describe measure entropy, we must take into account densities of subsets of a lattice. If $Y \subset \Lambda$, its *density* $\delta(Y)$ is defined by

$$\delta(Y) := \lim_{R \to \infty} \frac{|Y \cap B_R(\mathbf{0})|}{R^n v_n}, \tag{2}$$

when the limit exists; cf. [4] for related ways of defining densities of discrete point sets. In cases where the limit does not exist, we can still define an *upper density*, $\bar{\delta}(Y)$ and a *lower density*, $\underline{\delta}(Y)$, by replacing the limit in (2) by lim sup or lim inf. The *frequency*, $\nu(\mathcal{P})$, of a ρ-patch $\mathcal{P}$ of X is defined by

$$\nu(\mathcal{P}) := \delta(\{t \in \Lambda : \text{the } \rho\text{-patch of } X \text{ at } t \text{ is } \mathcal{P}\}), \tag{3}$$

when this density exists. In the absence of a well defined density, we can still define an *upper frequency*, $\bar{\nu}(\mathcal{P})$ and a *lower frequency*, $\underline{\nu}(\mathcal{P})$, by replacing δ by $\bar{\delta}$ or $\underline{\delta}$. The *measure entropy* of X, which can be thought of as corresponding to the metric entropy of a dynamical system, is now defined by

$$h_{\mathrm{meas}}(X) := \lim_{\rho \to \infty} \frac{1}{\rho^n v_n} \sum_{\mathcal{P} \in \mathcal{A}(\rho)} -\nu(\mathcal{P}) \log_2 \nu(\mathcal{P}), \tag{4}$$

with the convention that $\nu \log_2 \nu = 0$ when $\nu = 0$; see Sect. 6 for details. It is defined when every patch of X has a well defined frequency, in which case a subadditivity argument again shows that the limit exists. Since $\nu \log_2 \nu$ is a convex function of ν, the sum does not decrease if we replace the $\nu(\mathcal{P})$'s by their average value, $1/N(\rho)$, to make the right side the same as the right side of (1). Hence

$$h_{\mathrm{meas}}(X) \le h_{\mathrm{pc}}(X).$$

As a simple example where these entropies differ, consider the binary sequence consisting of the binary numbers in order $(0, 1, 10, 11, 100, \ldots)$ with $n, n+1$ separated by n 1's:

<u>01110</u>1<u>11</u>11<u>11</u><u>100</u>1111<u>10</u>1111<u>11</u><u>110</u>1111111<u>111</u>111111<u>100</u>01111 . . . .

Evidently, there are very few 0's to contribute variety here. In fact the sequence of 0's has density zero, and consequently any finite word that is not all 1's has frequency zero. So $h_{meas} = 0$. But since there are 2^l possible words of length l and every word occurs somewhere, $h_{pc} = 1$.

In general, h_{pc} is a combinatorial function of the set of finite configurations that occur, while h_{meas} is a geometric function of an infinite configuration and can differ among different infinite configurations built up from the same set of finite ones, with h_{pc} being an upper bound for the possible values it can take. Of the two entropies, h_{meas} would appear to carry more physical significance.

More generally, if we have a pattern formed by labelling the points of Λ with letters from an a-letter alphabet then we can again define ρ-patches, and the entropies of the pattern are given by (1) and (4) with 2 replaced by a as the base of logarithms. A subset of Λ corresponds to a 2-letter labelling indicating whether or not a site is occupied. The reason for the patch volume in the denominator and for the choice of base of logarithms is to normalize so that the integer lattice with random labelling has both entropies 1.

There are various ways in which the definition of measure entropy might be extended to sets X for which not all patch frequencies exist. A first step would be to replace the sum in (4) by

$$\lim_{R \to \infty} \sum_{\mathcal{P} \in \mathcal{A}(\rho)} -\frac{|L(\mathcal{P}) \cap B_R(\mathbf{0})|}{R^n v_n} \log_2 \Big(\frac{|L(\mathcal{P}) \cap B_R(\mathbf{0})|}{R^n v_n}\Big),$$

where $L(\mathcal{P})$ is the set appearing in (3). This delays taking the limit, so that it has a chance of existing even when some individual patch frequencies may fail to exist. We shall not need such extensions here, however, since Theorem 1 below guarantees that, for the sets studied in this paper, all patch frequencies exist.

3 k-Free Points

As a convenient context for our results, we shall use the set $V = V(\Lambda, k)$ of k-free points of a lattice Λ in $\mathbb{R}^n$. For a point $\boldsymbol{l} \neq \mathbf{0}$ in Λ define its k-content, $c_k(\boldsymbol{l})$, to be the largest integer c such that $\boldsymbol{l} \in c^k \Lambda$. Then $c_k(\boldsymbol{l})$ is also the least common multiple of the numbers d with $d^{-k}\boldsymbol{l} \in \Lambda$, i.e. $d^{-k}\boldsymbol{l} \in \Lambda$ if and only if $d \mid c_k(\boldsymbol{l})$. For consistency and convenience, we define $c_k(\mathbf{0}) = \infty$, with the understanding that $d \mid \infty$ for any number d. The k-free points, $V = V(\Lambda, k)$, of Λ are the points with $c_k(\boldsymbol{l}) = 1$. One can see that V is non-periodic, i.e. V has no nonzero translational symmetries. As particular cases we have the visible points of Λ (with $n \geq 2$ and $k = 1$), treated in [4], and the k-free integers (with $\Lambda = \mathbb{Z}$), treated in [4,15,16]. The more general context has the advantage of avoiding duplication of near-identical proofs. When $n = k = 1$, V consists of just the two points of Λ closest to $\mathbf{0}$ on either side, and we exclude this trivial case. Since Λ is a free Abelian group of rank n, its automorphism group, $\mathrm{Aut}(\Lambda)$, is isomorphic to the matrix group $\mathrm{GL}(n, \mathbb{Z})$. Explicit isomorphisms can be found by

taking coordinates with respect to any basis of Λ. Since the action of $\mathrm{GL}(n, \mathbb{Z})$ on Λ preserves k-content, the k-free points V are invariant under the action of $\mathrm{GL}(n, \mathbb{Z})$.

Proposition 1 *V is uniformly discrete, but has arbitrarily large holes. Moreover, for any $r > 0$, there is a set of holes in V of inradius at least r whose centres have positive density.*

Proof Since $V \subset \Lambda$, the uniform discreteness is trivial. Now let $C = \{a_1, \ldots, a_s\}$ be any finite configuration of points in Λ (e.g., all points in a ball or a cube). Choose s integers $m_1, \ldots, m_s > 1$ that are pairwise coprime (e.g., the first s primes). By (39), there is a point $a \in \Lambda$ with

$$a \equiv -a_i \pmod{m_i^k \Lambda}$$

for $i = 1, \ldots, s$. Now for any $x \equiv a \pmod{m_1^k \cdots m_s^k \Lambda}$ the configuration $C + x = \{a_1 + x, \ldots, a_s + x\}$ is congruent, in the geometric sense, to C but no point in $C + x$ is in V, since $a_i + x \in m_i^k \Lambda$ for $i = 1, \ldots, s$. The points x have density $1/((m_1 \cdots m_s)^{nk} \det(\Lambda)) > 0$ by (38) and (39). $\square$

For a natural number P, we define $V_P = V_P(\Lambda, k)$ to be the set of points $l \in \Lambda \setminus \{0\}$ with $(c_k(l), P) = 1$. Clearly V_P is fully periodic with a lattice of periods that contains $P^k \Lambda$. The V_P's are partially ordered inversely to the divisibility partial order on $\mathbb{N}$, that is, $V_{PQ} \subset V_P$ for all P, Q. In fact, more precisely, $V_{PQ} = V_P \cap V_Q$. The intersection of all the V_P's is V, so if P is divisible by all primes up to a large bound V_P can be regarded as a set of "potentially k-free" points.

For a finite subset $\mathcal{F}$ of Λ and a positive integer m, we shall use

$$\mathcal{F}/m\Lambda$$

to denote the set of cosets of $m\Lambda$ in Λ that are represented in $\mathcal{F}$. We also write

$$D(\mathcal{F}) := \max_{l, m \in \mathcal{F}} \|l - m\|$$

for the diameter of $\mathcal{F}$, where $\| \cdot \|$ denotes the Euclidean norm on $\mathbb{R}^n$.

Since entropies of sets in $\mathbb{R}^n$ vary under change of scale inversely as the nth power of the scaling constant, it is sufficient to consider lattices of determinant 1. (For other lattices the formula for the entropy of V must simply be divided by the determinant of Λ.) We fix the following notation for the rest of this paper:

Λ *is a lattice of determinant* 1 *in* $\mathbb{R}^n$, λ *is the length of its shortest nonzero vector, k is a natural number (with $k \geq 2$ if $n = 1$) and V is the set of k-free points in Λ.*

Also, for subsets $X, \mathcal{P}, \mathcal{Q}$ of Λ, with X infinite but $\mathcal{P}, \mathcal{Q}$ finite, we define the locator set

$$L(X; \mathcal{P}, \mathcal{Q}) := \{t \in \Lambda : \mathcal{P} + t \subset X, \ \mathcal{Q} + t \subset \Lambda \setminus X\}$$

consisting of those lattice translations that locate $\mathcal{P}$ totally inside X and $\mathcal{Q}$ totally outside X.

The genesis of our proof of positive, but non-maximal, patch-counting entropy for the visible points is the observation that, of the four corners of any unit square of the integer lattice in the plane, at least one is invisible (because both its coordinates are even) but each of the 15 possibilities for the visibility or not of the corners, when the possibility of their all being visible is excluded, can occur, depending on the position of the square within the lattice. This is the simplest example of the fact that, in general, every ρ-patch contains an irreducible minimum of points not in V but for the remaining points in the patch we can arrange that they are visible or not, independently of each other by choosing the position of the patch in the lattice. This leads to an exponentially large number of ρ-patches, the number of which can be estimated quite accurately.

Our aim with the following lemma is to concentrate most of the necessary inclusion–exclusion arguments into a single result from which ensuing results can be fairly readily derived. For this reason it has several parameters ($\mathcal{P}, m, \boldsymbol{m}, P$ and $\boldsymbol{x}$) and three components to its error term. Until the parameters are further specified, there is no assumption that the error terms are of smaller order than the main term. To keep the proof short we have not made the error terms as small as possible—in Sect. 8 we make use of the technique of [15] to vastly improve the last error term.

Lemma 1 *Let $\mathcal{P}$ be a finite subset of Λ, $m \in \mathbb{N}$, $\boldsymbol{m} \in \Lambda$, P be a natural number coprime to m and $\boldsymbol{x} \in \mathbb{R}^n$. Then*

$$|L(V_P; \mathcal{P}, \emptyset) \cap (\boldsymbol{m} + m\Lambda) \cap B_R(\boldsymbol{x})|$$

is estimated by a main term

$$\frac{R^n v_n}{m^n} \prod_{p|P} \left(1 - \frac{|\mathcal{P}/p^k\Lambda|}{p^{nk}}\right) \tag{5}$$

with error

$$O\left(R^{1/k} + R^{n-1}(\min\{\log\log P, \log S\})^{|\mathcal{P}|} + \min\{\tau_{|\mathcal{P}|+1}(P), (S/\lambda)^{|\mathcal{P}|/k}\}\right), \tag{6}$$

where $S := R + \|\boldsymbol{x}\| + \max_{\boldsymbol{p} \in \mathcal{P}} \|\boldsymbol{p}\|$, τ_r is the r-divisor function in (35), and the O-constant depends only on Λ, k and $\mathcal{P}$.

Proof We may clearly assume that P is squarefree. For each prime p the points $\boldsymbol{t}$ with $\boldsymbol{t} + \mathcal{P} \subset V_p$ consist of $p^{nk} - |\mathcal{P}/p^k\Lambda|$ cosets of $p^k\Lambda$ in Λ (those cosets $\boldsymbol{t} + p^k\Lambda$ with $(-\boldsymbol{t} + p^k\Lambda) \cap \mathcal{P} = \emptyset$). Clearly $|\mathcal{P}/p^k\Lambda| = |\mathcal{P}|$ when $p^k\lambda > D(\mathcal{P})$. Let Q be the product of those prime factors p of P with $|\mathcal{P}/p^k\Lambda| < |\mathcal{P}|$. By the Chinese Remainder Theorem (39), $L(V_Q; \mathcal{P}, \emptyset) \cap (\boldsymbol{m} + m\Lambda)$ consists of

$$\prod_{p|Q}(p^{nk} - |\mathcal{P}/p^k\Lambda|)$$

cosets of $mQ^k\Lambda$ in Λ. For each such coset $\boldsymbol{q} + mQ^k\Lambda$ we have

$$(\boldsymbol{q} + mQ^k\Lambda) \cap L(V_P; \mathcal{P}, \emptyset) = (\boldsymbol{q} + mQ^k\Lambda) \cap L(V_{P/Q}; \mathcal{P}, \emptyset). \tag{7}$$

Now write $\mathcal{P} = \{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_r\}$. Since $\boldsymbol{p}_i + \boldsymbol{t} \in V_{P/Q}$ if and only if $c_k(\boldsymbol{p}_i + \boldsymbol{t})$ is coprime to P/Q, it follows from (31) that for each of these cosets the cardinal of $L(V_P; \mathcal{P}, \emptyset) \cap (\boldsymbol{q} + mQ^k\Lambda) \cap B_R(\boldsymbol{x})$ is

$$\sum_{\substack{\boldsymbol{t} \in \Lambda \cap B_R(\boldsymbol{x}) \\ \boldsymbol{t} - \boldsymbol{q} \in mQ^k\Lambda}} \prod_{i=1}^{r} \sum_{\substack{d|P/Q \\ d|c_k(\boldsymbol{p}_i + \boldsymbol{t})}} \mu(d).$$

Reversing the order of summation gives

$$\sum_{d_1|P/Q} \sum_{\substack{d_2|P/Q \\ d_i^k < S/\lambda \text{ for each } i}} \cdots \sum_{d_r|P/Q} \mu(d_1 d_2 \cdots d_r) \sum_{\substack{\boldsymbol{t} \in \Lambda \cap B_R(\boldsymbol{x}) \\ \boldsymbol{t} \in \boldsymbol{q} + mQ^k\Lambda \\ \boldsymbol{t} \in -\boldsymbol{p}_i + d_i^k\Lambda}} 1, \tag{8}$$

where replacing $\mu(d_1) \cdots \mu(d_r)$ by $\mu(d_1 \cdots d_r)$ is justified by the fact that the d_i's are pairwise coprime since any common factor of $c_k(\boldsymbol{p}_i + \boldsymbol{t})$ and $c_k(\boldsymbol{p}_j + \boldsymbol{t})$ divides $c_k(\boldsymbol{p}_i - \boldsymbol{p}_j)$, all of whose prime factors divide Q. Writing $d_1 \cdots d_r = d$ and noting that $(mQ, d) = 1$, we can apply (38) with Λ replaced by $m(dQ)^k\Lambda$ to obtain, for the inner sum, the estimate

$$\frac{R^n v_n}{m^n (dQ)^{nk}} + O(R^{n-1}/m^{n-1}(dQ)^{(n-1)k}) + O(1). \tag{9}$$

Substituting this estimate in (8) gives a main term

$$\frac{R^n v_n}{m^n Q^{nk}} \prod_{p|P/Q} \left(1 - \frac{r}{p^{nk}}\right)$$

with error term (6). The main term arises by removing the conditions $d_i^k < S/\lambda$ from the sum of the main term in (9) then using the fact that $\mu(d)\tau_r(d)$ (where $\tau_r(d)$ is the number of ways of expressing d as a product of r natural numbers) is a multiplicative function, whose value is $-r$ at primes and 0 at prime powers, to express the extended sum as an Euler product, as in (37). The first error term in (6) comes from the extra terms included in the extended multiple sum, so is

$$\leq \frac{r R^n v_n}{m^n Q^{nk}} \sum_{d_1^k \geq S/\lambda} \frac{1}{d_1^{nk}} \sum_{d_2=1}^{\infty} \frac{1}{d_2^{nk}} \cdots \sum_{d_r=1}^{\infty} \frac{1}{d_r^{nk}} = O\left(R^n S^{-n+(1/k)}\right),$$

since each of the $r - 1$ complete sums is $\leq \zeta(2) < 2$. The other two error terms account for the sum over the error terms in (9). The logarithms in the middle error term are necessary only in the case $n = 2$, $k = 1$, when the series $\sum d_i^{-(n-1)k}$ diverge but the partial sums can be estimated by using (36) or the standard estimate for the partial sums of the harmonic series. In all other cases these series converge and the

middle error term can be taken as $O(R^{n-1})$. (When $n = 1$, there is no middle error term, since the first error term in (9) is then the same as the last.)

Finally, summing over the cosets of $m Q^k \Lambda$ that make up $L(V_Q; \mathcal{P}, \emptyset) \cap (m + m\Lambda)$ gives the main term (5) (since $|\mathcal{P}/p^k \Lambda| = r$ when $p \nmid Q$) and increases the error term by a factor at most Q^k, which is bounded in terms of k and $\mathcal{P}$.   □

Corollary 1 *If ρ is a positive radius and P is a natural number divisible by every prime less than* $\log \rho$, *then*

$$|V_P \cap B_\rho(\mathbf{0})| = \frac{\rho^n v_n}{\zeta(nk)} + o(\rho^n),   \tag{10}$$

$$|V \cap B_\rho(\mathbf{0})| = \frac{\rho^n v_n}{\zeta(nk)} + o(\rho^n),   \tag{11}$$

$$|(V_P \setminus V) \cap B_\rho(\mathbf{0})| = o(\rho^n)   \tag{12}$$

and, for any $\mathbf{x} \in \mathbb{R}^n$,

$$|V_P \cap B_\rho(\mathbf{x})| \leq \frac{\rho^n v_n}{\zeta(nk)} + o(\rho^n),   \tag{13}$$

where ζ is the Riemann ζ-function.

Proof For (10) we use the lemma with $\mathcal{P} = \{\mathbf{0}\}$, $m = 1$, $\mathbf{x} = \mathbf{0}$ and $R = \rho$, then replace the product by $1/\zeta(nk)$ using (33), with $N = \log \rho$, and (32). This gives (10) with error term $O(\rho^n / \log^{nk-1} \rho)$.

Clearly $V_Q \cap B_\rho(\mathbf{0}) = V \cap B_\rho(\mathbf{0})$ when Q is the product of all primes less than $(\rho/\lambda)^{1/k}$ (where λ is the length of the shortest nonzero vector in Λ), giving (11), and (12) results from subtracting this from (10).

For (13) we use the lemma with P replaced by P', the product of the primes less than $\log \rho$, together with (33) and (32), and note that $V_P \subset V_{P'}$. Then

$$|V_P \cap B_\rho(\mathbf{x})| \leq |V_{P'} \cap B_\rho(\mathbf{x})| = \frac{\rho^n v_n}{\zeta(nk)} + O(\rho^n / \log^{nk-1} \rho),$$

since $\log P' = O(\log \rho)$ and hence $\log \log P'$ and $\tau(P')$ are both $O(\rho^\epsilon)$ by (34) and (35).   □

We note that (11) tells us that V has density $1/\zeta(nk)$, generalizing Propositions 6 and 11 of [4] (though the error terms are not as good as those in [4] and much worse than those in [15,16]). Also, one might regard (13) as saying that V has a "uniform upper density" (or that $\Lambda \setminus V$ has a uniform lower density).

The following two theorems carry over to k-free points the results of Mirsky [15, 16] ([15] improves the error terms in [16][1] on k-free numbers. A weaker result for squarefree numbers goes back to Pillai [20]. Again, we make no attempt in Theorem 1 to match the error term of [15], postponing this to Sect. 8.

---

[1] As pointed out by Jörg Brüdern, the work of Tsang [24] can be extended to the case of k-free numbers and gives a further small improvement.

Theorem 1 *For any two disjoint finite subsets $\mathcal{P}$ and $\mathcal{Q}$ of Λ, $L(V; \mathcal{P}, \mathcal{Q})$ has a well defined density given by*

$$\sum_{\mathcal{F} \subset \mathcal{Q}} (-1)^{|\mathcal{F}|} \prod_p \left(1 - \frac{|(\mathcal{P} \cup \mathcal{F})/p^k \Lambda|}{p^{nk}}\right).$$

Proof By the inclusion–exclusion principle (30) applied to $L(V; \mathcal{P}, \emptyset) \cap B_R(\mathbf{0})$, with P_i being the property that $\mathbf{q}_i + \mathbf{t} \in V$ (where $\mathcal{Q} = \{\mathbf{q}_1, \mathbf{q}_2, \ldots\}$), we have

$$|L(V; \mathcal{P}, \mathcal{Q}) \cap B_R(\mathbf{0})| = \sum_{\mathcal{F} \subset \mathcal{Q}} (-1)^{|\mathcal{F}|} |L(V; \mathcal{P} \cup \mathcal{F}, \emptyset) \cap B_R(\mathbf{0})|.$$

Now Lemma 1 with P equal to the product of the primes less than $\log R$ gives the estimate

$$R^n v_n \prod_{p < \log R} \left(1 - \frac{|(\mathcal{P} \cup \mathcal{F})/p^k \Lambda|}{p^{nk}}\right) + O(R^{1/k} + R^{n-1+\varepsilon})$$

for $|L(V_P; \mathcal{P} \cup \mathcal{F}, \emptyset) \cap B_R(\mathbf{0})|$, the proof of (12) of Corollary 1 shows that V_P can be replaced by V at the expense of an extra error term $O(R^n/(\log R)^{nk-1})$, and (33) allows the product to be extended over all primes with a similar extra error term. Altogether, this gives the estimate

$$R^n v_n \sum_{\mathcal{F} \subset \mathcal{Q}} (-1)^{|\mathcal{F}|} \prod_p \left(1 - \frac{|(\mathcal{P} \cup \mathcal{F})/p^k \Lambda|}{p^{nk}}\right) + O(R^n/(\log R)^{nk-1})$$

for $|L(V; \mathcal{P}, \mathcal{Q}) \cap B_R(\mathbf{0})|$.   □

Theorem 2 *For disjoint finite subsets $\mathcal{P}$ and $\mathcal{Q}$ of Λ, the following statements are equivalent:*

  (i) *$|\mathcal{P}/p^k \Lambda| < p^{nk}$ for every prime p;*
  (ii) *$L(V; \mathcal{P}, \mathcal{Q})$ is non-empty;*
 (iii) *$L(V; \mathcal{P}, \mathcal{Q})$ has positive density.*

Proof Clearly (iii) implies (ii) and, almost as clearly, (ii) implies (i), since if $\mathcal{P}$ contains a complete set of coset representatives for $p^k \Lambda$ then, for every $\mathbf{t} \in \Lambda$, some point of $\mathcal{P} + \mathbf{t}$ is in $p^k \Lambda$ (so not in V).

Now assume (i) holds. For each $\mathbf{q} \in \mathcal{Q}$ choose a different prime $p(\mathbf{q}) > (D(\mathcal{P} \cup \mathcal{Q})/\lambda)^{1/k}$ and let m be the product of the $p(\mathbf{q})$'s. By the Chinese Remainder Theorem, there is an $\mathbf{m} \in \Lambda$ such that

$$\mathbf{t} \equiv \mathbf{m} \,(\mathrm{mod}\, m^k \Lambda) \iff \mathbf{t} \equiv -\mathbf{q} \,(\mathrm{mod}\, p(\mathbf{q})^k \Lambda) \quad \forall \mathbf{q} \in \mathcal{Q}.$$

Then for $\mathbf{t} \equiv \mathbf{m} \,(\mathrm{mod}\, m^k \Lambda)$ we have $\mathcal{Q} + \mathbf{t} \subset \Lambda \setminus V$ and $\mathcal{P} + \mathbf{t} \subset V_m$ (the latter using the fact that for every $\mathbf{p} \in \mathcal{P}$ and every prime factor $p(\mathbf{q})$ of m, $\mathbf{q} + \mathbf{t} \in p(\mathbf{q})^k \Lambda$

and $\|\boldsymbol{p} - \boldsymbol{q}\| < p(\boldsymbol{q})^k \lambda$, ensuring that $\boldsymbol{p} + \boldsymbol{t} \notin p(\boldsymbol{q})^k \Lambda$. Now Lemma 1, with P the product of the primes less than $\log R$ not dividing m, gives a main term $C R^n$ with error $O(\max\{R^{n-1} \log R, R^{1/k}\})$ for the cardinal of a subset of the points $\boldsymbol{t} \in \Lambda \cap B_R(\boldsymbol{0})$ with $\mathcal{P} + \boldsymbol{t} \subset V_{mP}$ and $\mathcal{Q} + \boldsymbol{t} \subset \Lambda \setminus V$, where the constant C is positive since the product in (5) has every term positive. By (12) of Corollary 1, the number of these points with $\mathcal{P} + \boldsymbol{t} \not\subset V$ is $o(R^n)$. Hence $L(V; \mathcal{P}, \mathcal{Q})$ has positive lower density, and so, by Theorem 1, positive density.   □

An interesting feature of Theorem 2 is that the criterion (i) is independent of $\mathcal{Q}$. This means, for example, that

$$L(V; \mathcal{P}, \emptyset) \neq \emptyset \quad \Rightarrow \quad \delta(L(V; \mathcal{P}, \mathcal{Q})) > 0 \quad \forall \mathcal{Q} \text{ with } \mathcal{P} \cap \mathcal{Q} = \emptyset,$$

which tells us, in particular, that every subset of a patch of V is a patch of V.

4 Examples

Theorem 1 allows us to calculate the frequencies of ρ-patches of V in terms of the products

$$\Pi_r(nk) := \prod_{p > r^{1/nk}} \left(1 - \frac{r}{p^{nk}}\right)$$

for $r = 0, 1, \ldots, |\Lambda \cap B_\rho(\boldsymbol{0})|$. Here, we give two simple examples that both have $nk = 2$ and that have $|\Lambda \cap B_\rho(\boldsymbol{0})| = 3$ and 5, respectively. So we need the products

$\Pi_1(2) = 1/\zeta(2) = 6/\pi^2 = 0.6079271\ldots,$

$\Pi_2(2) = 0.3226340\ldots$ (the Feller–Tornier constant),

$\Pi_3(2) = 0.1254869\ldots, \quad \Pi_4(2) = 0.3785994\ldots, \quad \Pi_5(2) = 0.2733455\ldots,$

whose values can be calculated efficiently by the method described in [19].

Our first example is to find the frequencies of all 2-patches when V is the set of squarefree numbers. Here $\Lambda = \mathbb{Z}$, $n = 1$, $k = 2$ and $|\Lambda \cap B_2(\boldsymbol{0})| = 3$. Since $-1, 0, 1$ are distinct mod p^k, for every p, $|(\mathcal{P} \cup \mathcal{F})/p^k| = |\mathcal{P} \cup \mathcal{F}|$ and $\nu(\mathcal{P})$ depends only on $|\mathcal{P}|$ in this case. Table 1 gives the frequencies of 2-patches of all possible sizes, both in terms of the above products and numerically, and Fig. 1 depicts the patches themselves, with their frequencies. There are three patches each of sizes 1 and 2, and we check that the sum, $\sum \nu(\mathcal{P})$, of the frequencies of all patches is 1 and that the average patch size, $\sum \nu(\mathcal{P})|\mathcal{P}|$, is $3\delta(V) = 18/\pi^2$. The patches of size 2 are the most frequent, as is to be expected since 2 is the closest integer to $3\delta(V)$: indeed, 59 % of all locations have patches of size 2. The empty patch is by far the rarest, occurring at less than 2 % of locations. The radius $\rho = 2$ is the largest for which every subset of $\Lambda \cap B_\rho(\boldsymbol{0})$ occurs as a patch of V: of the 32 subsets of $\Lambda \cap B_3(\boldsymbol{0})$ the 3 that contain 4 or 5 consecutive points do not occur as patches of V.

Table 1 Frequencies of the 2-patches of the squarefree numbers

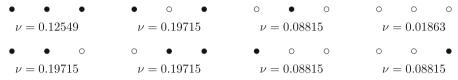| $|\mathcal{P}|$	$\nu(\mathcal{P})$	Numerical value
3	$\Pi_3(2)$	0.125486980905…
2	$\Pi_2(2) - \Pi_3(2)$	0.197147118033…
1	$\Pi_1(2) - 2\Pi_2(2) + \Pi_3(2)$	0.088145884881…
0	$1 - 3\Pi_1(2) + 3\Pi_2(2) - \Pi_3(2)$	0.018634010349…

$\nu = 0.12549$ $\nu = 0.19715$ $\nu = 0.08815$ $\nu = 0.01863$

$\nu = 0.19715$ $\nu = 0.19715$ $\nu = 0.08815$ $\nu = 0.08815$

Fig. 1 The 2-patches of the squarefree numbers ($\Lambda = \mathbb{Z}$, $k = 2$) with their frequencies accurate to five decimal places. The *black dots* are points of V and the *open circles* other lattice points. The *top row* contains the patches with mirror symmetry and the *bottom row* the two mirror image pairs

Our other example is the $\sqrt{2}$-patches of the visible points, V, in $\mathbb{Z}^2$, where $\Lambda = \mathbb{Z}^2$, $n = 2$, $k = 1$ and $|\Lambda \cap B_2(\mathbf{0})| = 5$. Figure 2 shows the different patches, up to symmetry, with their frequencies. The four patches in the top row have the full dihedral symmetry D_4; the two in the second row have symmetry D_2, and give rise to another patch on rotation through $\pi/2$; the remaining six patches have only reflection symmetry, and each gives rise to three others on rotation through $\pm\pi/2$ and π. We can again check that $\sum \nu(\mathcal{P}) = 1$ and $\sum \nu(\mathcal{P})|\mathcal{P}| = 5\delta(V) = 30/\pi^2$. This time, however, the frequencies do not depend only on $|\mathcal{P}|$, and indeed no two patches that are not symmetry related have the same frequency. Of the five patches with $|\mathcal{P}| = 4$, the symmetric one has frequency nearly five times that of each of the other four, and the ratio of the frequencies of two of the patches with $|\mathcal{P}| = 3$ is nearly 30. Surprisingly, one of the patches with $|\mathcal{P}| = 3$ (the commonest patch size) has frequency smaller than that of any patch except the empty one. The empty patch itself occurs at less than 1 in 900 locations. As in the previous example, $\sqrt{2}$ is the largest radius for which every subset of $\Lambda \cap B_\rho(\mathbf{0})$ is a patch: of the 512 subsets of $\Lambda \cap B_{\sqrt{3}}(\mathbf{0})$, the 135 that contain all four vertices of a lattice square do not occur as patches of V.

5 Entropy Calculations

Theorem 3 $h_{\mathrm{pc}}(V) = 1/\zeta(nk)$.

Proof For each radius $\rho > 0$ let $P = P(\rho)$ be the product of the primes p with $p^{nk} \leq |\Lambda \cap B_\rho(\mathbf{0})|$. Then P is divisible by every prime less than $\log \rho$ when ρ is large enough.

The ρ-patch of V at any point $t \in \Lambda$ is a subset of $(V_P - t) \cap B_\rho(\mathbf{0})$ and by (13) of Corollary 1 the cardinal of this set is at most $\rho^n v_n / \zeta(nk) + o(\rho^n)$. Also there are at most P^{nk} possibilities for $V_P - t$ as t varies, since $P^k \Lambda$ is the lattice of periods of V_P. So

0.06834   0.12096   0.01102   0.00111

0.13001   0.04129

0.02631   0.00452   0.03630
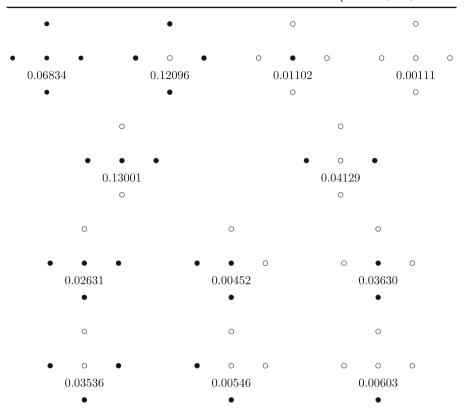
0.03536   0.00546   0.00603

Fig. 2 The $\sqrt{2}$-patches of the visible points of $\mathbb{Z}^2$, up to symmetry, with their frequencies, accurate to five decimal places

$$\log_2 N(\rho) \leq \frac{\rho^n v_n}{\zeta(nk)} + o(\rho^n) + nk \log_2 P$$

$$\leq \frac{\rho^n v_n}{\zeta(nk)} + o(\rho^n) + O(\rho^{1/k}), \tag{14}$$

since $\log_2 P = O(\rho^{1/k})$, by (34).

To bound $\log_2 N(\rho)$ below we note that every subset $\mathcal{P}$ of $V \cap B_\rho(\mathbf{0})$ is the ρ-patch of V at some point of Λ, by Theorem 2 with $\mathcal{Q} = \Lambda \cap B_\rho(\mathbf{0}) \setminus \mathcal{P}$. By (11) of Corollary 1, $|V \cap B_\rho(\mathbf{0})| = \rho^n v_n / \zeta(nk) + o(\rho^n)$, so

$$\log_2 N(\rho) \geq \frac{\rho^n v_n}{\zeta(nk)} + o(\rho^n). \tag{15}$$

On dividing by $\rho^n v_n$ and letting ρ tend to infinity, (14) and (15) give $h_{\mathrm{pc}}(V) = 1/\zeta(nk)$. $\qquad\square$

To bound the measure entropy we need the following lemma, which enables us to obtain good upper bounds for the frequency of "sparse" patches of V, i.e. patches that contain few points in comparison to their size.

Lemma 2 *Let $\mathcal{P}$ and $\mathcal{Q}$ be disjoint finite subsets of Λ, let Q be the product of all primes p with*

$$|\mathcal{Q}/p^k\Lambda| < |\mathcal{Q}| \tag{16}$$

and define

$$s := \min_{t \in L(V;\mathcal{P},\mathcal{Q})} |(\mathcal{Q}+t) \cap V_Q|.$$

Then

$$\delta(L(V;\mathcal{P},\mathcal{Q})) = O\big(4^{(D(\mathcal{Q})/\lambda)^{1/k}nk}/|\mathcal{Q}|^{s-s/nk}\big), \tag{17}$$

where the O-constant depends only on Λ.

Proof If $t \in L(V;\mathcal{P},\mathcal{Q})$ then for each $q \in \mathcal{Q} \cap (V_Q - t)$ there is a prime $p(q) \nmid Q$ with $q + t \in p(q)^k\Lambda$, and by the definition of Q these primes are distinct. By (12) of Corollary 1 with $\rho = R + \max_{q \in \mathcal{Q}} \|q\|$ and P the product of the primes less than $\log \rho$, the number of points $t \in L(V;\mathcal{P},\mathcal{Q}) \cap B_R(\mathbf{0})$ for which $p(q) \geq \log \rho$ for some $q \in \mathcal{Q} \cap (V_Q - t)$ is $o(\rho^n)$. The remaining t's in $L(V;\mathcal{P},\mathcal{Q}) \cap B_R(\mathbf{0})$ have $p(q) < \log \rho$ for each $q \in \mathcal{Q} \cap (V_Q - t)$. For the number of such t with a given set $\mathcal{Q} \cap (V_Q - t) = \{q_1, \ldots, q_t\}$ and a given ordered set of primes $\{p(q_1), \ldots, p(q_t)\}$, (38) with Λ replaced by $(p(q_1) \cdots p(q_t))^k\Lambda$, gives the estimate

$$\leq \frac{CR^n}{(p(q_1) \cdots p(q_t))^{nk}}$$

when R is large enough to ensure that $R^n > (\log \rho)^{nk|\mathcal{Q}|}$, where the constant C depends only on Λ. The sum of this over all sets of t primes not dividing Q is majorized by

$$CR^n\Big(\sum_{p \nmid Q} \frac{1}{p^{nk}}\Big)^t < CR^n\Big(\sum_{m \geq |\mathcal{Q}|^{1/nk}} \frac{1}{m^{nk}}\Big)^t < \frac{CR^n}{|\mathcal{Q}|^{s-s/nk}},$$

since $t \geq s$ and the least prime not dividing Q is $\geq |\mathcal{Q}|^{1/nk}$, by (16). There are at most Q^{nk} possibilities for $\mathcal{Q} \cap (V_Q - t)$, since $Q^k\Lambda$ is the lattice of periods of V_Q, so

$$|L(V;\mathcal{P},\mathcal{Q}) \cap B_R(\mathbf{0})| < \frac{Q^{nk}CR^n}{|\mathcal{Q}|^{s-s/nk}} + o(R^n)$$

$$< \frac{4^{(D(\mathcal{Q})/\lambda)^{1/k}nk}CR^n}{|\mathcal{Q}|^{s-s/nk}} + o(R^n)$$

for large R, where the second inequality results from (34) and the fact that $p^k \leq D(\mathcal{Q})/\lambda$ for every prime factor p of Q. The result follows on dividing by R^n and letting R tend to infinity (the existence of the limit on the left being guaranteed by Theorem 1). $\qquad\square$

Theorem 4 $h_{\text{meas}}(V) = 0$.

Proof Given $\rho > 0$ and a ρ-patch $\mathcal{P}$ of V, let $\mathcal{Q} := (\Lambda \cap B_\rho(\mathbf{0})) \setminus \mathcal{P}$ and, as in Lemma 2, define $Q = Q(\mathcal{P})$ to be the product of all primes p with $|\mathcal{Q}/p^k\Lambda| < |\mathcal{Q}|$ and

$$s = s(\mathcal{P}) := \min_{t \in L(V;\mathcal{P},\mathcal{Q})} |(\mathcal{Q} + t) \cap V_Q|.$$

By (13) of Corollary 1 and the fact that $V \subset V_P$ with P the product of primes less than $\log \rho$, we have

$$|\mathcal{Q}| > \left(1 - \frac{1}{\zeta(nk)}\right) v_n \rho^n - o(\rho^n) > \frac{v_n \rho^n}{2^{nk}} \tag{18}$$

for large enough ρ.

Now put $S = S(\rho) := \rho^n/\sqrt{\log_2 \rho}$. We shall calculate separately the contributions to the measure entropy h_{meas} of the ρ-patches $\mathcal{P}$ of V with $s(\mathcal{P}) \geq S$ and those with $s(\mathcal{P}) < S$. The former patches have small frequency and the latter are few in number.

For the ρ-patches with $s(\mathcal{P}) \geq S$, Lemma 2 and (18) give

$$-\log_2 \nu(\mathcal{P}) > \frac{n}{2} S \log_2 \rho - O\left(D(\mathcal{Q})^{1/k}\right) = \frac{n}{2} S \log_2 \rho - O\left(\rho^{1/k}\right) > \frac{S \log_2 \rho}{3}$$

for large enough ρ which, since $-\log_2 \nu$ is decreasing but $-\nu \log_2 \nu$ is increasing for $\nu \in (0, 1/e]$, gives the estimate

$$-\nu(\mathcal{P}) \log_2 \nu(\mathcal{P}) = O(2^{-S \log_2 \rho/3} S \log_2 \rho).$$

Since there are at most $2^{|\Lambda \cap B_\rho(\mathbf{0})|}$ ρ-patches in all, the contribution of the ρ-patches $\mathcal{P}$ with $s(\mathcal{P}) \geq S$ to the sum on the right of (4) is

$$O\left(2^{|\Lambda \cap B_\rho(\mathbf{0})| - (S \log_2 \rho)/3} S \log_2 \rho\right) = O\left(2^{-\frac{\rho^n}{4}\sqrt{\log_2 \rho}} \rho^n \sqrt{\log_2 \rho}\right) = o(1). \tag{19}$$

Turning to the ρ-patches with $s(\mathcal{P}) < S$, denote this set of patches by $\mathcal{B} \subset \mathcal{A}(\rho)$ and let F be their combined frequency. The contribution of these patches to the sum on the right of (4) is

$$\sum_{\mathcal{P} \in \mathcal{B}} -\nu(\mathcal{P}) \log_2 \nu(\mathcal{P})$$

which, since $\nu \log_2 \nu$ is a convex function of ν, does not decrease if we replace the $\nu(\mathcal{P})$'s by their average value, $F/|\mathcal{B}|$. So this contribution is

$$\leq F \log_2 |\mathcal{B}| - F \log_2 F \leq \log_2 |\mathcal{B}| + \frac{\log_2 e}{e}. \tag{20}$$

To bound $|\mathcal{B}|$ we note that if $\mathcal{P} \in \mathcal{B}$ then there is a $t \in L(V; \mathcal{P}, \mathcal{Q})$ with $|\mathcal{Q} \cap (V_Q - t)| < S$. Since $\mathcal{P} \subset V - t \subset V_Q - t$, $\mathcal{Q} = (\Lambda \cap B_\rho(\mathbf{0})) \setminus \mathcal{P}$, and $Q^k \Lambda$ is the lattice of periods of V_Q, $\mathcal{P}$ and $\mathcal{Q}$ are completely determined by this subset of $B_\rho(\mathbf{0})$ and by t modulo $Q^k \Lambda$. There are Q^{nk} cosets of $Q^k \Lambda$ in Λ and the number of subsets of $\Lambda \cap B_\rho(\mathbf{0})$ with fewer than S members is bounded above by

$$\sum_{i=0}^{\lfloor S \rfloor} \binom{|\Lambda \cap B_\rho(\mathbf{0})|}{i} \leq (\lfloor S \rfloor + 1) \binom{|\Lambda \cap B_\rho(\mathbf{0})|}{\lfloor S \rfloor} \leq 2S \left(\frac{e|\Lambda \cap B_\rho(\mathbf{0})|}{S} \right)^S$$

for large ρ, by (40). Hence the bound on the right of (20) is majorized by

$$S \log_2(e|\Lambda \cap B_\rho(\mathbf{0})|/S) + \log_2 2eS + nk \log_2 Q$$
$$= O\left(\frac{\rho^n \log_2 \log_2 \rho}{\sqrt{\log_2 \rho}} \right) + O(\log_2 \rho) + O(\rho^{1/k}) = o(\rho^n). \tag{21}$$

Since the contributions (19) and (21) are both $o(\rho^n)$, $h_{\text{meas}}(V) = 0$. $\square$

Note on patch shapes. On the general principle of the isotropy of space, we have used spherical patches throughout and measured densities and frequencies through expanding spherical regions; but the results we obtain are independent of the shapes of these patches and regions: all our point-counting estimates stem from (38) which remains valid for an arbitrary expanding region in place of the expanding ball, with main term the volume of the region (using the volume of the fundamental region of the lattice as a unit) and an error term of smaller order provided the boundary of the region has n-dimensional measure zero. It is not even necessary for the shape of the density-defining regions to be the same as the (also expanding) patch shape.

6 Variational Principle

Endowing the power set $\{0, 1\}^\Lambda$ of the lattice Λ with the product topology of the discrete topology on $\{0, 1\}$, it becomes a compact topological space (by Tychonov's theorem). This topology is in fact generated by the metric d defined by

$$d(X, Y) := \min \left\{ 1, \inf\{\varepsilon > 0 \mid X \cap B_{1/\varepsilon}(\mathbf{0}) = Y \cap B_{1/\varepsilon}(\mathbf{0})\} \right\}$$

for subsets X, Y of Λ. Then $(\{0, 1\}^\Lambda, \Lambda)$ is a *topological dynamical system*, i.e. the natural translational action of the group Λ on $\{0, 1\}^\Lambda$ is continuous.

Now let X be a subset of Λ. The closure $\mathbb{X}(X)$ of the set of lattice translations $t + X$ ($t \in \Lambda$) of X in $\{0, 1\}^\Lambda$ gives rise to the topological dynamical system $(\mathbb{X}(X), \Lambda)$, i.e. $\mathbb{X}(X)$ is a compact topological space on which the action of Λ is continuous; cf. [5] and references therein for details. Denote by $\mathcal{M}(\mathbb{X}(X), \Lambda)$ the set of Λ-invariant probability measures on $\mathbb{X}(X)$ with respect to the Borel σ-algebra on $\mathbb{X}(X)$, i.e. the smallest σ-algebra on $\mathbb{X}(X)$ which contains the open subsets of $\mathbb{X}(X)$. For a fixed such measure μ and a radius $\rho > 0$, let $h_\rho(\mu)$ be the entropy of μ restricted to $\mathcal{A}(\rho)$, i.e.

$$h_\rho(\mu) := \sum_{\mathcal{P} \in \mathcal{A}(\rho)} -\mu(C_\mathcal{P}) \log_2 \mu(C_\mathcal{P}),$$

where $\mathcal{A}(\rho)$ denotes the set of ρ-patches of X and $C_\mathcal{P}$ is the set of elements of $\mathbb{X}(X)$ whose ρ-patch at $\mathbf{0}$ is $\mathcal{P}$, the so-called cylinder set with respect to $\mathcal{P}$. The *metric entropy* of μ is then given by the limit

$$h(\mu) := \lim_{\rho \to \infty} \frac{h_\rho(\mu)}{\rho^n v_n},$$

which exists by a subadditivity argument; cf. [6] and also see [10,13,25]. As in Sect. 2, replacing the $\mu(C_\mathcal{P})$'s by their average value, $1/N(\rho)$, we see that

$$h(\mu) \le h_{\mathrm{pc}}(X) \quad \forall \mu \in \mathcal{M}(\mathbb{X}(X), \Lambda).$$

Since the topological entropy $h_{\mathrm{top}}(\mathbb{X}(X), \Lambda)$ of $(\mathbb{X}(X), \Lambda)$ coincides with $h_{\mathrm{pc}}(X)$ by [5, Theorem 1 and Remark 2], the variational principle for lattice actions on compact spaces here reads as follows; cf. [6] and [21, Sect. 6], the latter being an extension of the case $n = 1$ from [9,25]. An elementary proof can be found in [17]. Note that the additional statement follows from the expansiveness of the action of Λ on $\mathbb{X}(X)$.

Theorem 5 (Variational principle)

$$\sup_{\mu \in \mathcal{M}(\mathbb{X}(X), \Lambda)} h(\mu) = h_{\mathrm{pc}}(X).$$

Moreover, the supremum is achieved at some measure.   □

In case of V, $\mathbb{X}(V)$ will also contain the empty set (cf. Proposition 1) and various other subsets of Λ and thus admits many Λ-invariant probability measures. In fact, we shall now show that $\mathbb{X}(V)$ coincides with the set of *admissible* subsets A of Λ, i.e. subsets A of Λ having the property that every finite subset $\mathcal{P}$ of A satisfies criterion (i) of Theorem 2; compare [22, Theorem 8(i)]. We denote the set of all admissible subsets of Λ by $\mathbb{A}$.

Theorem 6 $\mathbb{X}(V) = \mathbb{A}$.

Proof Since $V \in \mathbb{A}$ (otherwise some point of V is in $p^k \Lambda$ for some prime p, a contradiction) and since $\mathbb{A}$ is a Λ-invariant and closed subset of $\{0, 1\}^\Lambda$, it follows that $\mathbb{A}$ contains $\mathbb{X}(V)$. For the other inclusion, let $A \in \mathbb{A}$. Then, for any $\rho > 0$, Theorem 2 applied to the finite subset $A_\rho = A \cap B_\rho(\mathbf{0})$ of A implies the existence of a $t_\rho \in L(V; A_\rho, \Lambda \cap B_\rho(\mathbf{0}) \setminus A_\rho)$. It follows that $A \in \mathbb{X}(V)$.   □

Moreover, one has $h_{\mathrm{pc}}(V) = 1/\zeta(nk)$ by Theorem 3. Consider the frequency function ν from above which gives the frequencies $\nu(\mathcal{P})$ of occurence of ρ-patches $\mathcal{P}$ of V in space. The function ν, regarded as a function on the cylinder sets by setting $\nu(C_\mathcal{P}) := \nu(\mathcal{P})$, is finitely additive on the cylinder sets with $\nu(\mathbb{X}(V)) = \sum_{\mathcal{P} \in \mathcal{A}(\rho)} \nu(C_\mathcal{P}) = 1$. Since the family of cylinder sets is a (countable) semi-algebra that generates the Borel

σ-algebra on $\mathbb{X}(V)$, one can use the method from [25, Sect. 0.2] to show that ν extends uniquely to a probability measure on $\mathbb{X}(V)$. Moreover, this probability measure can be seen to be Λ-invariant. This shows that the measure entropy $h_{\text{meas}}(V)$ is indeed a metric entropy of a Λ-invariant probability measure on $\mathbb{X}(V)$. Certainly, an explicit characterisation of $(\mathcal{M}\mathbb{X}(V), \Lambda)$ together with the corresponding metric entropies (in particular those measures $\mu \in \mathcal{M}(\mathbb{X}(V), \Lambda)$ with maximal entropy, i.e. $h(\mu) = 1/\zeta(nk)$) would be desirable (but not simple).

7 Diffraction Spectrum

In the following, we assume that the reader is acquainted with the mathematics of diffraction as carefully laid out in [4]; see also [2] and references therein for a review. We shall also use the notation and results from that text. In fact, the proofs presented below are straightforward modifications of the corresponding proofs in [4] and are only included for the reader's convenience. For an alternative derivation of the diffraction spectrum in case of the visible lattice points, see [23, Sect. 5a].

A Dirichlet series we shall encounter below is

$$\xi(s) := \sum_{m=1}^{\infty} \frac{\mu(m)\tau(m)}{m^s} = \prod_p \left(1 - \frac{2}{p^s}\right), \tag{22}$$

which is absolutely convergent for $\Re(s) > 1$, where τ is the ordinary divisor function in (35).

Theorem 7 *The natural autocorrelation of V exists and is supported on Λ, the weight of a point $\boldsymbol{a} \in \Lambda$ in the autocorrelation of V being given by*

$$w(\boldsymbol{a}) = \xi(nk) \prod_{p | c_k(\boldsymbol{a})} \left(1 + \frac{1}{p^{nk} - 2}\right),$$

with error term equal to $O(R^{-(1-(1/k))^2})$ for $n = 1$ and $k \geq 2$, $O(R^{-1/2})$ for $n = 2$ and $k = 1$ and $O(R^{-1})$ otherwise, where, in any case, the implied constant depends on $\boldsymbol{a}$ as well as on Λ. (For lattices Λ with determinant $\neq 1$ the weights above must be divided by $\det(\Lambda)$.)

Proof Since the cases $n = 1$, $k \geq 2$ and $n \geq 2$, $k = 1$ were already treated in [4, Theorems 1, 2 and 4], we may assume that $n, k \geq 2$. Since $V - V \subset \Lambda$, the autocorrelation of V (if it exists) can only be supported on Λ. The weight of a point $\boldsymbol{a} \in \Lambda$ in the autocorrelation of V is the limit as $R \to \infty$ of

$$\frac{1}{R^n v_n} \sum_{\boldsymbol{x}, \boldsymbol{x} - \boldsymbol{a} \in V \cap B_R(\mathbf{0})} 1 \tag{23}$$

and, by [4, Lemma 1], the existence of this limit for each $\boldsymbol{a} \in \Lambda$ is sufficient to ensure the existence of the autocorrelation.

It is convenient to drop the condition $x - a \in B_R(0)$ in (23), which then becomes

$$\frac{1}{R^n v_n} \sum_{\substack{x, x-a \in V \\ x \in B_R(0)}} 1 . \tag{24}$$

The difference between these sums is $O(1/R)$ by (38), due to the extra lattice points x within a constant distance $\|a\|$ of the boundary of $B_R(0)$ that are included in the latter. By (31), this can be written as

$$\frac{1}{R^n v_n} \sum_{x \in \Lambda \cap B_R(0) \setminus \{0, a\}} \sum_{l | c_k(x)} \mu(l) \sum_{m | c_k(x-a)} \mu(m) .$$

Reversing the order of summation gives

$$\frac{1}{R^n v_n} \sum_{1 \le l < S^{\frac{1}{k}}} \sum_{1 \le m < S^{\frac{1}{k}}} \mu(l)\mu(m) \sum_{\substack{x \in \Lambda \cap B_R(0) \setminus \{0, a\} \\ x \in l^k \Lambda \\ x - a \in m^k \Lambda}} 1 ,$$

where $S := (R + \|a\|)/\lambda$. Collecting terms with the same value of $d = (l, m)$, noting that all x in the inmost sum belong to $d^k \Lambda$ and that there is no such x unless $a \in d^k \Lambda$, and putting $l' := l/d, m' := m/d, x' := x/d^k, a' := a/d^k$, we obtain

$$\frac{1}{R^n v_n} \sum_{d | c_k(a)} \sum_{1 \le l' < S^{\frac{1}{k}}/d} \sum_{\substack{1 \le m' < S^{\frac{1}{k}}/d \\ (l', m') = 1}} \mu(l'd)\mu(m'd) \sum_{\substack{x' \in \Lambda \cap B_{R/d^k}(0) \setminus \{0, a'\} \\ x' \in l'^k \Lambda \\ x' - a' \in m'^k \Lambda}} 1 . \tag{25}$$

Since l' and m' are bound variables of summation and x' and 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 a is the original a.

By (38) with Λ replaced by $(lm)^k \Lambda$, the inmost sum is

$$v_n \Big(\frac{R}{(dlm)^k}\Big)^n + O\Big(\frac{R}{(dlm)^k}\Big)^{n-1} + O(1) .$$

These three terms give a main term and two error terms in (25).

The first error term is majorized by

$$O\Big(\frac{1}{R} \sum_{1 \le l < S^{\frac{1}{k}}} \frac{1}{l^{k(n-1)}} \sum_{1 \le m < S^{\frac{1}{k}}} \frac{1}{m^{k(n-1)}}\Big) = O(1/R)$$

since the sums are convergent due to $k(n-1) \ge 2$.

The second error term is majorized by

$$O\Big(\frac{S^{2/k}}{R^n}\Big) = O\Big(\frac{1}{R^{n-2/k}}\Big) = O(1/R)$$

since $S = O(R)$ and $k(n-1) \geq 2$. So both error terms are $O(1/R)$ and thus tend to 0 as $R \to \infty$.

The main term is

$$\sum_{d|c_k(\boldsymbol{a})} \quad \sum_{\substack{1 \leq l < S^{\frac{1}{k}}/d \\ (l,m)=1}} \quad \sum_{1 \leq m < S^{\frac{1}{k}}/d} \frac{\mu(ld)\mu(md)}{(dlm)^{nk}}$$

$$= \sum_{d|c_k(\boldsymbol{a})} \quad \sum_{\substack{1 \leq l < S^{\frac{1}{k}}/d \\ (l,d)=1}} \quad \sum_{\substack{1 \leq m < S^{\frac{1}{k}}/d \\ (m,d)=1 \\ (l,m)=1}} \frac{\mu(ld)\mu(md)}{(dlm)^{nk}}$$

$$= \sum_{d|c_k(\boldsymbol{a})} \frac{\mu^2(d)}{d^{nk}} \quad \sum_{\substack{1 \leq l < S^{\frac{1}{k}}/d \\ (l,d)=1}} \quad \sum_{\substack{1 \leq m < S^{\frac{1}{k}}/d \\ (m,d)=1}} \frac{\mu(lm)}{(lm)^{nk}}$$

since μ is multiplicative and $\mu(lm) = 0$ when $(l, m) \neq 1$. Since the last double sum is absolutely convergent, this converges to

$$\sum_{d|c_k(\boldsymbol{a})} \frac{\mu^2(d)}{d^{nk}} \sum_{\substack{r=1 \\ (r,d)=1}}^{\infty} \frac{\mu(r)\tau(r)}{r^{nk}} \tag{26}$$

as $R \to \infty$. The difference between this limit and the partial sum above is $O(1/R^{nk-1})$, so falls within the error estimate $O(1/R)$.

Using (22), the expression for the limit (26) can be rearranged as

$$\sum_{\substack{d|c_k(\boldsymbol{a}) \\ d \text{ squarefree}}} \frac{1}{d^{nk}} \prod_{p \nmid d} \Big(1 - \frac{2}{p^{nk}}\Big)$$

$$= \xi(nk) \sum_{\substack{d|c_k(\boldsymbol{a}) \\ d \text{ squarefree}}} \frac{1}{d^{nk}} \prod_{p | d} \Big(1 - \frac{2}{p^{nk}}\Big)^{-1}$$

$$= \xi(nk) \prod_{p|c_k(\boldsymbol{a})} \Big(1 + \frac{1}{p^{nk}}\Big(1 - \frac{2}{p^{nk}}\Big)^{-1}\Big)$$

$$= \xi(nk) \prod_{p|c_k(\boldsymbol{a})} \Big(1 + \frac{1}{p^{nk} - 2}\Big).$$

This completes the proof. $\square$

Corollary 2 $V - V = \Lambda$.

Proof Trivially, one has $V - V \subset \Lambda$. From Theorem 7, one gets $w(\boldsymbol{a}) > 0$ for all $\boldsymbol{a} \in \Lambda$ and thus also $\Lambda \subset V - V$ by [4, Lemma 1]. □

The *dual* or *reciprocal lattice* Λ^* of Λ is

$$\Lambda^* := \{\boldsymbol{y} \in \mathbb{R}^n \mid \boldsymbol{y} \cdot \boldsymbol{x} \in \mathbb{Z} \text{ for all } \boldsymbol{x} \in \Lambda\}$$

By definition, the *denominator* q of a point $\boldsymbol{p} \in \mathbb{Q}\Lambda^*$ is the smallest number $a \in \mathbb{N}$ with $a\boldsymbol{p} \in \Lambda^*$. This is also the greatest common divisor of the numbers $a \in \mathbb{N}$ with $a\boldsymbol{p} \in \Lambda^*$, i.e. $a\boldsymbol{p} \in \Lambda^*$ if and only if $q \mid a$.

Theorem 8 *The diffraction measure $\widehat{\gamma}$ of the autocorrelation γ of V exists and is a positive, pure-point, translation-bounded measure which is concentrated on the set of points in $\mathbb{Q}\Lambda^*$ with $(k+1)$-free denominator and whose intensity at a point with such a denominator q is given by*

$$\frac{1}{\zeta^2(nk)} \prod_{p|q} \frac{1}{(p^{nk} - 1)^2} . \tag{27}$$

This measure can also be interpreted as

$$\widehat{\gamma} = \xi(nk) \sum_{\substack{d=1 \\ d \text{ squarefree}}}^{\infty} \left(\prod_{p|d} \frac{1}{p^{2nk} - 2p^{nk}}\right) \omega_{\Lambda^*/d^k} , \tag{28}$$

a weak*-convergent sum (in fact, even $\|\cdot\|_{\mathrm{loc}}$-convergent sum) of Dirac combs. (For lattices Λ with determinant $\neq 1$ the above formulas must be divided by the square of $\det(\Lambda)$.)

Proof Let γ be the autocorrelation of V. As shown in the proof of Theorem 7, one has

$$w(\boldsymbol{a}) = \xi(nk) \sum_{\substack{d=1 \\ d \text{ squarefree} \\ \boldsymbol{a} \in d^k \Lambda}}^{\infty} \frac{1}{d^{nk}} \prod_{p|d} \left(1 - \frac{2}{p^{nk}}\right)^{-1} .$$

So by Theorem 7 and [4, Lemma 1] one obtains

$$\gamma = \xi(nk) \sum_{\substack{d=1 \\ d \text{ squarefree}}}^{\infty} \frac{1}{d^{nk}} \prod_{p|d} \left(1 - \frac{2}{p^{nk}}\right)^{-1} \omega_{d^k \Lambda} .$$

Since $\|\omega_{d^k\Lambda}\|_{\mathrm{loc}} = O(1)$ and the coefficient of $\omega_{d^k\Lambda}$ is $O(1/d^{nk})$, this sum of tempered distributions is convergent in the weak*-topology by [4, Lemma 2]. By the

Poisson summation formula for lattice Dirac combs [4, Eq. 31], its term-by-term Fourier transform is

$$\widehat{\gamma} = \xi(nk) \sum_{\substack{d=1 \\ d \text{ squarefree}}}^{\infty} \frac{1}{d^{2nk}} \prod_{p|d} \left(1 - \frac{2}{p^{nk}}\right)^{-1} \omega_{\Lambda^*/d^k}$$

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

which weak*-converges to the diffraction measure of V, since the Fourier transform operator is weak*-continuous. Since $\|\omega_{\Lambda^*/d^k}\|_{\text{loc}} = O(d^{nk})$ and the coefficient of ω_{Λ^*/d^k} is $O(1/d^{2nk})$, the weak*-sum is a translation-bounded pure-point measure equal to the pointwise sum of its terms by [4, Lemma 2].[2] This establishes the series form (28) for the diffraction spectrum.

The explicit values of the intensities can now be calculated as follows. Let $\boldsymbol{p}$ be a point in $\mathbb{Q}\Lambda^*$ with denominator q. We can assume that q is $(k + 1)$-free, since otherwise there is no contribution to (28) at all. The terms in (28) that contribute to the intensity at $\boldsymbol{p}$ are those with $d = mq^*$, where q^* is the squarefree kernel of q and $m \in \mathbb{N}$ is squarefree and coprime to q. Thus the intensity at $\boldsymbol{p}$ is

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

Using the Euler products in (32) and (22) this simplifies to

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

$$= \xi(nk) \prod_{p|q} \frac{1}{p^{2nk}} \left(1 - \frac{2}{p^{nk}}\right)^{-1} \prod_{p \nmid q} \left(1 - \frac{1}{p^{nk}}\right)^2 \left(1 - \frac{2}{p^{nk}}\right)^{-1}$$

$$= \frac{1}{\zeta^2(nk)} \prod_{p|q} \frac{1}{p^{2nk}} \left(1 - \frac{1}{p^{nk}}\right)^{-2},$$

which agrees with (27). □

One explicitly sees that $\widehat{\gamma}$ above is fully translation invariant, with lattice of periods Λ^*, in accordance with Theorem 1 of [1]. Moreover, since the action of the group of

---

[2] Note that $\widehat{\gamma}$ is even a $\| \cdot \|_{\text{loc}}$-convergent sum of Dirac combs. Since convergence with respect to the local norm preserves the spectral type, it is thus clear that $\widehat{\gamma}$ is a pure-point measure; cf. [3, Theorem 8.4].

automorphisms of Λ^*, $\mathrm{Aut}(\Lambda^*) \simeq \mathrm{GL}(n, \mathbb{Z})$, on $\mathbb{Q}\Lambda^*$ preserves the denominator, $\widehat{\gamma}$ is $(\Lambda^* \rtimes \mathrm{Aut}(\Lambda^*))$-symmetric. In particular, both $\widehat{\gamma}$ and the set V itself are $\mathrm{GL}(n, \mathbb{Z})$-symmetric.

8 Improving the Error Terms

What has kept the error term large in the argument as we have presented it so far is the last term of (6), with $|\mathcal{P}|$ in the exponent of S (in the second component of the minimum). This arose from the $O(1)$ error term in (9), when (9) was substituted for the inner sum in (8). The $O(1)$ error term was not even a boundary effect: it was caused solely by lattices whose determinants are much larger than the volume of the region in which points are being counted. The result of this was to put the burden of keeping the last term of (6) small onto P (which occurs in the first component of the minimum), causing an increase in the error due to the tail of the ζ-function product. Mirsky's idea in [16] and [15] was to show that the terms with some d_i large contribute a negligible amount to (8) and can be discarded before the substitution of (9) is made. The remaining terms have the individual d_i's so well bounded that the second component of the minimum can take over the role of providing a respectable error term, freeing P to be assigned a much larger value and thus reducing the size of the tail of the ζ-function product.

Let $r \in \mathbb{Z}^+$, let $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_r \in \Lambda$, and let $m_1, \ldots, m_r \in \mathbb{N}$. Define the symbol $E\left({m_1, \ldots, m_r \atop \boldsymbol{p}_1, \ldots, \boldsymbol{p}_r}\right)$ as 1 or 0 according to the system of congruences in $\boldsymbol{t} \in \Lambda$,

$$\boldsymbol{t} + \boldsymbol{p}_i \in m_i \Lambda \quad (1 \le i \le r), \tag{29}$$

being solvable or not. Further, for a positive real number R and a point $\boldsymbol{x} \in \mathbb{R}^n$, let $T\left(\boldsymbol{x}; R; {m_1, \ldots, m_r \atop \boldsymbol{p}_1, \ldots, \boldsymbol{p}_r}\right)$ denote the number of points $\boldsymbol{t} \in \Lambda$ such that

$$\boldsymbol{t} \in B_R(\boldsymbol{x}),$$
$$\boldsymbol{t} + \boldsymbol{p}_i \in m_i \Lambda \quad (1 \le i \le r).$$

We denote by $[m_1, \ldots, m_r]$ the least common multiple of $m_1, \ldots, m_r$. Further, (m_i, m_j) denotes the greatest common divisor of m_i and m_j. For brevity, let $c(\boldsymbol{l})$ denote the 1-content of a nonzero point $\boldsymbol{l} \in \Lambda$.

Lemma 3 *The system* (29) *of congruences is soluble if and only if*

$$(m_i, m_j) \mid c(\boldsymbol{p}_i - \boldsymbol{p}_j) \quad (1 \le i < j \le r).$$

In the case of solubility, the solutions form precisely one residue class

$$(\mathrm{mod} \ [m_1, \ldots, m_r]\Lambda).$$

Proof This is an immediate consequence of [15, Lemma 1] applied to each coordinate with respect to a basis of Λ. □

If $m_1, \ldots, m_r$ are pairwise coprime, the last result boils down to the Chinese Remainder Theorem (39). In fact, only this special case will be needed in Theorem 9 below. However, since the subsequent lemmas may be of independent interest, we prefer to stick to the general case.

Lemma 4

$$T\left(\boldsymbol{x}; R; \begin{matrix} m_1, \ldots, m_r \\ \boldsymbol{p}_1, \ldots, \boldsymbol{p}_r \end{matrix}\right) = R^n v_n \frac{E\left(\begin{matrix} m_1, \ldots, m_r \\ \boldsymbol{p}_1, \ldots, \boldsymbol{p}_r \end{matrix}\right)}{[m_1, \ldots, m_r]^n} + O\left(\frac{R^{n-1}}{[m_1, \ldots, m_r]^{n-1}}\right) + O(1),$$

where the implied O-constants depend only on Λ.

Proof This is an immediate consequence of Lemma 3 together with (38) applied to the lattice $[m_1, \ldots, m_r]\Lambda$. □

Lemma 5

$$\frac{E\left(\begin{matrix} m_1, \ldots, m_r \\ \boldsymbol{p}_1, \ldots, \boldsymbol{p}_r \end{matrix}\right)}{[m_1, \ldots, m_r]} \leq \frac{K}{m_1 \cdots m_r}.$$

where K depends only on $r, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_r$.

Proof This follows along the same lines as Lemma 3 of [15] by employing Lemma 3 instead of [15, Lemma 1]. □

For points $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_r \in \Lambda$ and positive real numbers R and α, denote by $L(\boldsymbol{x}; R; \boldsymbol{p}_1, \ldots, \boldsymbol{p}_r; \alpha)$ the cardinality of systems $(\boldsymbol{t}, a_1, \ldots, a_r)$ of lattice points $\boldsymbol{t} \in \Lambda$ and numbers $a_1, \ldots, a_r \in \mathbb{N}$ such that

$$\begin{aligned} &\boldsymbol{t} \in B_R(\boldsymbol{x}), \\ &\boldsymbol{t} + \boldsymbol{p}_i \in a_i^k \Lambda \quad (1 \leq i \leq r), \\ &a_1 \cdots a_r > R^\alpha. \end{aligned}$$

Lemma 6

$$L(\boldsymbol{x}; R; \boldsymbol{p}_1, \ldots, \boldsymbol{p}_r; \alpha) = O\left(R^{n-\alpha(nk-1)+\varepsilon}\right) + O\left(R^{n-1+\frac{2}{nk+1}+\varepsilon}\right),$$

where the implied O-constants depend only on $\Lambda, k, r, \boldsymbol{p}_1, \ldots, \boldsymbol{p}_r$.

Proof The proof is by induction on r. For $r = 1$, we can apply Lemma 4 to obtain

$$
L(\boldsymbol{x}; R; \boldsymbol{p}_1; \alpha) = \sum_{\substack{t \in B_R(\boldsymbol{x}) \\ t + \boldsymbol{p}_1 \in a_1^k \Lambda \\ a_1 > R^\alpha}} 1 = \sum_{\substack{a_1 > R^\alpha \\ a_1 < ((R + \|\boldsymbol{p}_1\|)/\lambda)^{1/k}}} T\left(\boldsymbol{x}; R; \frac{a_1^k}{\boldsymbol{p}_1}\right)
$$

$$
= \sum_{\substack{a_1 > R^\alpha \\ a_1 < ((R + \|\boldsymbol{p}_1\|)/\lambda)^{1/k}}} \left(\frac{R^n v_n}{a_1^{nk}} + O\left(\frac{R^{n-1}}{a_1^{(n-1)k}}\right) + O(1)\right)
$$

$$
= O\left(R^{n - \alpha(nk-1)}\right) + O\left(R^{n-1} \log R\right) + O\left(R^{1/k}\right),
$$

where there is no middle term when $n = 1$ and the logarithm in the middle term is only needed in the case $n = 2$, $k = 1$, when the corresponding harmonic series $\sum 1/a_1^{(n-1)k}$ diverges. In all other cases, these series converge, and the middle term can be taken as $O(R^{n-1})$. Thus the lemma holds for $r = 1$. Assume now that the assertion holds for some $r \geq 1$. Let β be a positive real parameter to be fixed later. Writing $a = a_1 \cdots a_{r+1}$, for symmetry reasons one has

$$
L(\boldsymbol{x}; R; \boldsymbol{p}_1, \ldots, \boldsymbol{p}_{r+1}; \alpha) = O\left(\sum_{\substack{t \in B_R(\boldsymbol{x}) \\ t + \boldsymbol{p}_1 \in a_1^k \Lambda \\ \cdots \\ t + \boldsymbol{p}_{r+1} \in a_{r+1}^k \Lambda \\ a > R^\alpha \\ \frac{a}{a_1}, \ldots, \frac{a}{a_{r+1}} \leq x^\beta}} 1\right) + O\left(\sum_{\substack{t \in B_R(\boldsymbol{x}) \\ t + \boldsymbol{p}_1 \in a_1^k \Lambda \\ \cdots \\ t + \boldsymbol{p}_{r+1} \in a_{r+1}^k \Lambda \\ a > R^\alpha \\ a_1 \cdots a_r > x^\beta}} 1\right) = L_1 + L_2,
$$

say. Employing Lemmas 4 and 5, one obtains

$$
L_1 = O\left(\sum_{\substack{t \in B_R(\boldsymbol{x}) \\ t + \boldsymbol{p}_1 \in a_1^k \Lambda \\ \cdots \\ t + \boldsymbol{p}_{r+1} \in a_{r+1}^k \Lambda \\ R^\alpha < a \leq R^{\beta(r+1)/r}}} 1\right) = O\left(\sum_{R^\alpha < a \leq R^{\beta(r+1)/r}} T\left(\boldsymbol{x}; R; \frac{a_1^k, \ldots, a_{r+1}^k}{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_{r+1}}\right)\right)
$$

$$
= O\left(\sum_{R^\alpha < a \leq R^{\beta(r+1)/r}} \left(R^n v_n \frac{E\left(\frac{a_1^k, \ldots, a_{r+1}^k}{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_{r+1}}\right)}{[a_1^k, \ldots, a_{r+1}^k]^n} + R^{n-1} + 1\right)\right)
$$

$$
= O\left(R^n \sum_{a > R^\alpha} \frac{1}{(a_1 \cdots a_{r+1})^{nk}}\right) + O\left(R^{n-1+\beta(r+1)/r + \varepsilon}\right)
$$

$$
= O\left(R^{n - \alpha(nk-1) + \varepsilon}\right) + O\left(R^{n-1+\beta(r+1)/r + \varepsilon}\right).
$$

With τ denoting the ordinary divisor function, one further obtains

$$L_2 = O\Big(\sum_{\substack{t \in B_R(x) \\ t+p_1 \in a_1^k \Lambda \\ \dots \\ t+p_r \in a_r^k \Lambda \\ a_1 \cdots a_r > x^\beta}} \sum_{t+p_{r+1} \in a_{r+1}^k \Lambda} 1\Big) = O\Big(\sum_{\substack{t \in B_R(x) \\ t+p_1 \in a_1^k \Lambda \\ \dots \\ t+p_r \in a_r^k \Lambda \\ a_1 \cdots a_r > x^\beta}} \tau\big(\|t+p_{r+1}\|/\lambda\big)\Big)$$

$$= O\Big(\sum_{\substack{t \in B_R(x) \\ t+p_1 \in a_1^k \Lambda \\ \dots \\ t+p_r \in a_r^k \Lambda \\ a_1 \cdots a_r > x^\beta}} R^\varepsilon\Big) = O(R^\varepsilon L(x; R; p_1, \dots, p_r; \beta))$$

$$= O\big(R^{n-\beta(nk-1)+2\varepsilon}\big) + O\big(R^{n-1+\frac{2}{nk+1}+2\varepsilon}\big),$$

by assumption. Setting $\beta := \frac{r}{rnk+1}$, we obtain

$$L(x; R; p_1, \dots, p_{r+1}; \alpha) = O\big(R^{n-\alpha(nk-1)+\varepsilon}\big) + O\big(R^{n-1+\frac{2}{nk+1}+2\varepsilon}\big),$$

which proves the lemma. □

We are now in a position to improve the error term of Lemma 1.

Theorem 9 *Let $\mathcal{P}$ be a finite subset of Λ, $m \in \mathbb{N}$, $\boldsymbol{m} \in \Lambda$, P be a natural number coprime to m and $\boldsymbol{x} \in \mathbb{R}^n$. Then*

$$|L(V_P; \mathcal{P}, \emptyset) \cap (\boldsymbol{m} + m\Lambda) \cap B_R(\boldsymbol{x})|$$

is

$$\frac{R^n v_n}{m^n} \prod_{p|P} \Big(1 - \frac{|\mathcal{P}/p^k \Lambda|}{p^{nk}}\Big) + O\big(R^{n-1+\frac{2}{nk+1}+\varepsilon}\big),$$

where the O-constant depends only on Λ, k and $\mathcal{P}$.

Proof This follows from the following modification of the proof of Lemma 1. We shall also use the notation from that proof. It suffices to show that (8) is

$$\frac{R^n v_n}{m^n Q^{nk}} \prod_{p|P/Q} \Big(1 - \frac{r}{p^{nk}}\Big) + O\big(R^{n-1+\frac{2}{nk+1}+\varepsilon}\big).$$

To this end, divide (8) as $C_1 + C_2$, where

$$C_1 = \sum_{d_1|P/Q} \sum_{\substack{d_2|P/Q \\ d_1 \cdots d_r \le R^{\frac{1}{nk}}}} \cdots \sum_{d_r|P/Q} \mu(d_1 \cdots d_r) \sum_{\substack{t \in \Lambda \cap B_R(x) \\ t \in q+m Q^k \Lambda \\ t \in -p_i+d_i^k \Lambda}} 1,$$

and C_2 consists of the terms with $d_1 \cdots d_r > R^{\frac{1}{nk}}$. By Lemma 6,

$$C_2 = O\left(R^{n-\frac{nk-1}{nk}+\varepsilon}\right) + O\left(R^{n-1+\frac{2}{nk+1}+\varepsilon}\right) = O\left(R^{n-1+\frac{2}{nk+1}+\varepsilon}\right).$$

One further obtains

$$C_1 = \sum_{d_1|P/Q} \sum_{d_2|P/Q} \cdots \sum_{\substack{d_r|P/Q \\ d_1 \cdots d_r \leq R^{\frac{1}{nk}}}} \mu(d_1 \cdots d_r) \sum_{\substack{t \in \Lambda \cap B_R(x) \\ t \in q+mQ^k\Lambda \\ t \in -p_i+d_i^k\Lambda}} 1$$

$$= \sum_{d_1|P/Q} \sum_{d_2|P/Q} \cdots \sum_{\substack{d_r|P/Q \\ d_1 \cdots d_r \leq R^{\frac{1}{nk}}}} \mu(d_1 \cdots d_r) T\left(x; R; \begin{matrix} d_1^k, \ldots, d_r^k, mQ^k \\ p_1, \ldots, p_r, q \end{matrix}\right)$$

Since the $d_1, \ldots, d_r$ are pairwise coprime and since $(mQ, d) = 1$, Lemma 4 in conjunction with Lemma 3 shows that

$$T\left(x; R; \begin{matrix} d_1^k, \ldots, d_r^k, mQ^k \\ p_1, \ldots, p_r, q \end{matrix}\right) = \frac{R^n v_n}{m^n (dQ)^{nk}} + O\left(\frac{R^{n-1}}{m^{n-1}(dQ)^{(n-1)k}}\right) + O(1)$$

$$= \frac{R^n v_n}{m^n (dQ)^{nk}} + O(R^{n-1}).$$

Just as in the proof of Lemma 1, substituting this in the above expression for C_1 and removing the condition $d \leq R^{\frac{1}{nk}}$ from the sum over $R^n v_n/(m^n (dQ)^{nk})$ gives the main term

$$\frac{R^n v_n}{m^n Q^{nk}} \prod_{p|P/Q} \left(1 - \frac{r}{p^{nk}}\right)$$

The error from the extra terms included in the extended multiple sum is

$$O\left(R^n \sum_{d>R^{\frac{1}{nk}}} \frac{\tau_r(d)}{d^{nk}}\right) = O(R^{n-1+\frac{1}{nk}+\varepsilon}) = O\left(R^{n-1+\frac{2}{nk+1}+\varepsilon}\right).$$

Similarly, the sum over the error term can be seen to be $O(R^{n-1+\frac{1}{nk}+\varepsilon})$. Altogether, this proves the assertion. □

One can now employ Theorem 9 instead of Lemma 1 to see that the error terms in Corollary 1 and Theorem 1 of the form $O(R^n/(\log R)^{nk-1})$ can indeed be improved to

$$O(R^{n-1+\frac{2}{nk+1}+\varepsilon}).$$

More precisely, Theorem 9 allows one to choose P as large as the product of primes less than $R^{1/(nk+1)}$ (instead of $\log R$) in the modified proofs.

Appendix A: Facts from Number Theory

We have used a number of standard facts from number theory in this paper, which we collect here with proper references for convenience.

The *inclusion–exclusion principle* says that if we have a set of N elements and a list of properties $P_1, P_2, \ldots$, with N_i elements having property P_i, N_{ij} having both properties P_i and P_j, and so on, then the number of elements having none of the properties is [11, Theorem 260]

$$N - N_1 - N_2 - \cdots + N_{12} + \cdots - N_{123} - \cdots . \tag{30}$$

The *Möbius function* is defined for $m \in \mathbb{N}$ by

$$\mu(m) := \begin{cases} 1 & \text{when } m = 1, \\ (-1)^r & \text{when } m \text{ is the product of } r \text{ distinct primes,} \\ 0 & \text{otherwise,} \end{cases}$$

and has the property that, for any $m \in \mathbb{N}$ [11, Theorem 263], [12, Theorem 6.3.1],

$$\sum_{d|m} \mu(d) = \begin{cases} 1 \text{ if } m = 1, \\ 0 \text{ otherwise,} \end{cases} \tag{31}$$

(the basis of the Möbius inversion formula), derived by applying the inclusion–exclusion principle to the singleton set $\{m\}$ with the property P_i being divisibility by the ith prime.

The *Riemann ζ-function* is defined for $\Re(s) > 1$ by

$$\zeta(s) := \sum_{m=1}^{\infty} \frac{1}{m^s} = \prod_{p \text{ prime}} \left(1 - \frac{1}{p^s}\right)^{-1} \tag{32}$$

and, as a result of Möbius inversion [11, Theorem 287], [12, Sect. 6.14],

$$\sum_{m=1}^{\infty} \frac{\mu(m)}{m^s} = \frac{1}{\zeta(s)}.$$

We needed to approximate partial Euler products, slightly more general than that on the right of (32). Let $r \geq 0$ and $s > 1$ be fixed. Then

$$0 \geq \log \prod_{p \geq N} \left(1 - \frac{r}{p^s}\right) = \sum_{p \geq N} \log \left(1 - \frac{r}{p^s}\right)$$

$$\geq \sum_{p \geq N} \frac{\frac{r}{p^s}}{\frac{r}{p^s} - 1} \quad \text{if} N^s > r$$

$$\geq \frac{r}{\frac{r}{N^s} - 1} \sum_{p \geq N} \frac{1}{p^s}$$

$$\geq \frac{r}{\frac{r}{N^s} - 1} \frac{1}{(N-1)^{s-1}}.$$

Hence, on exponentiating,

$$\prod_{p \geq N} \left(1 - \frac{r}{p^s}\right) = 1 - O(N^{1-s}). \tag{33}$$

We also had to estimate the product of primes up to a given bound [11, Theorem 415]:

$$\prod_{p \leq N} p < 4^N. \tag{34}$$

The *r-divisor function* $\tau_r(m)$, for $r \geq 2$, is the number of ways of expressing m as an ordered product of r natural numbers. The special case $r = 2$ is the ordinary divisor function

$$\tau(m) := \tau_2(m) = \sum_{d \mid m} 1.$$

which satisfies $\tau(m) = O(m^\varepsilon)$ for every $\varepsilon > 0$ [11, Theorem 315], [12, Theorem 6.5.2], from which we deduce that also

$$\tau_r(m) \leq \tau(m)^r = O(m^\varepsilon) \tag{35}$$

Another divisor sum estimate we have used is [11, Theorem 323]

$$\sum_{d \mid m} \frac{1}{d} = O(\log \log m). \tag{36}$$

An arithmetic function $f(m)$ (defined on natural numbers m) is called *multiplicative* if $f(m_1 m_2) = f(m_1) f(m_2)$ whenever $(m_1, m_2) = 1$. For example, the functions $\mu(m)$ and $\tau_r(m)$ are clearly multiplicative. A *Dirichlet series*, $\sum_{m=1}^{\infty} f(m)/m^s$, whose coefficients $f(m)$ are multiplicative can be expressed as an *Euler product* over the primes p,

$$\sum_{m=1}^{\infty} \frac{f(m)}{m^s} = \prod_p \left(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots\right), \qquad (37)$$

for all values of s for which the sum is absolutely convergent.

An estimate we have used frequently is

$$|B_\rho(\boldsymbol{x}) \cap \Lambda| = \frac{\rho^n v_n}{\det \Lambda} + O\left(\left(\frac{\rho^n}{\det \Lambda}\right)^{1-1/n}\right) + O(1), \qquad (38)$$

approximating the number of points of an n-dimensional lattice Λ in a large ball $B_\rho(\boldsymbol{x})$ (the last error term being required only when $\det \Lambda$ is bigger than ρ^n). This is obtained by dividing $B_\rho(\boldsymbol{x})$ into fundamental regions for Λ, each of volume $\det \Lambda$ and containing one point of Λ, with the error terms arising from fundamental regions that overlap the boundary of $B_\rho(\boldsymbol{x})$. The O-constants depend on the shape of Λ, but not on its size, and are independent of $\boldsymbol{x}$. A more precise version is given as Proposition 1 of [4].

We have also made much use of the Chinese Remainder Theorem in the form that if $\boldsymbol{m}_1, \boldsymbol{m}_2, \ldots, \boldsymbol{m}_r$ are points of a lattice Λ and $m_1, m_2, \ldots, m_r$ is a set of natural numbers that are pairwise coprime, then there is a point $\boldsymbol{t} \in \Lambda$ such that for all points $\boldsymbol{x} \in \Lambda$

$$\boldsymbol{x} \equiv \boldsymbol{m}_i (\mathrm{mod}\ m_i \Lambda) \text{ for } i = 1, \ldots, r \quad \Longleftrightarrow \quad \boldsymbol{x} \equiv \boldsymbol{t} (\ \mathrm{mod}\ m_1 \cdots m_r \Lambda). \quad (39)$$

This is given as Proposition 2 of [4] and is proved by applying Theorem 2.7.2 of [12] (or Theorem 121 of [11]) to each coordinate relative to a basis of Λ.

We needed a simple upper bound for binomial coefficients. By comparison with $\int_1^s \log x\, dx$, we have, for $s \in \mathbb{N}$,

$$\log s! \geq s \log s - s + 1$$

(a weak, one-sided version of Stirling's formula), so $s! > (s/e)^s$ and hence

$$\binom{m}{s} = \frac{m(m-1)\cdots(m-s+1)}{s!} \leq \frac{m^s}{s!} < \left(\frac{em}{s}\right)^s. \qquad (40)$$

References

1. Baake, M.: Diffraction of weighted lattice subsets. Can. Math. Bull. **45**(4), 483–498 (2002). http://arXiv/abs/math.MG/0106111

2. Baake, M., Grimm, U.: Kinematic diffraction from a mathematical viewpoint. Z. Kristallogr. **226**: 711–725 (2011). http://arXiv/abs/math-ph/1105.0095
3. Baake, M., Grimm, U.: Theory of Aperiodic Order: A Mathematical Invitation. Cambridge University Press, Cambridge (2013) (in preparation)
4. Baake, M., Moody, R.V., Pleasants, P.A.B.: Diffraction from visible lattice points and kth power free integers. Discrete Math. **221**, 3–42 (2000). http://arXiv/abs/math.MG/9906132
5. Baake, M., Lenz, D., Richard, C.: Pure point diffraction implies zero entropy for Delone sets with uniform cluster frequencies. Lett. Math. Phys. **82**, 61–77 (2007). http://arXiv/abs/math.DS/0706.1677
6. Burton, R.M., Steif, J.E.: Some 2-d symbolic dynamical systems: entropy and mixing. In: Pollicott, M., Schmidt, K. (eds.) Ergodic Theory of $\mathbb{Z}^d$-Actions. London Mathematical Society Lecture Note Series, vol. 228, pp. 297–305. Cambridge University Press, Cambridge (1996)
7. Cellarosi, F., Sinai, Y.G.: Ergodic properties of square-free numbers. J. Eur. Math. Soc. **15**, 1343–1374 (2013). http://arXiv/abs/math.DS/1112.4691
8. Cellarosi, F., Vinogradov, I.: Ergodic properties of m-free integers in number fields (submitted). http://arXiv/abs/math.DS/1304.0214
9. Denker, M., Grillenberger, C., Sigmund, K.: Ergodic Theory on Compact Spaces. Lecture Notes in Mathematics, vol. 527. Springer, Berlin (1976)
10. Einsiedler, M., Lindenstrauss, E., Ward, T.: Entropy in dynamics (in preparation)
11. Hardy, G.H., Wright, E.M.: An Introduction to the Theory of Numbers, 5th edn. Clarendon Press, Oxford (1979)
12. Hua, L.K.: Introduction to Number Theory. Springer, Berlin (1982)
13. Keller, G.: Equilibrium States in Ergodic Theory. London Mathematical Society Student Texts, vol. 42. Cambridge University Press, Cambridge (1998)
14. Lagarias, J.C., Pleasants, P.A.B.: Repetitive Delone sets and quasicrystals. Ergod. Theory Dyn. Syst. **23**, 831–867 (2003). http://arXiv/abs/math.DS/9909033
15. Mirsky, L.: Note on an asymptotic formula connected with r-free integers. Q. J. Math. **18**, 178–182 (1947)
16. Mirsky, L.: Arithmetical pattern problems relating to divisibility by rth powers. Proc. Lond. Math. Soc. **50**(2), 497–508 (1949)
17. Misiurewicz, M.: A short proof of the variational principle for a $\mathbb{Z}_+^N$-action on a compact space. Asterisque **40**, 147–157 (1975)
18. Moody, R.V.: Model sets: a survey. In: Axel, F., D énoyer, F., Gazeau, J.P. (eds.) From Quasicrystals to More Complex Systems, pp. 145–166. EDP Sciences, Les Ulis, and Springer, Berlin (2000). http://arXiv/abs/math.MG/0002020
19. Moree, P.: Approximation of singular series and automata. Manuscr. Math. **101**, 385–399 (2000)
20. Pillai, S.S.: On sets of square-free numbers. J. Indian Math. Soc. **2**, 116–118 (1936)
21. Ruelle, D.: Thermodynamic Formalism, 2nd edn. Cambridge University Press, Cambridge (2004)
22. Sarnak, P.: Three lectures on the Möbius function randomness and dynamics (Lecture 1) (2010). http://publications.ias.edu/sites/default/files/MobiusFunctionsLectures(2).pdf
23. Sing, B.: Pisot substitutions and beyond. PhD Thesis, Universität Bielefeld (2006). BieSOn: http://bieson.ub.uni-bielefeld.de/volltexte/2007/1155/
24. Tsang, K.M.: The distribution of r-tuples of square-free numbers. Mathematika **32**, 265–275 (1985)
25. Walters, P.: An Introduction to Ergodic Theory (reprint). Springer, New York (2000)