Available online at www.sciencedirect.com
ScienceDirect
Journal of Geometry and Physics 58 (2008) 506–541
JOURNAL OF GEOMETRY AND PHYSICS
www.elsevier.com/locate/jgp

Dworkin's argument revisited: Point processes, dynamics, diffraction, and correlations

Xinghua Deng, Robert V. Moody*

Department of Mathematics and Statistics, University of Victoria, Victoria, BC, V8W3P4, Canada

Received 30 July 2007; accepted 11 December 2007
Available online 27 December 2007

Abstract

The setting is an ergodic dynamical system (X, μ) whose points are themselves uniformly discrete point sets Λ in some space $\mathbb{R}^d$ and whose group action is that of translation of these point sets by the vectors of $\mathbb{R}^d$. Steven Dworkin's argument relates the diffraction of the typical point sets comprising X to the dynamical spectrum of X. In this paper we look more deeply at this relationship, particularly in the context of point processes.

We show that there is an $\mathbb{R}^d$-equivariant, isometric embedding, depending on the scattering strengths (weights) that are assigned to the points of $\Lambda \in X$, that takes the L^2-space of $\mathbb{R}^d$ under the diffraction measure into $L^2(X, \mu)$. We examine the image of this embedding and give a number of examples that show how it fails to be surjective. We show that full information on the measure μ is available from the weights and set of *all* the correlations (that is, the two-point, three-point, ..., correlations) of the typical point set $\Lambda \in X$.

We develop a formalism in the setting of random point measures that includes multi-colour point sets, and arbitrary real-valued weightings for the scattering from the different colour types of points, in the context of Palm measures and weighted versions of them. As an application we give a simple proof of a square-mean version of the Bombieri–Taylor conjecture, and from that we obtain an inequality that gives a quantitative relationship between the autocorrelation, the diffraction, and the ϵ-dual characters of typical element of X. The paper ends with a discussion of the Palm measure in the context of defining pattern frequencies.

1. Introduction

Imagine a point set representing the positions of an infinite set of scatterers in some idealized solid of possibly infinite extent. In practice such a set would be in two- or three-dimensional space, but for our purposes we shall simply assume that it lies in some Euclidean space $\mathbb{R}^d$. Suppose this point set satisfies the hard core condition that there is a positive lower bound r to the separation distance between the individual scatterers (uniform discreteness). Consider the set X of all possible configurations Λ of the scatterers. Assume that X is invariant under the translation action of $\mathbb{R}^d$ and assume also that there is a translation invariant ergodic probability measure μ on X. In [10], Steven Dworkin pointed out an important connection between the spectrum of the dynamical system $(X, \mathbb{R}^d, \mu)$ and the diffraction of the scattering sets $\Lambda \in X$.

* Corresponding author. Tel.: +1 250 370 2579.
E-mail addresses: xdeng@math.ualberta.ca (X. Deng), rmoody@uvic.ca (R.V. Moody).

0393-0440/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.geomphys.2007.12.006

Dworkin's argument, as it is called (see Cor. 1 of Theorem 3, below), has proven to be very fruitful, particularly in the case of pure point dynamical systems and pure point diffraction, where his argument for making the connection can be made rigorously effective; see for example [12,19,28,2]. Nonetheless, the precise relationship between the diffraction and dynamics is quite elusive. One of the purposes of this paper is to clarify this connection.

The diffraction of Λ, which is the Fourier transform of its autocorrelation (also called the two-point correlation), is not necessarily the same for all $\Lambda \in X$, whereas there is only one obvious measure, namely μ, on the dynamical system side with which to match it. However, the autocorrelation of Λ is the same for μ-almost all $\Lambda \in X$. In fact, as Jean-Baptiste Gouéré [11] has pointed out, using concepts from the theory of point processes, there is a canonical construction for this almost-everywhere-the-same autocorrelation through the use of the associated Palm measure $\dot{\mu}$ of μ. Under the hypotheses above, the first moment $\dot{\mu}_1$ of the Palm measure, which is a measure on the ambient space $\mathbb{R}^d$, is μ-almost surely the autocorrelation of $\Lambda \in X$. We offer another proof of this in Theorems 1 and 2 below.

Put in these terms, we can see that what underlies Dworkin's argument is a certain isometric embedding θ of the Hilbert space $L^2(\mathbb{R}^d, \widehat{\dot{\mu}_1})$ into $L^2(X, \mu)$. Both of these Hilbert spaces afford natural unitary representations of $\mathbb{R}^d$, call them U_t and T_t respectively ($t \in \mathbb{R}^d$). Representation T arises from the translation action of $\mathbb{R}^d$ on X and U is a multiplication action which we define in Eq. (18). The embedding θ intertwines the representations. However, θ is not in general surjective, and in fact it can fail to be surjective quite badly.

The fact is that the diffraction, or equivalently the autocorrelation measure of a typical point set $\Lambda \in X$, does not usually contain enough information to determine the measure μ, even qualitatively; see for example an explicit discussion of this in [30]. We will give a number of other examples which show that outside the situation of pure point diffraction, one must assume that this is the normal state of affairs. In fact, even in the pure point case, θ can fail to be surjective. However, we shall show in Theorem 5 that, pure point or not, the knowledge of *all* the correlations of Λ (two-point, three-point, etc.) is enough to determine μ. This is one of the principal results of the paper and depends very much on the assumption of uniform discreteness.

From a more realistic point of view, a material solid will be constituted from a number of different types of atoms and these will each have their own scattering strengths. We have incorporated this possibility into the paper by allowing there to be different types of points, labelled by indices $1, 2, \ldots, m$, and allowing each type (or colour, as we prefer to say) to have its own scattering weight w_i. There is an important distinction to be made here. We view the measure μ on X as depending only on the geometry of the point sets (including the colour information) and not on the scattering weights, which only come into consideration of the diffraction. Thus $(X, \mathbb{R}^d, \mu)$ is independent of the weighting scheme w, whereas the diffraction is not. The effect of this is that the diffraction is described in terms of a weighted version of the first moment of the Palm measure and the embedding $\theta^w : L^2(\mathbb{R}^d, \widehat{\dot{\mu}_1^w}) \longrightarrow L^2(X, \mu)$ depends on w (Theorem 3). This allows us to study the significant effect that weighting has on this mapping.

As we have already suggested, an interesting and revealing point of view is to consider our dynamical system $(X, \mathbb{R}^d, \mu)$ as a point process, in which case we think in terms of a random variable whose outcomes are the various point sets Λ of X. The theory of point processes is very well developed and has its own philosophy and culture. Although the theory is perfectly applicable to the situation that we are considering, this is nonetheless an unusual setting for it. More often random point processes are built around some sort of renewal or branching processes and the points sets involved do not satisfy any hard core property like the one we are imposing. Moreover, diffraction is not a central issue. From the point of view of the theory of diffraction of tilings or Delone point sets, which are often derived in completely deterministic ways, it is not customary to think of these in terms of random point processes. But the randomness is not in the individual point sets themselves (though that is not disallowed, e.g. [14] or [1]) but rather in the manner in which we choose them from X and the way in which the measure μ of the dynamical system on X can be viewed as a probability measure. The primary building blocks of the topology on X are the cylinder sets A of point sets Λ that have a certain colour pattern in a certain finite region of space, and $\mu(A)$ is the probability that a point set Λ, randomly chosen from X, will lie in A.

The main purposes of this paper can be seen as continuing to build bridges between the study of uniformly discrete point sets (in the context of long-range order) and point processes that was started by Gouéré in [11], and to provide a formalism of sufficient generality that the diffraction of point sets and the dynamics of their hulls can be studied together.

It is standard in the theory of point processes to model the point sets Λ involved as point measures $\lambda = \sum_{x \in \Lambda} \delta_x$, so that it is the supports of the measures that correspond to the actual point sets. This turns out to be very convenient for several reasons. The most natural topology for measures, the vague topology, exactly matches the natural topology (local topology), which is used for the construction of dynamical systems in the theory of tilings and Delone point sets (Proposition 3). Ultimately, to discuss diffraction, one ends up in measures and the vague topology anyway, so having them from the outset is useful. It is easy to build in the notion of colouring and weightings into measures.

In fact there are a number of precedents for the study of diffraction in the setting of measures rather than point sets [3,2]. However, we note that the way in which weightings are used here does not allow one to simply start from arbitrary weighted point measures at the outset. The point process itself knows about colours but nothing about weights. As we have pointed out, the weighting only enters with the correlations.

The paper first lays out the basic notions of point processes, Palm measures, and moment measures, leading to the first embedding result, Proposition 9, mentioned above. We have chosen to develop this in the non-coloured version first, since this allows the essential ideas to be more transparent. The additional complications of colour and weightings are relatively easy to add in afterwards, leading to the main embedding theorem Theorem 3, which establishes a mapping $\theta^w : L^2(\mathbb{R}^d, \widehat{\mu_1^w}) \to L^2(X, \mu)$, $\dot{\mu}_1^w$ being the weighted autocorrelation. The key to proving that the knowledge of all the higher correlations is enough to completely determine the law of the process is based on a result that shows that although θ^w need not be surjective, the algebra generated by the image under θ^w of the space of rapidly decreasing functions on $\mathbb{R}^d$ is dense in $L^2(X, \mu)$ (with some restrictions on the weighting system w). This is Theorem 4. Here the uniform discreteness seems to play a crucial role.

Section 8 provides a number of examples that fit into the setting discussed here and that illustrate a variety of things that can happen. The reader may find it useful to consult this section in advance, as the paper proceeds.

As an application of our methods, we give a simple proof of a square-mean version of the Bombieri–Taylor conjecture[1] (see Theorem 6). Using this we obtain, under the assumption of finite local complexity, an inequality that gives a quantitative relationship between three fundamental notions: the autocorrelation, the diffraction, and the ϵ-dual characters of typical elements of X. The proof of this does not involve colour and depends only on the embedding theorem Proposition 9.

The paper ends with a discussion of the problem of defining pattern frequencies for elements of X, which arises because of the built-in laxness of the local topology. We will find that the Palm measure provides a solution to the problem.

2. Point sets and point processes

2.1. Point sets and measures

Start with $\mathbb{R}^d$ with its usual topology, and metric given by the Euclidean distance $|x - y|$ between points $x, y \in \mathbb{R}^d$. We let B_R and C_R denote the open ball of radius R and the open cube of edge length R about 0 in $\mathbb{R}^d$. Lebesgue measure will be indicated by ℓ.

We are interested in closed discrete point sets Λ in $\mathbb{R}^d$, but, as explained in the Introduction, we wish also to be able to deal with different types, or colours, of points. Thus we introduce $\boldsymbol{m} := \{1, \ldots, m\}$, $m = 1, 2, 3, \ldots$, with the discrete topology and take as our basic space the set $\mathbb{E} := \mathbb{R}^d \times \boldsymbol{m}$ with the product topology, so that any point $(x, i) \in \mathbb{E}$ refers to the point x of $\mathbb{R}^d$ with colour i. When $m = 1$ we simply identify $\mathbb{E}$ and $\mathbb{R}^d$.

Closures of sets in $\mathbb{R}^d$ and $\mathbb{E}$ are denoted by overline symbols. The overline also represents complex conjugation in this paper, but there is little risk of confusion.

There is the natural translation action of $\mathbb{R}^d$ on $\mathbb{E}$ given by

$$T_t : (t, (x, i)) \mapsto t + (x, i) := (t + x, i).$$

Given $\Lambda \subset \mathbb{E}$, and $B \subset \mathbb{R}^d$, we define

$$\begin{aligned} B + \Lambda &:= \bigcup_{b \in B} T_b \Lambda \subset \mathbb{E} \\ B \cap \Lambda &:= \{(x, i) \in \Lambda : x \in B\} \subset \mathbb{E} \\ \Lambda^{\downarrow} &:= \bigcup_{(x,i) \in \Lambda} \{x\} \subset \mathbb{R}^d. \end{aligned} \tag{1}$$

$\Lambda^{\downarrow}$ is called the **flattening** of Λ.

[1] For more on the history of this see Hof's discussion of it in [13].

Let $\mathcal{O} := \{(0, 1), \ldots, (0, m)\} \subset \mathbb{E}$. Then $C_R^{(m)} := C_R + \mathcal{O}$ is a 'rainbow' cube that consists of the union of the cubes (C_R, i), $i = 1, \ldots, m$. Its closure is $\overline{C_R^{(m)}}$.

Let $r > 0$. A subset $\Lambda \subset \mathbb{E}$ is said to be r-**uniformly discrete** if for all $a \in \mathbb{R}^d$,

$$\mathrm{card}((a + C_r) \cap \Lambda) \leq 1. \tag{2}$$

In particular this implies that points of distinct colours cannot coincide.[2] The family of all the r-uniformly discrete subsets of $\mathbb{E}$ will be denoted by $\mathcal{D}_r^{(m)}$.

As we have pointed out, it is not just individual discrete point sets that we wish to discuss, but rather translation invariant families of such sets which collectively can be construed as dynamical systems.

A very convenient way to deal with countable point subsets Λ of $\mathbb{E}$ and families of them is to put them into the context of measures by replacing them by pure point measures, where the atoms correspond to the points of the set(s) in question. To this end we introduce the following objects on any locally compact space S:

- $\mathcal{S}$, the set of all Borel subsets of S;
- $\mathcal{B}(S)$, the set of all relatively compact Borel subsets of S;
- $BM_c(S)$, the space of all bounded measurable $\mathbb{C}$-valued functions of compact support on S;
- $C_c(S)$, the continuous $\mathbb{C}$-valued functions with compact support on S. If S is known to be compact, we can write $C(S)$ instead.

Following Karr [17] we let M denote the set of all **positive Radon measures** on $\mathbb{E}$, that is all positive regular Borel measures λ on $\mathbb{E}$ for which $\lambda(A) < \infty$ for all $A \in \mathcal{B}(\mathbb{E})$. Equivalently, we may view these measures as linear functionals on the space $C_c(\mathbb{E})$. We give M the vague topology. This is the topology for which a sequence $\{\lambda_n\} \in M$ converges to $\lambda \in M$ if and only if $\{\lambda_n(f)\} \to \lambda(f)$ for all $f \in C_c(\mathbb{E})$. This topology has a number of useful characterizations, some of which we give below.

Within M we have the subset M_p of **point measures** λ, those for which $\lambda(A) \in \mathbb{N}$ for all $A \in \mathcal{B}$. (Here $\mathbb{N}$ is the set of natural numbers, $\{0, 1, 2, \ldots\}$.) These measures are always pure point measures in the sense that they are countable (possibly finite) sums of delta measures:

$$\lambda = \sum a_x \delta_x, \qquad x \in \mathbb{E}, a_x \in \mathbb{N}.$$

Within M_p we also have the set M_s of **simple point measures** λ, those satisfying $\lambda(\{x\}) \in \{0, 1\}$, which are thus of the form

$$\lambda = \sum_{x \in \Lambda} \delta_x$$

where the support Λ is a countable subset of $\mathbb{E}$. Evidently for these measures, for $x \in \mathbb{E}$,

$$\lambda(\{x\}) > 0 \Leftrightarrow \lambda(\{x\}) = 1 \Leftrightarrow x \in \Lambda.$$

The Radon condition prevents the support of a point measure from having accumulation points in $\mathbb{E}$. Thus, the correspondence $\lambda \longleftrightarrow \Lambda$ provides a bijection between M_s and the closed discrete point sets of $\mathbb{E}$, i.e. the discrete point sets with no accumulation points. This is the connection between point sets and measures that we wish to use.[3] We note that the translation action of $\mathbb{R}^d$ on $\mathbb{E}$ produces an action of $\mathbb{R}^d$ on functions by $T_t f(x) = f(T_{-t}x)$, and on the spaces M, M_p, M_s of measures by $(T_t(\lambda))(A) = \lambda(-t + A)$, $(T_t\lambda)(f) = \lambda(T_{-t}(f))$ for all $A \in \mathcal{B}(\mathbb{E})$, and for all measurable functions f on $\mathbb{E}$.

Here are some useful characterizations of the vague topology and some of its properties. These are cited in [17], Appendix A, and appear with proofs in [6], Appendix A2.

[2] It is more customary to define r-uniformly discreteness by using balls rather than cubes. This makes no intrinsic difference to the concept. However, in this paper, we find that the use of cubes makes certain ideas more transparent and some of our constructions less awkward.

[3] Note that the point sets that we are considering here are simple in the sense that the multiplicity of each point in the set is just 1. However, it is not precluded that the same point in $\mathbb{R}^d$ may occur more than once in such a point set, though necessarily it would have to occur with different colours. Very soon, however, we shall also preclude this.

Proposition 1 (*The Vague Topology*).

(i) *For* $\{\lambda_n\}, \lambda \in M$ *the following are equivalent:*
  (a) $\{\lambda_n(f)\} \to \lambda(f)$ *for all* $f \in C_c(\mathbb{E})$ *(definition of vague convergence).*
  (b) $\{\lambda_n(f)\} \to \lambda(f)$ *for all* $f \in BM_c(\mathbb{E})$ *for which the set of points of discontinuity of* f *has* λ*-measure 0.*
  (c) $\{\lambda_n(A)\} \to \lambda(A)$ *for all* $A \in \mathcal{B}(\mathbb{E})$ *for which* λ *vanishes on the boundary of* A, *i.e.* $\lambda(\partial A) = 0$.

(ii) *In the vague topology,* M *is a complete separable metric space and* M_p *is a closed subspace.*

(iii) *A subspace* L *of* M *is relatively compact in the vague topology if and only if for all* $A \in \mathcal{B}(\mathbb{E})$, $\{\lambda(A) : \lambda \in L\}$ *is bounded, which again happens if and only if for all* $f \in C_c(\mathbb{E})$, $\{\lambda(f) : \lambda \in L\}$ *is bounded.*

Note that M_s is not a closed subspace of M_p: a sequence of measures in M_s can converge to point measure with multiplicities.

Proposition 2 (*The Borel Sets of* M). *The following* σ*-algebras are equal:*

(i) *The* σ*-algebra* $\mathcal{M}$ *of Borel sets of* M *under the vague topology.*
(ii) *The* σ*-algebra generated by requiring that all the mappings* $\lambda \mapsto \lambda(f)$, $f \in C_c(\mathbb{E})$ *are measurable.*
(iii) *The* σ*-algebra generated by requiring that all the mappings* $\lambda \mapsto \lambda(A)$, $A \in \mathcal{B}(\mathbb{E})$ *are measurable.*
(iv) *The* σ*-algebra generated by requiring that all the mappings* $\lambda \mapsto \lambda(f)$, $f \in BM_c(\mathbb{E})$ *are measurable.*

A measure $\lambda \in M$ is **translation bounded** if for all bounded sets $K \in \mathcal{B}(\mathbb{E})$, $\{\lambda(a+K) : a \in \mathbb{R}^d\}$ is bounded. In fact, a measure is translation bounded if this condition holds for a single set of the form $K = K_0 \times \boldsymbol{m}$ where $K_0 \subset \mathbb{R}^d$ has a non-empty interior. For such a K and for any positive constant n, we define the space $M_p(K, n)$ of translation bounded measures $\lambda \in M_p$ for which

$$\lambda(a+K) \le n$$

for all $a \in \mathbb{R}^d$. Evidently $M_p(K, n)$ is closed if K is open, and by Proposition 1 it is relatively compact, and hence compact. See also [2], where this is proved in a more general setting.

If $r > 0$ then $M_p(C_r^{(m)}, 1)$ is the set of point measures λ whose support Λ satisfies the uniform discreteness condition (2). In particular, $M_p(C_r^{(m)}, 1) \subset M_s$ and is compact.

If $\lambda \in M_s$ is a translation bounded measure on $\mathbb{R}^d$ we shall often write expressions like $\sum_{x \in B} \lambda(\{x\})$ where B is some uncountable set (like $\mathbb{R}^d$ itself). Such sums only have a countable number of terms and so sum to a non-negative integer if B is bounded, or possibly to $+\infty$ otherwise.

2.2. Point processes

By definition, a **point process** on $\mathbb{E}$ is a measurable mapping

$$\xi : (\Omega, \mathcal{A}, P) \longrightarrow (M_p, \mathcal{M}_p)$$

from some probability space into M_p with its σ-algebra of Borel sets $\mathcal{M} \cap M_p$. That is, it is a random point measure. Sometimes, when $m > 1$, it is called a **multivariate** point process. The **law** of the point process is the probability measure which is the image $\mu := \xi(P)$ of P. The point process is **stationary** if μ is invariant under the translation action of $\mathbb{R}^d$ on M_p.

Thus from the stationary point process ξ we arrive at a measure-theoretical dynamical system $(M_p, \mathbb{R}^d, \mu)$. Conversely, any such system may be interpreted as a stationary point process (by choosing $(\Omega, \mathcal{A}, P)$ to be $(M_p, \mathbb{R}^d, \mu)$).

There is no indication in the definition what the support of the law μ of the process might look like. In most cases of interest, this will be something, or be inside something, considerably smaller. In the sequel we shall assume that we have a point process $\xi : (\Omega, \mathcal{A}, P) \longrightarrow (M_p, \mathcal{M}_p)$ that satisfies the following conditions:

(PPI) the support of the measure $\mu = \xi(P)$ is a closed subset X of $M_p(C_r^{(m)}, 1)$ for some $r > 0$.
(PPII) μ is stationary and has positive intensity (see below for the definition).
(PPIII) μ is ergodic.

These are examples of what are called translation bounded measure dynamical systems in [2], although it should be noted that there the space of measures is not restricted to point measures, or even positive measures.

Obviously under (PPI) and (PPII), X is compact, and $(X, \mathbb{R}^d, \mu)$ is both a measure-theoretic and a topological dynamical system.

Condition (PPI) implies that the point process is simple and so we may identify the measures of the point process as the actual (uniformly discrete) point sets in $\mathbb{E}$ that are their supports. Write $\ddot{X}$ for the subset of $\mathcal{D}_r^{(m)}$ given by the supports of the measures of X. We call a point process satisfying (PPI) (and (PPII)) a **uniformly discrete (stationary) point process**. We will make considerable use of these two ways of looking at a point process — either as being formed of point measures or of uniformly discrete point sets.

The ergodic hypothesis eventually becomes indispensable, but for our initial results it is not required. Usually we simplify the terminology and speak of a point process ξ and assume implicitly the accompanying notation $(X, \mathbb{R}^d, \mu)$ and so on. We denote the family of all Borel subsets of X by $\mathcal{X}$.

A key point is that the vague topology on the space X of a uniformly discrete point process is precisely the topology most commonly used in the study of point set dynamical systems [26]. Sometimes this is called the **local topology** since it implies a notion of closeness that depends on the local configuration of points (as opposed to other topologies that depend only on the long-range average structure of the point set).

The local topology is most easily described as the uniform topology on $\mathcal{D}_r^{(m)}$ generated by the entourages

$$U(C_R, \epsilon) := \{(\Lambda, \Lambda') \in \mathcal{D}_r^{(m)} : C_R \cap \Lambda \subset C_\epsilon + \Lambda', \quad C_R \cap \Lambda' \subset C_\epsilon + \Lambda\}, \tag{3}$$

where R, ϵ vary over the positive real numbers.

Note that in (3), Λ and Λ' are subsets of $\mathbb{E}$ and we are using the conventions of (1). Intuitively two sets are close if on large cubes their points can be paired, taking colour into account, so that they are all within ϵ-cubes of each other. It is easy to see that $\mathcal{D}_r^{(m)}$ is closed in this topology.

Given any $\Lambda' \in \mathcal{D}_r^{(m)}$ we define the open set

$$U(C_R, \epsilon)[\Lambda'] := \{\Lambda \in \mathcal{D}_r^{(m)} : (\Lambda, \Lambda') \in U(C_R, \epsilon)\}.$$

Proposition 3 (*See also [2]*). *Let ξ be a uniformly discrete point process. Then under the correspondence $\lambda \leftrightarrow \Lambda$ between measures in X and the point sets in $\ddot{X}$, the vague and local topologies are the same.*

Proof. Let r be the constant of the uniform discreteness. Let $\{\lambda_n\}$ be a sequence of elements of X for which the corresponding sequence $\{\Lambda_n\} \subset \ddot{X}$ converges in the local topology to some point set $\Lambda \in \mathcal{D}_r^{(m)}$. Choose any positive function $f \in C_c(\mathbb{E})$ and suppose that its support is in C_R, and choose any $\epsilon > 0$. Let $N_0 := 1 + \sup_{n\in\mathbb{N}} \lambda_n(C_R)$ and find $\eta > 0$ so that $\eta < r$ and for all $x, y \in C_R$,

$$|x - y| < \eta \Longrightarrow |f(x) - f(y)| < \epsilon/N_0.$$

Let $\{x_1, \ldots, x_N\} = C_R \cap \Lambda \subset \mathbb{E}$. Then for all large n, $C_R \cap \Lambda_n \subset \{C_\eta + x_1, \ldots C_\eta + x_N\}$ with exactly one point in each of these cubes. Then

$$|\lambda_n(f) - \lambda(f)| = \left| \sum_{y\in C_R\cap\Lambda_n} f(y) - \sum_{x\in C_R\cap\Lambda} f(x) \right| \le N_0\epsilon/N_0 = \epsilon.$$

Thus $\{\lambda_n(f)\} \to \lambda(f)$, and since $f \in C_c(\mathbb{E})$ was arbitrary, $\{\lambda_n\} \to \lambda$ in X.

Now, going the other way, suppose that $\{\lambda_n\} \to \lambda$ in X. Let $R > 0$ and let $C_R \cap \Lambda = \{x_1, \ldots, x_N\}$. Choose any $0 < \epsilon < r$ small enough that for all $i \le N$, $C_\epsilon + x_i \subset C_R^{(m)}$, and let

$$f_\epsilon := \sum_{i=1}^{N} \mathbf{1}_{C_\epsilon + x_i}.$$

Then $\{\lambda_n(f_\epsilon)\} \to \lambda(f_\epsilon) = N = \lambda(C_R) \leftarrow \{\lambda_n(C_R)\}$, so for all $n \gg 0$, $\lambda_n(f_\epsilon) = N = \lambda_n(C_R)$ (see Proposition 1). Since each cube $C_\epsilon + x_i$ can contain at most one point of any element of $\mathcal{D}_r^{(m)}$, then for all $n \gg 0$, and for all $i \le N$, there is a $y_i^{(n)} \in (C_\epsilon + x_i) \cap \Lambda_n$. This accounts for all the points of $C_R \cap \Lambda_n$. Thus $\Lambda_n \in U(C_R, \epsilon)[\Lambda]$. This proves that $\{\Lambda_n\} \to \Lambda$. $\square$

Remark 1. Let ξ be a uniformly discrete point process. By Proposition 3, the two topological spaces X and $\ddot{X}$ are homeomorphic. In particular, $\ddot{X}$ is compact in the local topology (a fact that can be seen directly from its definition). The σ-algebras of their Borel sets $\mathcal{X}$ and $\ddot{\mathcal{X}}$ are isomorphic and we obtain a measure $\mu_{\ddot{X}}$ on $\ddot{X}$. Geometrically it is often easier to work in $\ddot{X}$ rather than X, and we will frequently avail ourselves of the two different points of view. Notationally it is convenient to use the same symbols X and μ for both and to use upper and lower case symbols to denote elements from X according to whether we are treating them as sets or measures.

3. The moments and counting functions

In this section we work in the one-colour case $m = 1$. Thus $\mathbb{E} = \mathbb{R}^d$. We let $\xi : (\Omega, \mathcal{A}, P) \longrightarrow (X, \mathcal{X})$ be a uniformly discrete stationary point process on $\mathbb{E}$ with law μ. We assume that $X \subset M_p(C_r, 1) \subset M_s$.

According to Proposition 2, for each $A \in \mathcal{B}(\mathbb{R}^d)$ and for each $f \in BM_c(\mathbb{R}^d)$, the mappings

$$N_A : M_p(C_r, 1) \longrightarrow \mathbb{Z}, \qquad N_A(\lambda) = \lambda(A) \tag{4}$$

$$N_f : M_p(C_r, 1) \longrightarrow \mathbb{C}, \qquad N_f(\lambda) = \lambda(f) \tag{5}$$

are measurable functions on $M_p(C_r, 1)$, and by restriction, measurable functions on X. The first of these simply counts the number of points of the support of λ that lie in the set A, and N_f is its natural extension from sets to functions. Hence the name **counting functions**. They may also be considered as functions on M_p. They may also be viewed as functions on the space X viewed as the space of corresponding point sets.

Thus, for example, in this notation we have, for all $f \in BM_c(\mathbb{R}^d)$,

$$\begin{aligned} \int_X \lambda(f) \mathrm{d}\mu(\lambda) &= \int_X N_f(\lambda) \mathrm{d}\mu(\lambda) = \int_X \sum_{x \in \mathbb{R}^d} \lambda(\{x\}) f(x) \mathrm{d}\mu(\lambda) \\ &= \int_X N_f(\Lambda) \mathrm{d}\mu(\Lambda) = \int_X \sum_{x \in \Lambda} f(x) \mathrm{d}\mu(\Lambda). \end{aligned} \tag{6}$$

This is the first moment of the measure μ, henceforth denoted as μ_1. More generally, the nth **moments**, $n = 1, 2, \ldots$, of a finite positive measure ω on X are the unique measures on $(\mathbb{R}^d)^n$ defined by

$$\begin{aligned} \omega_n(A_1 \times \cdots \times A_n) &= \int_X \lambda(A_1) \ldots \lambda(A_n) \mathrm{d}\omega(\lambda) \\ &= \int_X N_{A_1} \ldots N_{A_n} \mathrm{d}\omega, \end{aligned}$$

where $A_1, \ldots A_n$ run through all $\mathcal{B}(\mathbb{R}^d)$. Alternatively, for all $f_1, \ldots, f_n \in BM_c(\mathbb{R}^d)$,

$$\omega_n((f_1, \ldots, f_n)) = \int_X N_{f_1} \ldots N_{f_n} \mathrm{d}\omega.$$

Since ω is a finite measure and the values of $\lambda(f) = N_f(\lambda)$ are uniformly bounded for any $f \in BM_c(\mathbb{R}^d)$ as λ runs over X, these expressions define translation bounded measures on $(\mathbb{R}^d)^n$.

If the measure ω is stationary (invariant under the translation action of $\mathbb{R}^d$) then the nth moment of ω is invariant under the action of simultaneous translation of all n variables. Thus, if the point process ξ is also stationary then the first moment of the law of ξ is invariant, and hence a multiple of the Lebesgue measure:

$$\mu_1(A) = \int_X \lambda(A) \mathrm{d}\mu(\lambda) = I\ell(A). \tag{7}$$

This non-negative constant I, which is finite because of our assumption of uniform discreteness, is the expectation for the number of points per unit volume of λ in A and is called the **intensity** of the point process. We shall always assume (see PPII) that the intensity is positive, i.e. not zero.

The meaning of N_f can be extended well beyond $BM_c(\mathbb{R}^d)$. To make this extension we introduce the usual L^p-spaces $L^p(\mathbb{R}^d, \ell)$, $L^p(X, \mu)$ together with their norms which we shall indicate by $\|\cdot\|_p$ in either case. In fact, we need these only for $p = 1, 2$. We shall also make use of the sup-norms $\|\cdot\|_\infty$.

Proposition 4. *The mapping (4) uniquely defines a continuous mapping (also called N)*

$$N : L^1(\mathbb{R}^d, \ell) \longrightarrow L^1(X, \mu)$$
$$f \mapsto N_f$$

satisfying $\|N_f\|_1 \le \sqrt{2} I \|f\|_1$. Moreover, for all $f \in L^1(\mathbb{R}^d, \ell)$,

$$N_f(\lambda) = \lambda(f) \quad \text{for } \mu \text{ almost surely all } \lambda \in X.$$

Proof. Let $A \subset \mathbb{R}^d$ be a bounded and measurable set, let $\mathbf{1}_A$ be the characteristic function of A on $\mathbb{R}^d$, and define $N_{\mathbf{1}_A}$ on X by $N_{\mathbf{1}_A}(\lambda) = \lambda(\mathbf{1}_A) = \lambda(A) = N_A(\lambda)$; see (4). From (7), $\|N_{\mathbf{1}_A}\|_1 = \int_X N_A(\lambda)\mathrm{d}\mu(\lambda) = I\ell(A) = I\|\mathbf{1}_A\|_1$. This shows that the result holds for N defined on these basic functions.

For simple functions of the form $f = \sum_{k=1}^n c_k \mathbf{1}_{A_k}$, where the sets $A_k \subset \mathcal{B}(\mathbb{R}^d)$ are mutually disjoint and the $c_k \ge 0$, define

$$N_f = \sum_{k=1}^n c_k N_{\mathbf{1}_{A_k}} = \sum_{k=1}^n c_k N_{A_k}.$$

Then

$$\|N_f\|_1 = \sum_{k=1}^n c_k \|N_{\mathbf{1}_{A_k}}\|_1 = \sum_{k=1}^n c_k I\ell(A_k) = I\|f\|_1,$$

and $N_f(\lambda) = \lambda(f)$ for all $\lambda \in X$.

The extension, first to arbitrary positive measurable functions and then to arbitrary real-valued functions f, goes in the usual measure-theoretical way, and need not be reproduced here.

Finally we use the linearity to go to complex-valued integrable f. If $f = f_r + \sqrt{-1} f_i$ is the splitting of f into real and imaginary parts, then $N_f = N_{f_r} + \sqrt{-1} N_{f_i}$, so

$$\|N_f\|_1 \le I(\|f_r\|_1 + \|f_i\|_1) = I \int_{\mathbb{R}^d} (|f_r| + |f_i|)\mathrm{d}\ell.$$

Using the inequality $(|f_r| + |f_i|)^2 \le 2(|f_r|^2 + |f_i|^2)$, we have

$$\|N_f\|_1 \le \sqrt{2} I \int_{\mathbb{R}^d} \sqrt{|f_r|^2 + |f_i|^2}\mathrm{d}\ell = \sqrt{2} I \|f\|_1.$$

It is clear that if f and g differ on sets of measure 0 then likewise so do N_f and N_g, so this establishes the existence of the mapping. □

Proposition 5. *Let f_n, $n = 1, 2, 3, \ldots$, and f be measurable $\mathbb{C}$-valued functions on $\mathbb{R}^d$ with supports all contained within a fixed compact set K. Suppose that $\|f_n\|_\infty, \|f\|_\infty < M$ for some $M > 0$ and $\{f_n\} \to f$ in the L^1-norm on $\mathbb{R}^d$. Then $\{N_{f_n}\} \to N_f$ in the L^2-norm on X.*

Proof. Because of the uniform discreteness, $\lambda(K)$ is uniformly bounded on X by a constant $C(K) > 0$. Then for $g = f$ or $g = f_n$ for some n, $|N_g(\lambda)| < MC(K)$.

$$\begin{aligned}
\|N_f - N_{f_n}\|_2^2 &= \int_X |N_f(\lambda) - N_{f_n}(\lambda)|^2 \mathrm{d}\mu(\lambda) \\
&\le \int_X (|N_f(\lambda)| + |N_{f_n}(\lambda)|)|N_f(\lambda) - N_{f_n}(\lambda)|\mathrm{d}\mu(\lambda) \\
&\le 2MC(K) \int_X |N_f(\lambda) - N_{f_n}(\lambda)|\mathrm{d}\mu(\lambda),
\end{aligned}$$

which, by Proposition 4, tends to 0 as $n \to \infty$. □

4. Averages, the Palm measure and autocorrelation: one-colour case

In this section we work in the one-colour case $m = 1$. Thus $\mathbb{E} = \mathbb{R}^d$. We let $\xi : (\Omega, \mathcal{A}, P) \longrightarrow (X, \mathcal{X})$ be a uniformly discrete stationary point process on $\mathbb{E}$ with law μ.

4.1. The Palm measure

The **Campbell measure** is the measure c' on $\mathbb{R}^d \times X$, defined by

$$c'(B \times D) = \int_D \lambda(B)\mathrm{d}\mu(\lambda) = \int_X \sum_{x\in B} \lambda(\{x\})\mathbf{1}_D(\lambda)\mathrm{d}\mu(\lambda) \tag{8}$$

for all $B \times D \in \mathcal{E} \times \mathcal{X}$.

We note that c' is invariant with respect to simultaneous translation of its two variables. By introducing the measurable mapping

$$\phi : \mathbb{R}^d \times X \longrightarrow \mathbb{R}^d \times X : \quad (x, \lambda) \mapsto (x, T_{-x}\lambda)$$

we obtain a twisted version c of c', also defined on $\mathbb{R}^d \times X$:

$$\begin{aligned} c(B \times D) &= \int_X \sum_{x\in B} \lambda(\{x\})\mathbf{1}_D(T_{-x}\lambda)\mathrm{d}\mu(\lambda) \\ &= \int_X \sum_{x\in B\cap\Lambda} (\mathbf{1}_D)(-x + \Lambda)\mathrm{d}\mu(\Lambda), \end{aligned}$$

and this is invariant under translation of the first variable:

$$\begin{aligned} c((t + B) \times D) &= \int_X \sum_{x\in(t+B)} \lambda(\{x\})\mathbf{1}_D(T_{-x}\lambda)\mathrm{d}\mu(\lambda) \\ &= \int_X \sum_{y\in B} T_{-t}\lambda(\{y\})\mathbf{1}_D(T_{-y}T_{-t}\lambda)\mathrm{d}\mu(T_{-t}\lambda) \\ &= c(B \times D), \end{aligned}$$

using the translation invariance of μ.

Hence for D fixed, c is a multiple $\dot{\mu}(D)\ell(B)$ of the Lebesgue measure and $D \mapsto \dot{\mu}(D) := c(B \times D)/\ell(B)$ is a non-negative measure on X that is independent of the choice of $B \in \mathcal{B}(\mathbb{R}^d)$ (assuming that B has positive measure). This measure is called the *Palm measure* of the point process. See [6] for more details.

$$\begin{aligned} \dot{\mu}(D) &= \frac{1}{\ell(B)} \int_X \sum_{x\in B} \lambda(\{x\})\mathbf{1}_D(T_{-x}\lambda)\mathrm{d}\mu(\lambda) \\ &= \frac{1}{\ell(B)} \int_X \sum_{x\in B\cap\Lambda} \mathbf{1}_D(-x + \Lambda)\mathrm{d}\mu(\Lambda). \end{aligned} \tag{9}$$

We note that $\dot{\mu}(X) = \int_X \lambda(B)\mathrm{d}\mu(\lambda)/\ell(B) = I$, which is the intensity of the point process. Some authors normalize the Palm measure by the intensity in order to render it a probability measure, and then call this probability measure the Palm measure. We shall not do this. However, we note that the normalized Palm measure is often viewed as being the conditional probability

$$\frac{1}{I}\dot{\mu}(D) = \mu(\{\lambda \in D\} \mid \lambda(\{0\}) = 1\}),$$

that is, the probability conditioned by the assumption that 0 is in the support of the point measures that we are considering. In fact the conditional probability defined in this way is meaningless in general since the probability that $\lambda(\{0\}) \neq 0$ is usually 0. But the intuition of what is desired is contained in the definition. Taking B as an arbitrarily small neighbourhood of 0, we see that in effect we are only looking at points of λ very close to 0 and then translating λ so that 0 is in the support. The result is averaged over the volume of B.

If the point process falls into the subspace X of $\mathcal{M}$ then the support of the Palm measure is also in X. However, the Palm measure is not stationary in general, since the translation invariance of μ has, in effect, been taken out.

Its first moment, sometimes called the intensity of the Palm measure, is

$$\dot{\mu}_1 : \dot{\mu}_1(A) = \int_X \lambda(A)\mathrm{d}\dot{\mu}(\lambda) \quad \text{or equivalently} \tag{10}$$
$$\dot{\mu}_1(f) = \int_X \lambda(f)\mathrm{d}\dot{\mu}(\lambda) = \int_X N_f(\lambda)\mathrm{d}\dot{\mu}(\lambda).$$

The first moment of the Palm measure, and also the higher moments to be defined later, play a crucial role in the development of the paper, since they are, in an almost sure sense, the two-point and higher point correlations of the elements of X.

As with μ, we will, consider the Palm measure interchangeably as a measure on X or on $\ddot{X}$ (as we have already done implicitly in Eq. (9)).

The importance of the Palm measure is its relation to the average value of a function over a typical point set $\Lambda \in X$, and from there to pattern frequencies in Λ and its direct involvement in the autocorrelation of Λ. To explain this we need to develop the Palm theory a little further.

Lemma 1 (*Campbell Formula*). *For any measurable function $F : \mathbb{R}^d \times X \longrightarrow \mathbb{R}$,*

$$\int_{\mathbb{R}^d}\int_X F(x,\lambda)\mathrm{d}\dot{\mu}(\lambda)\mathrm{d}x = \int_X \sum_{x\in\mathbb{R}^d} \lambda(\{x\})F(x, T_{-x}\lambda)\mathrm{d}\mu(\lambda).$$

Proof. This can be proven easily by checking it on simple functions. Let $F = \mathbf{1}_B \times \mathbf{1}_D$. Then

$$\begin{aligned}\int_{\mathbb{R}^d}\int_X \mathbf{1}_B(x) \times \mathbf{1}_D(\lambda)\mathrm{d}\dot{\mu}(\lambda)\mathrm{d}x &= \ell(B)\dot{\mu}(D) = c(B\times D)\\ &= \int_X \sum_{x\in\mathbb{R}^d} \lambda(\{x\})\mathbf{1}_B(x)\mathbf{1}_D(T_{-x}\lambda)\mathrm{d}\mu(\lambda)\\ &= \int_X \sum_{x\in\mathbb{R}^d} \lambda(\{x\})F(x, T_{-x}\lambda)\mathrm{d}\mu(\lambda). \quad \square\end{aligned}$$

Let ν_R be the function on X defined by

$$\nu_R(\lambda) = \frac{1}{\ell(C_R)}N_{C_R}(\lambda),$$

for all $R > 0$. We treat ν_R as the Radon–Nikodym density of an absolutely continuous measure on X (with respect to μ).

Proposition 6. *In vague convergence,*

$$\{\nu_R\} \to \dot{\mu} \quad \textit{as} \quad R \to 0.$$

Proof. Use the definition of the Palm measure in (9) with B replaced by C_R. Then for any continuous function G on X,

$$\dot{\mu}(G) = \frac{1}{\ell(C_R)}\int_X \sum_{y\in C_R} \lambda(\{y\})G(T_{-y}\lambda)\mathrm{d}\mu(\lambda).$$

If we require that $R < r$ then

$$\sum_{y\in C_R} \lambda(\{y\})G(T_{-y}\lambda) = N_{C_R}(\lambda)G(T_{-x}\lambda),$$

where x is the unique point in $\Lambda \cap C_R$ when it is not empty, and then

$$\dot{\mu}(G) = \frac{1}{\ell(C_R)} \int_X N_{C_R}(\lambda) G(T_{-x}\lambda) \mathrm{d}\mu(\lambda).$$

On the other hand

$$\nu_R(G) = \frac{1}{\ell(C_R)} \int_X N_{C_R}(\lambda) G(\lambda) \mathrm{d}\mu(\lambda).$$

Thus

$$\begin{aligned} |\dot{\mu}(G) - \nu_R(G)| &= \left| \frac{1}{\ell(C_R)} \int_X N_{C_R}(\lambda) \{G(T_{-x}\lambda) - G(\lambda)\} \mathrm{d}\mu(\lambda) \right| \\ &\leq \frac{1}{\ell(C_R)} \int_X N_{C_R}(\lambda) |\{G(T_{-x}\lambda) - G(\lambda)\}| \mathrm{d}\mu(\lambda). \end{aligned} \tag{11}$$

The rest follows from the uniform continuity of G (X is compact). From the inequality (11),

$$|\dot{\mu}(G) - \nu_R(G)| \leq \frac{\epsilon_R}{\ell(C_R)} \int_X N_{C_R}(\Lambda) \mathrm{d}\mu(\Lambda) = \epsilon_R I,$$

for some $\epsilon_R \to 0$ as $R \to 0$, where I is the intensity of the point process.

Therefore, we have that $\nu_R \to \dot{\mu}$ vaguely. □

4.2. Averages

Let ξ be a uniformly discrete ergodic stationary point process, with corresponding dynamical system $(X, \mathbb{R}^d, \mu)$. Let $F \in C(X)$. The **average** of F at $\lambda \in X$ is

$$\mathrm{Av}(F)(\lambda) = \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x \in C_R} \lambda(\{x\}) F(T_{-x}\lambda),$$

if it exists. Thus $\mathrm{Av}(F)$ is a function defined at certain points of X. Alternatively, we may think of F as a function on point sets and write this as

$$\mathrm{Av}(F)(\Lambda) = \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x \in \Lambda \cap C_R} F(-x + \Lambda).$$

We will prove the almost sure existence of averages.

Proposition 7. *Let $F \in C(X)$. The average value of* $\mathrm{Av}(F)(\lambda)$ *of F exists μ-almost surely for $\lambda \in X$ and it is almost surely equal to $\dot{\mu}(F)$. In particular* $\mathrm{Av}(F)$ *exists as a measurable function on X. If μ is uniquely ergodic then the average value always exists everywhere and is equal to $\dot{\mu}(F)$.*

Proof. It is clear that the average value is constant along the orbit of any point λ for which it exists.

Let $\epsilon > 0$. Since F is uniformly continuous, there is a compact set K and an $s > 0$ so that $|F(\lambda') - F(\lambda'')| < \epsilon$ whenever $(\lambda', \lambda'') \in U(K, s)$. In particular $|F(-x + \lambda) - F(-u + \lambda)| < \epsilon$ whenever $|x - u| < s$. We can assume that $s < r$.

Let $\nu_s := \frac{1}{\ell(C_s)} N_{C_s} : X \longrightarrow \mathbb{C}$, as above. For $x \in \mathbb{R}^d$ and $\lambda \in X$, $N_{C_s}(T_{-x}\lambda) = 1$ if and only if $x \in u + C_s$ for some $u \in \Lambda$. Thus

$$\frac{1}{\ell(C_R)} \int_{C_R} F(T_{-x}\lambda) \nu_s(T_{-x}\lambda) \mathrm{d}x \tag{12}$$

$$\sim \frac{1}{\ell(C_R)} \sum_{u \in C_R} \lambda(\{u\}) \frac{1}{\ell(C_s)} \int_{u + C_s} F(T_{-x}\lambda) \mathrm{d}x, \tag{13}$$

where the $\sim$ comes from boundary effects only and becomes equality in the limit.

There is a constant $a > 0$ so that $\mathrm{card}(\lambda(C_R))/\ell(C_R) < a$, independent of R or which $\lambda \in X$ is taken. Using this and our choice of s, we obtain

$$\left| \lim_{R\to\infty} \frac{1}{\ell(C_R)} \int_{C_R} F(T_{-x}\lambda)\nu_s(T_{-x}\lambda)\mathrm{d}x - \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{u\in C_R} \lambda(\{u\})F(T_{-u}\lambda) \right| < a\epsilon. \tag{14}$$

The right-hand term is just the average value of F at λ, if the limit exists. However, by the Birkhoff ergodic theorem the left integral exists almost surely and is equal to $\int_X F\nu_s \mathrm{d}\mu = \nu_s(F)$, ν_s being treated as a measure.

Now making $\epsilon \to 0$, so $s \to 0$ also, and using Proposition 6 we have

$$\dot{\mu}(F) = \lim_{s\to 0} \nu_s(F) = \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{u\in C_R} \lambda(\{u\})F(T_{-u}\lambda) = \mathrm{Av}(F)(\lambda).$$

Thus the average value of F on λ exists almost surely.

In the uniquely ergodic case, the conclusion of Birkhoff's theorem is true everywhere in X. □

4.3. *The autocorrelation and the Palm measure*

Again, let $\xi : (\Omega, \mathcal{A}, P) \longrightarrow (X, \mathcal{X})$ be a uniformly discrete stationary ergodic point process on $\mathbb{R}^d$ with law μ. For each $\lambda \in X$ we define $\tilde{\lambda}$ to be the point measure on $\mathbb{R}^d$ defined by $\tilde{\lambda}(\{x\}) = \overline{\lambda(\{-x\})}$ (though at this point we are only dealing with real measures). Then the autocorrelation of λ is defined as

$$\gamma_\lambda := \lim_{R\to\infty} \frac{1}{\ell(C_R)} \left(\lambda|_{C_R} * \widetilde{\lambda}|_{C_R}\right) = \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x,y\in\Lambda\cap C_R} \delta_{y-x}.$$

Here the limit, which may or may not exist, is taken in the vague topology.

A simple consequence of the van Hove property of cubes is

$$\gamma_\lambda = \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x\in\Lambda\cap C_R, y\in\Lambda} \delta_{y-x}. \tag{15}$$

Namely, for any $f \in C_c(\mathbb{R}^d)$, say with support K, and any $x \in C_R$, $f(y-x) = 0$ unless $y \in C_R + K$, and thus for large R the only relevant y which are not in C_R are in the K-boundary of C_R, which is vanishingly small in relative volume as $R \to \infty$.

Theorem 1. *The first moment $\dot{\mu}_1$ of the Palm measure is a positive, positive definite, translation bounded measure. Furthermore, μ-almost surely, $\lambda \in X$ admits an autocorrelation γ_λ and it is equal to $\dot{\mu}_1$. If X is uniquely ergodic then $\dot{\mu}_1 = \gamma_\lambda$ for all $\lambda \in X$.*

Proof. We begin with the statement about the autocorrelation measures γ_λ. Let $f \in C_c(\mathbb{R}^d)$. The autocorrelation of λ at f, if it exists, is

$$\begin{aligned}\gamma_\lambda(f) &= \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x\in C_R, y\in\Lambda} \lambda(\{x\})f(y-x) \\ &= \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x\in C_R} \lambda(\{x\})N_f(T_{-x}\lambda) \\ &= \dot{\mu}(N_f) = \dot{\mu}_1(f)\end{aligned} \tag{16}$$

for $\lambda \in X$, μ-almost surely, where we have used Proposition 7 and (10).

This is basically what we want, but we must show that it holds for *all* $f \in C_c(\mathbb{R}^d, \mathbb{R})$ for almost all $\lambda \in X$. This is accomplished by using a countably dense (in the sup-norm) set of elements of $C_c(\mathbb{R}^d, \mathbb{R})$. We can get (16) simultaneously for this countable set, and this is enough to get it for all $f \in C_c(\mathbb{R}^d, \mathbb{R})$. Then γ_λ exists and is equal to $\dot{\mu}_1$ for almost all $\lambda \in X$. For more details see [11].

Finally, it is clear that γ_λ is a positive and positive definite measure whenever it exists, and hence also $\dot{\mu}_1$ is positive and positive definite. All positive and positive definite measures are translation bounded; [4] Prop. 4.4., or [12]. □

4.4. Diffraction and the embedding theorem

It is a consequence of the positivity and positive definiteness of the autocorrelation that it is Fourier transformable and that its Fourier transform is likewise positive, positive definite, and translation bounded [4].

We recall that the Fourier transform of such a measure ω on $\mathbb{R}^d$ can be defined by the formula

$$\widehat{\omega}(f) = \omega(\widehat{f}) \tag{17}$$

for all f in the space $\mathbb{S}$ of rapidly decreasing functions of $\mathbb{R}^d$. In fact, it will suffice to have this formula on the space $\mathbb{S}_c$ of compactly supported functions in $\mathbb{S}$, since they are dense in $\mathbb{S}$ in the standard topology on $\mathbb{S}$ ([29]). The key point is that if $\{f_n\} \in \mathbb{S}_c$ converges to $f \in \mathbb{S}$, then $\{\widehat{f_n}\}$ converges to $\widehat{f}$ and one can use the translation boundedness of ω to see then that $\{\omega(\widehat{f_n})\}$ converges to $\{\omega(\widehat{f})\}$, i.e. $\widehat{\omega}(f)$ is known from the values of $\{\widehat{\omega}(f_n)\}$.

The measure $\widehat{\gamma_\lambda}$ is the **diffraction** of λ, when it exists. Our results show that the first moment of the Palm measure, $\dot{\mu}_1$, must also be a positive, positive definite transformable translation bounded measure and that almost surely $\widehat{\dot{\mu}_1}$ is the diffraction of $\lambda \in X$.

The next result appears, in a slightly different form, in [11]. For complex-valued functions h on $\mathbb{E}$ define $\tilde{h}$ by $\tilde{h}(x) = \overline{h(-x)}$. We denote the standard inner product defined by $\|\cdot\|_2$ on $L^2(X, \mu)$ by $(\cdot, \cdot)$.

Proposition 8. *Let $g, h \in BM_c(\mathbb{R}^d)$ and suppose that $g * \tilde{h} * \dot{\mu}_1$ is a continuous function on $\mathbb{R}^d$. Then for all $t \in \mathbb{R}^d$,*

$$g * \tilde{h} * \dot{\mu}_1(-t) = (T_t N_g, N_h).$$

Proof. It suffices to prove the result when g, h are real-valued functions. By Proposition 2, N_g, N_h are measurable functions on X, and they are clearly L^1-functions (Proposition 4).

$$\begin{aligned}
g * \tilde{h} * \dot{\mu}_1(-t) &= \int_{\mathbb{R}^d} (g * \tilde{h})(-t-u)\mathrm{d}\dot{\mu}_1(u) = \int_{\mathbb{R}^d} (\tilde{g} * h)(t+u)\mathrm{d}\dot{\mu}_1(u) \\
&= \int_{\mathbb{R}^d} \widetilde{T_t g} * h(u)\mathrm{d}\dot{\mu}_1(u) \\
&= \int_X \left(\sum_{x \in \mathbb{R}^d} \lambda(\{x\})(\widetilde{T_t g} * h)(x) \right) \mathrm{d}\dot{\mu}(\lambda) \\
&= \int_{\mathbb{R}^d} \int_X \sum_{x \in \mathbb{R}^d} \lambda(\{x\})(T_t g)(u) h(x+u)\mathrm{d}\dot{\mu}(\lambda)\mathrm{d}u \\
&= \int_{\mathbb{R}^d} \int_X (T_t g)(u) T_{-u} N_h(\lambda)\mathrm{d}\dot{\mu}(\lambda)\mathrm{d}u
\end{aligned}$$

where we have used (6) and the dominated convergence theorem to rearrange the sum and the integral. Now using the Campbell formula we may continue:

$$\begin{aligned}
\dot{\mu}_1 * g * \tilde{h}(-t) &= \int_X \sum_{u \in \mathbb{R}^d} \lambda(\{u\})(T_t g)(u) N_h(\lambda)\mathrm{d}\mu(\lambda) \\
&= \int_X N_{T_t g}(\lambda) N_h(\lambda)\mathrm{d}\mu = (T_t N_g, N_h). \quad \square
\end{aligned}$$

We are now at the point where we can prove the embedding theorem (in the unweighted case). This involves the two Hilbert spaces $L^2(\mathbb{R}^d, \widehat{\dot{\mu}_1})$ and $L^2(X, \mu)$. Since the translation action of $\mathbb{R}^d$ on X is measure preserving, it gives rise to a unitary representation T of $\mathbb{R}^d$ on $L^2(X, \mu)$ by the usual translation action of $\mathbb{R}^d$ on measures.

We also have a unitary representation U of $\mathbb{R}^d$ on $L^2(\mathbb{R}^d, \widehat{\dot{\mu}_1})$ defined by

$$U_t f(x) = \mathrm{e}^{-2\pi \mathrm{i} t.x} f(x) = \chi_{-t}(x) f(x), \tag{18}$$

where the characters χ_k are defined by

$$\chi_k(x) = \mathrm{e}^{2\pi \mathrm{i} k.x}. \tag{19}$$

We denote the inner product of $L^2(\mathbb{R}^d, \widehat{\mu_1})$ by $\langle\cdot,\cdot\rangle$ and note that with respect to it U is a unitary representation of $\mathbb{R}^d$.

Proposition 9. *If $g, h \in \mathbb{S}$ are rapidly decreasing functions then*

$$g * \tilde{h} * \dot{\mu}_1(-t) = \langle U_t(\hat{g}), \hat{h}\rangle.$$

In particular,

$$\langle U_t(\hat{g}), \hat{h}\rangle = (T_t N_g, N_h).$$

Thus there is an isometric embedding intertwining U and T,

$$\theta : L^2(\mathbb{R}^d, \widehat{\mu_1}) \longrightarrow L^2(X, \mu),$$

under which

$$\hat{f} \mapsto N_f$$

for all $f \in \mathbb{S}$.

Proof. As we have pointed out, it will suffice to show the first result for $g, h \in \mathbb{S}_c$ since it is dense in $\mathbb{S}$ under the standard topology of $\mathbb{S}$. We note that the hypotheses of Proposition 8 are satisfied, so, starting as in its proof and denoting the inverse Fourier transform by $f \mapsto \check{f}$, we have

$$g * \tilde{h} * \dot{\mu}_1(-t) = \int_{\mathbb{E}} \widetilde{T_t g} * h(u)\mathrm{d}\dot{\mu}_1(u) = \int_{\mathbb{E}} (\widetilde{T_t g})^{\vee} h^{\vee} \mathrm{d}\widehat{\mu_1}.$$

The first result follows from $\check{h} = \check{\bar{h}} = \overline{\hat{h}}$ and $(\widetilde{T_t g})^{\vee} = \widehat{T_t g} = \chi_{-t}\hat{g}$.

The second part of the proposition follows from Proposition 8 and the observation that $\mathbb{S}_c$ is dense in $C_c(\mathbb{R}^d)$ in the sup-norm ([29], Thm. 1), and hence certainly in the $\|\cdot\|_2$-norm, and $C_c(\mathbb{R}^d)$ is dense in $L^2(\mathbb{R}^d, \widehat{\mu_1})$ in the $\|\cdot\|_2$-norm (see [27], Appendix E).

Thus we have the existence of the embedding on a dense subset of $L^2(\mathbb{R}^d, \widehat{\mu_1})$ and it extends uniquely to the closure. □

5. Adding colour

We now look at the changes required to Section 4 in order to include colour, i.e. to have $m > 1$. The colour enters in two ways. First of all, the dynamics, that is to say the dynamical hull X and the measure μ, depend on colour since closeness in the local topology depends on simultaneous closeness of points of like colours. Secondly the autocorrelation, and then the diffraction, depends on colour.

Diffraction depends on how scattering waves from different points (atoms) superimpose upon each other. However, physically, different types of atoms will have different scattering strengths, and so we wish to incorporate this into the formalism. This is accomplished by specifying a vector w of weights to be associated with the different colours and introducing for each point measure λ of our hull X a weighted version of it, λ^w. This will be a measure on $\mathbb{R}^d$. It will be important that the weighting is kept totally separate from the topology and geometry of X. The geometry of the configuration and the weighting of points, which enters into the diffraction, are different things. The measures describing our point sets are measures on $\mathbb{E}$, but the diffraction always takes place on the flattened point sets.

On the geometrical side we have treated the full colour situation from the start. In this section we introduce it into the autocorrelation/diffraction side. This affects almost every result in Section 4. However, we shall see that every proof then generalizes quite easily, and we simply outline the new situation and the generalized results, leaving the reader to do the easy modifications to the proofs.

5.1. Weighting systems

Let $\xi : (\Omega, \mathcal{A}, P) \longrightarrow (X, \mathcal{X})$ be a uniformly discrete stationary multi-variate point process, where $X \subset M_p$, $\mathcal{X} = \mathcal{M} \cap X$, and $(X, \mathbb{R}^d, \mu)$ is the resulting dynamical system. We let $\mathbb{E} = \mathbb{R}^d \times \boldsymbol{m}$, with $\mathbb{E}^i = \mathbb{R}^d \times \{i\}$ and $\mathbb{E} = \bigcup_{i\le m} \mathbb{E}^i$. For each i we have the restriction

$$\mathrm{res}^i : \lambda \mapsto \lambda^i$$

of measures on $\mathbb{E}$ to measures on $\mathbb{E}^i$. We will simply treat these restricted measures as being measures on $\mathbb{R}^d$. If $\lambda \leftrightarrow \Lambda$ then we also think of res^i as the mapping $\Lambda \mapsto \Lambda^i := \{x \in \mathbb{R}^d : (x, i) \in \Lambda\}$.[4]

The same argument that led to (7) gives

$$\int_X \lambda^i(A)\mathrm{d}\mu(\lambda) = I^{(i)}\ell(A) \tag{20}$$

for some $I^{(i)} \geq 0$, for each i. We shall always assume:
(PPIIw) $I^{(i)} > 0$ for all $i \leq m$.

A **system of weights** is a vector $w = (w_1, \ldots, w_m)$ of real numbers.[5] We define a mapping

$$X \to M_s(\mathbb{R}^d) \quad \lambda \mapsto \lambda^w := \sum_{i \leq m} w_i \lambda^i.$$

The quantity

$$I^w := \sum_{i=1}^m w_i I^{(i)} \tag{21}$$

is called the **weighted intensity** of the weighted point process.

We also have the flattening map:

$$X \to M_s(\mathbb{R}^d) \quad \lambda \mapsto \lambda^{\downarrow} := \sum_{i \leq m} \lambda^i.$$

First introduce the measure c^w on $\mathbb{R}^d \times X$:

$$c^w(B \times D) = \int_X \sum_{x \in B} \lambda^w(\{x\}) T_x \mathbf{1}_D(\lambda)\mathrm{d}\mu(\lambda).$$

Since $(T_x\lambda)^w = T_x(\lambda^w)$ this measure is invariant under translation of the first variable and we have

$$c^w(B \times D) = \ell(B)\dot{\mu}^w(D).$$

This determines the w-**weighted Palm measure** $\dot{\mu}^w$ on X. This is not a Palm measure in the normal sense of the word. However, it plays the same role as the Palm measure in much of what follows. For example, there is a corresponding Campbell formula:

$$\int_{\mathbb{R}^d} \int_X F(x, \lambda)\mathrm{d}\dot{\mu}^w(\lambda)\mathrm{d}x = \int_X \sum_{x \in \mathbb{R}^d} \lambda^w(\{x\}) F(x, T_{-x}\lambda)\mathrm{d}\mu(\lambda)$$

for all measurable $F : \mathbb{R}^d \times X \longrightarrow \mathbb{C}$.

We note the formula for the weighted intensity:

$$\begin{aligned} I^w l(A) &= \int_X \lambda^w \mathrm{d}\mu(\lambda) = \int_X \sum_{x \in A} \lambda^w(\{x\})\mathrm{d}\mu(\lambda) \\ &= c^w(A \times X) = l(A)\dot{\mu}^w(X), \quad \text{whence} \end{aligned}$$

$$I^w = \dot{\mu}^w(X). \tag{22}$$

For all $i \leq m$, for all $A \in \mathcal{B}(\mathbb{R}^d)$, and for all $f \in BM_c(\mathbb{R}^d)$ define

$$N_A^w : X \longrightarrow \mathbb{N} \quad N_A^w(\lambda) = \lambda^w(A) = \sum_{x \in A} \lambda^w(\{x\}) \tag{23}$$

$$N_f^w : X \longrightarrow \mathbb{N} \quad N_f^w(\lambda) = \lambda^w(f) = \sum_{x \in \mathbb{R}^d} \lambda^w(\{x\}) f(x).$$

4 It is also possible to define associated dynamical systems X^i and with them Palm measures. However, it is important here that everything will always refer back to the full colour situation encoded in the geometry of X.

5 One could have complex numbers here, but it makes things easier, and more natural for higher correlations, if the weights are real.

Thus, for example,

$$N_A^w(\lambda) = \sum w_i \lambda^i(A) = \sum w_i N_A(\mathrm{res}^i(\lambda)) = \sum w_i N_A \circ \mathrm{res}^i(\lambda). \tag{24}$$

Define

$$\nu_R^w : X \longrightarrow \mathbb{R}, \quad \nu_R^w(\lambda) = \frac{1}{\ell(C_R)} N_{C_R}^w(\lambda)$$

or equivalently, $\nu_R^w(\Lambda) = \frac{1}{\ell(C_R)} N_{C_R}^w(\Lambda)$. In vague convergence,

$$\{\nu_R^w\} \to \dot{\mu}^w \quad \text{as } R \to 0.$$

These auxiliary measures are used, as before, to prove the existence of averages. Let $F \in C(X)$. The w-**average value** of F **on** X is

$$\mathrm{Av}^w(F)(\lambda) = \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x\in C_R} \lambda^w(\{x\}) T_x F(\lambda)$$

$$\mathrm{Av}^w(F)(\Lambda) = \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x\in C_R} \lambda^w(\{x\}) F(-x+\Lambda),$$

if it exists.

Proposition 7 becomes:

Proposition 10. *The w-average value of $F \in C(X)$ is defined at $\Lambda \in X$, μ-almost surely and is almost surely equal to $\dot{\mu}^w(F)$. If μ is uniquely ergodic then the average value always exists and is equal to $\dot{\mu}^w(F)$.*

We now come to the w-weighted autocorrelation. This is the measure on $\mathbb{R}^d$ defined by

$$\begin{aligned}
\gamma_\lambda^w(f) &= \lim_{R\to\infty} \frac{1}{\ell(C_R)} \lambda^w|_{C_R} * \widetilde{\lambda^w}|_{C_R}(f) \\
&= \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x\in C_R, y\in\mathbb{R}^d} \lambda^w(\{x\}) \overline{\lambda^w(\{y\})} f(y-x) \\
&= \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x\in C_R} \lambda^w(\{x\}) N_f^w(T_{-x}\lambda) \\
&= \mathrm{Av}^w(N_f^w) = \dot{\mu}^w(N_f^w) =: \dot{\mu}_1^w(f)
\end{aligned} \tag{25}$$

for all $f \in C_c(\mathbb{R}^d)$ and for μ-almost all $\lambda \in X$.

We call $\dot{\mu}_1^w$ the **weighted first moment of the weighted Palm measure**.

Theorem 2. *The weighted first moment $\dot{\mu}_1^w$ of the weighted Palm measure is a positive definite measure. It is Fourier transformable and its Fourier transform $\widehat{\dot{\mu}_1^w}$ is a positive translation bounded measure on $\mathbb{R}^d$. Furthermore, μ-almost surely, $\lambda \in X$ admits a w-weighted autocorrelation γ_λ^w and it is equal to $\dot{\mu}_1^w$. If X is uniquely ergodic then $\dot{\mu}_1^w = \gamma_\lambda^w$ for all $\lambda \in X$.* □

Remark 2. Regarding the statements about the transformability and translation boundedness of the Fourier transform, this is a consequence of the positive definiteness of the Palm measure; see [4] Thm. 4.7, Prop. 4.9.

Proposition 8 has the weighted form: Let $g, h \in BM_c(\mathbb{R}^d)$ and suppose that $g * \tilde{h} * \dot{\mu}_1^w$ is a continuous function on $\mathbb{R}^d$. Then for all $t \in \mathbb{R}^d$,

$$g * \tilde{h} * \dot{\mu}_1^w(-t) = (T_t N_g^w, N_h^w). \tag{26}$$

Our interest now shifts to $L^2(\mathbb{R}^d, \widehat{\dot{\mu}_1^w})$, its inner product $\langle\cdot,\cdot\rangle^w$, and the unitary representation U^w of $\mathbb{R}^d$ on it which is given by the same formula as (18).

5.2. The embedding theorem

From Eq. (26) we obtain our embedding theorem, which is the full colour version of Proposition 9.

Theorem 3. *For each system of weights $w = (w_1, \ldots, w_m)$, the mapping*

$$\hat{f} \mapsto N_f^w, \tag{27}$$

defined for all $f \in \mathbb{S}$, extends uniquely to an isometric embedding

$$\theta^w : L^2(\mathbb{R}^d, \widehat{\mu_1^w}) \longrightarrow L^2(X, \mu)$$

which intertwines the representations U and T.

We note here that the space on the left-hand side depends on w while the space on the right-hand side does not. The question of the image of θ^w is then an interesting one. We come to this later.

We also note that the formula for $\theta^w(f)$ in (27), though true for $f \in \mathbb{S}$, and no doubt many other functions too, is not true in general, and in particular not true for some functions that we will need to consider in the discussion of spectral properties, e.g. see Corollary 3.

Theorem 3 gives an isometric embedding of $L^2(\mathbb{R}^d, \widehat{\mu_1^w})$ into $L^2(X, \mu)$ and along with it a correspondence of the spectral components of $L^2(\mathbb{R}^d, \widehat{\mu_1^w})$ and its image in $L^2(X, \mu)$. Now the point is that the spectral information of $L^2(\mathbb{R}^d, \widehat{\mu_1^w})$ can be read directly off that of the measure $\widehat{\mu_1^w}$. Specifically, let $\widehat{\mu_1^w} = (\widehat{\mu_1^w})_{pp} + (\widehat{\mu_1^w})_{sc} + (\widehat{\mu_1^w})_{ac}$ be the decomposition of $\widehat{\mu_1^w}$ into its pure point, singular continuous, and absolutely continuous parts. For $f \in L^2(\mathbb{R}^d, \widehat{\mu_1^w})$, the associated spectral measure σ_f^w on $\mathbb{R}^d$ is given by

$$\langle f, U_t f\rangle^w = \int \mathrm{e}^{2\pi \mathrm{i} x.t} \mathrm{d}\sigma_f^w(x).$$

However,

$$\langle f, U_t f\rangle^w = \int \mathrm{e}^{2\pi \mathrm{i} x.t} f(x)\overline{f(x)} \mathrm{d}\widehat{\mu_1^w}(x),$$

so we have

$$\sigma_f^w = |f|^2 \widehat{\mu_1^w} = |f|^2 (\widehat{\mu_1^w})_{pp} + |f|^2 (\widehat{\mu_1^w})_{sc} + |f|^2 (\widehat{\mu_1^w})_{ac}, \tag{28}$$

which is the spectral decomposition of the measure σ_f. With $\Box$ standing for pp, sc, or ac, we have

$$\begin{aligned} L^2(\mathbb{R}^d, \widehat{\mu_1^w})_{\Box} &:= \{f \in L^2(\mathbb{R}^d, \widehat{\mu_1^w}) : \sigma_f^w \text{ is of type } \Box\} \\ &= \{f \in L^2(\mathbb{R}^d, \widehat{\mu_1^w}) : \operatorname{supp}(f) \subset \operatorname{supp}((\widehat{\mu_1^w})_{\Box})\}. \end{aligned}$$

This explains how information about the spectrum of the diffraction can be inferred from the nature of the dynamical spectrum and vice versa. Since the mapping θ depends on w and is not always surjective, the correspondence between the two has to be treated with care. Some examples of what can happen are given in Section 8.

Combining (28) with Theorem 3, we have S. Dworkin's theorem:

Corollary 1. *Let $f \in C_c(\mathbb{R}^d)$. Then for μ-almost all $\lambda \in X$, $\widehat{\gamma_{f*\lambda}^w}$ is the spectral measure $\sigma_{N_f^w}$ on $L^2(X, \mu)$.*

Proof. $\gamma_{f*\lambda}^w = f * \tilde{f} * \gamma_\lambda^w$, so $\widehat{\gamma_{f*\lambda}^w} = |\widehat{f}|^2 \widehat{\gamma_\lambda^w} = |\widehat{f}|^2 \widehat{\mu_1^w} = \sigma_{\widehat{f}}^w$ almost surely. Now,

$$\langle \widehat{f}, U_t \widehat{f}\rangle^w = (N_f^w, T_t N_f^w)_{L^2(X,\mu)}$$

so the spectral measure $\sigma_{\widehat{f}}^w$ computed for $L^2(\mathbb{R}^d, \widehat{\mu_1^w})$ is the same as the spectral measure $\sigma_{N_f^w}$ computed for $L^2(X, \mu)$. $\Box$

Corollary 2. *For all $f, g \in L^2(\mathbb{R}^d, \widehat{\mu_1^w})$, the spectral measures $(\langle U_t f, g\rangle^w)^\vee$ and $(T_t\theta^w(f), \theta^w(g))^\vee$ on $\mathbb{R}^d$ are equal, and in particular of the same spectral type: absolutely continuous, singular continuous, pure point, or a mix of these.*

Corollary 3. *For $k \in \mathbb{R}^d$, χ_k is in the point spectrum*[6] *of U_t if and only if $\widehat{\mu_1^w}(k) \neq 0$. The corresponding eigenfunction is $\mathbf{1}_{\{-k\}}$. When this holds, χ_k is in the point spectrum of T_t, the eigenfunction corresponding to it is $\theta^w(\mathbf{1}_{\{-k\}})$, and $\|\theta^w(\mathbf{1}_{\{-k\}})\| = \widehat{\mu_1^w}(k)^{1/2}$.*

Proof. The first statement is clear from (18) and our remarks above. For the second, suppose that $k \in \mathbb{R}^d$ and $\widehat{\mu_1^w}(k) \neq 0$. Let $f \in L^1(\mathbb{R}^d, \widehat{\mu_1^w})$ be an eigenfunction for k. Then

$$\exp(2\pi \mathrm{i} k.t) f(x) = U_t f(x) = \exp(-2\pi \mathrm{i} t.x) f(x)$$

for all $x \in \mathbb{R}^d$. For x with $f(x) \neq 0$, $\exp(2\pi \mathrm{i} k.t) = \exp(-2\pi \mathrm{i} x.t)$ for all $t \in \mathbb{R}^d$, so $x = -k$. Thus $f = f(-k)\mathbf{1}_{\{-k\}}$. By Theorem 3, $\theta^w(f) \in L^2(X, \mu)$ with $T_t(\theta^w(f)) = \chi_k(t)\theta^w(f)$ for all $t \in \mathbb{R}^d$. $\square$

Remark 3. One should note that the eigenvalues always occur in pairs $\pm k$ since μ_1 is positive definite and $\widehat{\mu_1^w}(-k) = \widehat{\mu_1^w}(k)$. How does one work out $\theta^w(\mathbf{1}_{\{-k\}})$? This is the content of the L^2-mean form Bombieri–Taylor conjecture that we shall establish in Section 9.

6. The algebra generated by the image of θ

6.1. The density of $\Theta^w(\mathbb{S})$

Theorem 4. *Let (X, μ) be an m-coloured stationary uniformly discrete ergodic point process and w a system of weights. Suppose that the weights w_i, $i = 1, \ldots, m$, are all different from one another and also none of them is equal to 0. Then the algebra Θ^w generated by $\theta^w(\mathbb{S})$ and the identity function 1_X is dense in $L^2(X, \mu)$.*

Remark 4. If $\widehat{\mu_1^w}(0) \neq 0$ then $\theta^w(\mathbb{S})$ already contains 1_X by Corollary 3.

The remainder of this subsection is devoted to the proof of this theorem.

We begin with the construction of certain basic types of finite partitions of X. Here we will find it easier to deal with coloured point sets than with their corresponding measures.

Let $r > 0$ be fixed so that $X \subset M_p(C_r^{(m)}, 1)$. For each pair of measurable sets $K, V \subset \mathbb{R}^d$, with K bounded and V a neighbourhood of 0, we define

$$U(K, V) := \{(\Lambda, \Lambda') \in \mathcal{D}_r : K \cap \Lambda \subset V + \Lambda' \quad \text{and} \quad K \cap \Lambda' \subset V + \Lambda\}, \tag{29}$$

which is just a variation on (3), and serves to define another fundamental system of entourages for the same uniformity, and then the same topology, on X as we have been using all along. For any $\Phi \in \mathcal{D}_r^{(m)}$ we define

$$U(K, V)[\Phi] := \{\Lambda \in X : (\Lambda, \Phi) \in U(K, V)\}.$$

We begin by choosing a finite grid in $\mathbb{R}^d$ and partitioning X according to the colour patterns that it makes in this grid. Here are the details. Let $K \subset \mathbb{R}^d$ be a half-open cube of the form $[a_1, a_1 + R) \times \cdots \times [a_d, a_d + R)$, $R > 0$, and V be a half-open cube of diameter less than r, centred on 0, which is so sized that its translates can tile K without overlaps. The set of translation vectors used to make up this tiling is denoted by Ψ, so in fact this set is the set of centres of the tiles of the tiling. Each centre locates a tile and in each of these tiles we can have at most one coloured point of Λ, that is, at most one pair (x, i) with $x \in \mathbb{R}^d$ and $i \leq m$. Let

$$\mathfrak{P} := \{\Phi = (\Phi_1, \ldots, \Phi_m) : (\Phi_0, \Phi_1, \ldots, \Phi_m) \text{ is an ordered partition of } \Psi\}; \tag{30}$$

that is, we take all possible ordered partitions of Ψ into $m + 1$ pieces, which we interpret as all the various coloured patterns of cells of our tiling. Φ_i designates the cells containing the points of colour i (second component i), $i = 1, \ldots m$, and Φ_0 designates all the cells which contain no points of the pattern.

The inclusion relation $\subset$ on $\mathfrak{P}$ given by $\Phi = (\Phi_1, \ldots, \Phi_m) \subset \Phi' = (\Phi'_1, \ldots, \Phi'_m)$ if and only if $\Phi_i \subset \Phi'_i$, for all $1 \leq i \leq m$, provides a natural partial ordering on $\mathfrak{P}$. Using the notation established in (1), for each $\Phi \in \mathfrak{P}$ define

$$P[\Phi] := \{\Lambda \in X : K \cap \Lambda \subset V + \Phi, K \cap \Lambda \not\subset V + \Phi' \quad \text{for any } \Phi' \subsetneq \Phi\}.$$

[6] One often simply says that k is in the point spectrum, with the understanding that it means χ_k.

Because of the choice of V, an element of X can have at most one point in any one of the cubes making up the tiling of K. Each $P[\Phi]$ is the set of elements of X which make the coloured pattern Φ inside the cube K.

Lemma 2.

$$X = \bigcup_{\Phi \in \mathfrak{P}} P[\Phi]$$

is a partition of X. Furthermore, for all $\Phi \in \mathfrak{P}$,

$$U(K, V^\circ)[\Phi] \cap X \subset P[\Phi] \subset U(K, \overline{V})[\Phi] \cap X.$$

Proof. By construction the $P[\Phi]$ form a partition of X. Let $\Phi \in \mathfrak{P}$ and let $\Lambda \in P[\Phi]$. Then $K \cap \Lambda \subset V + \Phi$. Also, for each s lying in some component Φ_i of Φ there is an $x \in K \cap \Lambda_i$ with $x \in V + s$, whence $s \in -V + x \subset \overline{V} + x$. This shows that $K \cap \Phi \subset \overline{V} + \Lambda$, so $\Lambda \in U(K, \overline{V})[\Phi]$.

On the other hand, if $\Lambda \in U(K, V^\circ)[\Phi]$ then $K \cap \Lambda \subset V^\circ + \Phi \subset V + \Phi$, which is the first condition for $\Lambda \subset P[\Phi]$. Since also $\Phi = K \cap \Phi \subset V^\circ + \Lambda$, for each s in some component Φ_i there is an $x \in \Lambda_i$ with $x = -v + s \in V^\circ + s \subset V + s$. By the construction of the tiling of K, no other set $V^\circ + t$, $t \in \Psi$, can contain x. Thus Λ meets every tile centred on a point of Φ and $\Lambda \in P[\Phi]$. □

We know that $\theta^w(\mathbb{S})$ contains all the functions N^w_f, $f \in \mathbb{S}$, in particular all the N^w_f, $f \in \mathbb{S}_c$, and so its L^2-closure contains N^w_f, $f \in C_c(\mathbb{R}^d)$ (use Proposition 5). Again, using Proposition 5 we can conclude that $\overline{\theta^w(\mathbb{S})}$ contains all the functions N^w_A, where A is a bounded open or closed subset of $\mathbb{R}^d$. We start with these functions and work to produce more complicated ones.

Lemma 3. *Let $s \in \Psi$ and let $i \le m$. Then the functions $N_{V^\circ+s} \circ \mathrm{res}^i(\Lambda)$ and $N_{\overline{V}+s} \circ \mathrm{res}^i(\Lambda)$ are in $\overline{\Theta^w}$.*

Proof. $N^w_{V^\circ+s} \in \overline{\Theta^w}$. From (24) and $\mathrm{diam}(V) < r$,

$$N^w_{V^\circ+s}(\Lambda) = \sum_{i=1}^{m} w_i N_{V^\circ+s} \circ \mathrm{res}^i(\Lambda) \tag{31}$$

$$= \sum_{i=1}^{m} w_i N_{V^\circ+s}(\Lambda^i) = 0 \quad \text{or} \quad w_j \tag{32}$$

according as $(V^\circ + s) \cap \Lambda$ is empty or contains a (necessarily unique) point x of some colour j. Write $\mathbf{F}$ for $N^w_{V^\circ+s}$ and F for $N_{V^\circ+s}$. The first is a function on X, the second a function on r-uniformly discrete subsets of $\mathbb{R}^d$ (see (4)). Then[7] $\mathbf{F}^j(\Lambda) = \sum_{i=1}^m w_i^j F(\Lambda_i)$ since always $F^j(\Lambda_i) = F(\Lambda_i)$ and $F(\Lambda_i)F(\Lambda_k) = 0$ whenever $i \neq k$.

Let W be the $m \times m$ matrix defined by $W_{jk} = w_k^j$, $1 \le j, k \le m$. By the hypotheses on the weights it has an inverse Y. Then

$$\sum_{j=1}^{m} Y_{ij} \mathbf{F}^j(\Lambda) = \sum_{j=1}^{m} Y_{ij} \sum_{k=1}^{m} w_k^j F(\Lambda_k) = F(\Lambda_i).$$

This proves that the functions $\Lambda \mapsto N_{V^\circ+s}(\Lambda_i) = N_{V^\circ+s} \circ \mathrm{res}_i(\Lambda)$ are all in Θ^w. The same argument applies in the case of $\overline{V}$. □

Lemma 4. *For all $\Phi \in \mathfrak{P}$, $\mathbf{1}_{P[\Phi]} \in \overline{\Theta^w}$.*

Proof. Let $\Phi \in \mathfrak{P}$ and assume $\Phi \neq \emptyset$. Let

$$f_1 := \prod_{i=1}^{m} \prod_{s \in \Phi_i} N_{V^\circ+s} \circ \mathrm{res}_i$$

$$f_2 := \prod_{i=1}^{m} \prod_{s \in \Phi_i} N_{\overline{V}+s} \circ \mathrm{res}_i,$$

[7] Here the superscripts really mean powers!

all in Θ^w. These functions take the value 1 only on sets Λ which hit all the cells $V^\circ + s$ (respectively $\overline{V} + s$) centred on the points and with the colours specified by Φ. However, such Λ may hit other cells also; hence

$$f_1 \le \sum_{\Phi \subset \Phi' \in \mathfrak{P}} \mathbf{1}_{P[\Phi']} \le f_2.$$

However, for any fixed i,

$$\int |N_{\overline{V}+s}(\Lambda_i) - N_{V^\circ+s}(\Lambda_i)|^2 \mathrm{d}\mu(\Lambda) = \int |N_{\overline{V}+s}(\Lambda_i) - N_{V^\circ+s}(\Lambda_i)| \mathrm{d}\mu(\Lambda)$$
$$= \int |N_{(\overline{V}\setminus V^\circ)+s}(\Lambda_i)| \mathrm{d}\mu(\Lambda) = I^i \ell((\overline{V} \setminus V^\circ + s)) = 0$$

showing that $N_{\overline{V}+s}$ and $N_{V^\circ+s}$ are equal as L^2 functions, whence also f_1 and f_2 are equal. This shows that

$$\sum_{\Phi' \supset \Phi} \mathbf{1}_{P[\Phi']} \in \overline{\Theta^w}.$$

In the case that Φ is empty,

$$\sum_{\Phi' \supset \Phi} \mathbf{1}_{P[\Phi']} = \sum_{\Phi' \subset \Psi} \mathbf{1}_{P[\Phi']} = \mathbf{1}_X \in \overline{\Theta^w}.$$

Now by Möbius inversion on the partially ordered on the subsets of $\mathfrak{P}$, $\mathbf{1}_{P[\Phi]} \in \overline{\Theta^w}$ for all $\Phi \in \mathfrak{P}$. □

Lemma 5. *Let $F : X \longrightarrow \mathbb{R}$ be a continuous function and let $\epsilon > 0$. Then there exist half-open cubes K, V as above so that for the corresponding partition of X,*

$$\|F - \sum_{\Phi \in \mathfrak{P}} m_\Phi \mathbf{1}_{P[\Phi]}\|_\infty \le \epsilon,$$

where $m_\Phi := \inf\{F(\Lambda) : \Lambda \subset P[\Phi]\}$.

Proof. Since X is compact, F is uniformly continuous. Then given $\epsilon > 0$ there exist a compact set $K \subset \mathbb{R}^d$ and a neighbourhood V' of $0 \in \mathbb{R}^d$ so that $|F(\Lambda) - F(\Lambda')| < \epsilon$ for all $(\Lambda, \Lambda') \in U(K, V') \cap (X \times X)$. We can increase K to some half-open cube of the type above without spoiling this and then choose some half-open cube V, centred on 0 and of diameter less than r, which tiles K and also satisfies $2\overline{V} \subset V'$. We let $\mathfrak{P}$ be the corresponding set of partitions.

Let $\Phi \in \mathfrak{P}$. If $\Lambda, \Lambda' \in U(K, \overline{V})[\Phi] \cap X$ then $K \cap \Lambda \subset \overline{V} + \Phi$ and $\Phi \subset \overline{V} + \Lambda'$. Thus for any $x \in K \cap \Lambda$, $x = v + s = v' + v + x'$ where $s \in \Phi$, $x' \in \Lambda'$ (both with the same colour as x), and $v, v' \in \overline{V}$, from which we conclude $K \cap \Lambda \subset 2\overline{V} + \Lambda'$. In the same way $K \cap \Lambda' \subset 2\overline{V} + \Lambda$, so $(\Lambda, \Lambda') \in U(K, V')$ and $|F(\Lambda) - F(\Lambda')| < \epsilon$. In particular this holds for all $\Lambda, \Lambda' \in P[\Phi]$, since it is contained in $U(K, \overline{V})[\Phi]$, and so F varies by less than ϵ on $P[\Phi]$. The result follows at once from this. □

The proof of Theorem 4 is an immediate consequence of this. $\overline{\Theta^w(\mathbb{S})}$ contains the functions $\mathbf{1}_{P[\Phi]}$ and so also all their limit points, and hence all continuous functions on X. Finally the continuous functions are dense in $L^2(X, \mu)$. □

For the case $m = 1$, recall that the nth moment of μ is the measure μ_n on $(\mathbb{R}^d)^n$, defined by $\mu_n(f_1, \ldots, f_n) = \mu(N_{f_1} \ldots N_{f_n})$. Since Theorem 4 says that the linear span of all the product functions $N_{f_1} \ldots N_{f_n}$ is dense in $L^2(X, \mu)$, we see that μ is entirely determined by its moment measures.

In the general case we may define the nth weighted moments by

$$\mu_n^w(f_1, \ldots, f_n) = \mu(N_{f_1}^w \cdots N_{f_n}^w). \tag{33}$$

For example, using Eqs. (20) and (21),

$$\mu_1^w(\mathbf{1}_A) = \mu(N_A^w) = \int_X \sum w_i \lambda^i(A) \mathrm{d}\mu(\lambda)$$
$$= \sum w_i I^{(i)} \ell(A) = I^w \ell(A), \tag{34}$$

which shows that $\mu_1^w = I^w \ell$ (also see Eq. (22)).

Then the same argument as in Theorem 4 leads to:

Proposition 11. *Let (X, μ) be an m-coloured stationary uniformly discrete ergodic point process and w a system of weights in which w_i, $i = 1, \ldots, m$, are all different from one another and also none of them is equal to 0. Then the measure μ is determined entirely by its set of nth weighted moments, $n = 1, 2, \ldots$.* □

We will relate this to higher correlations in the next section.

Corollary 4. *Let (X, μ) and w be as in Proposition 11. Then the measure $\widehat{\mu_1^w}$ (which is the almost sure diffraction for the members of X when the weighting is w) is pure point if and only if the dynamical system (X, μ) is pure point, i.e., the linear span of the eigenfunctions is dense in $L^2(X, \mu)$.*

Remark 5. This is the principal result of [19]. See also [11].

Proof. The 'if' direction is a consequence of Corollary 2 of Theorem 3.

The idea behind the 'only if' direction is simple enough. The assumption is that the linear space of the eigenfunctions $L^2(\mathbb{R}^d, \widehat{\mu_1^w})$ is dense in $L^2(\mathbb{R}^d, \widehat{\mu_1^w})$, and eigenfunctions of $L^2(\mathbb{R}^d, \widehat{\mu_1^w})$ map to eigenfunctions of $L^2(X, \mu)$ under θ^w. However, products of eigenfunctions of (X, μ) are again eigenfunctions. We know that the algebra generated by the image of $\mathbb{S}(\mathbb{R}^d)$ in $L^2(X, \mu)$ is dense. So the linear space that we get by taking the algebra generated by the eigenfunctions ought also to be both dense and linearly generated by eigenfunctions. The trouble is that the eigenfunctions of $L^2(\mathbb{R}^d, \widehat{\mu_1^w})$ are not in $\mathbb{S}(\mathbb{R}^d)$ and the space $L^2(X, \mu)$ is not closed under multiplication, so we need to be careful.

The set $BL^2(X, \mu)$ of measurable square integrable functions on X that are bounded on a subset of full measure form an algebra (i.e. the products of such functions are also bounded), and $\theta^w(\mathbb{S})$ is contained in it. In fact any bounded function of $L^2(\mathbb{R}^d, \widehat{\mu_1^w})$ is mapped by θ^w into $BL^2(X, \mu)$, as we can see from Theorem 3 and Eq. (23) and taking approximations by elements of $\mathbb{S}$. For $F \subset BL^2(X, \mu)$, let $L(F)$ denote its linear span and $\langle F \rangle_{\mathrm{alg}}$ the subalgebra of $BL^2(X, \mu)$ generated by F.

By Corollary 3, χ_k is in the point spectrum of U_t if and only if $\widehat{\mu_1^w}(k) \neq 0$, and the eigenfunction corresponding to χ_k is $\theta^w(\mathbf{1}_{\{-k\}})$. Denote by E the set of $\{\mathbf{1}_{\{-k\}} : \widehat{\mu_1^w}(k) \neq 0\}$ and by $L(E)$ its linear span. By Theorem 3, $\theta^w(E)$ is a set of eigenfunctions of T_t, and by what we just saw $\theta^w(L(E)) \subset BL^2(X, \mu)$. By assumption, $L(E)$ is dense in $L^2(\mathbb{R}^d, \widehat{\mu_1^w})$.

Then

$$\overline{L(\theta^w(E))} \supset \theta^w(\overline{L(E)}) \supset \theta^w(\mathbb{S})$$

and

$$BL^2(X, \mu) \supset \langle \theta^w(L(E)) \rangle_{\mathrm{alg}}.$$

Thus,

$$\overline{\langle \theta^w(E) \rangle_{\mathrm{alg}}} = \overline{\langle \theta^w(\overline{L(E)}) \rangle_{\mathrm{alg}}} \supset \overline{\langle \theta^w(\mathbb{S}) \rangle_{\mathrm{alg}}} = L^2(X, \mu),$$

which shows the denseness of the linear span of the eigenfunctions of $L^2(X, \mu)$. □

7. Higher correlations and higher moments

Let $\xi : (\Omega, \mathcal{A}, P) \longrightarrow (X, \mathcal{X})$ be a uniformly discrete stationary multi-variate point process with accompanying dynamical system $(X, \mathbb{R}^d, \mu)$, and let w be a system of weights.

The $n + 1$-point **correlation** ($n = 1, 2, \ldots$) of $\lambda \in X$ is the measure on $(\mathbb{R}^d)^n$ defined by

$$\begin{aligned}
\gamma_\lambda^{(n+1)}(f) &= \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{y_1,\ldots,y_n,x\in C_R} \lambda^w(\{x\}) \prod_{i=1}^n \lambda^w(\{y_i\}) T_x f(y_1, \ldots, y_n) \\
&= \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{\substack{x\in C_R \\ y_1,\ldots,y_n\in\mathbb{R}^d}} \lambda^w(\{x\}) \prod_{i=1}^n \lambda^w(\{y_i\}) T_x f(y_1, \ldots, y_n),
\end{aligned}$$

for all $f \in C_c((\mathbb{R}^d)^n)$. In particular for $f = (f_1, \dots, f_n) \in (C_c(\mathbb{R}^d))^n$, where each $f_i \in C_c(\mathbb{R}^d)$,

$$\begin{aligned}\gamma_\lambda^{(n+1)}(f) &= \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{\substack{x\in C_R\\ y_1,\dots,y_n\in\mathbb{R}^d}} \lambda^w(\{x\}) \prod_{i=1}^{n} \lambda^w(\{y_i\}) T_x f_1(y_1)\dots T_x f_n(y_n)\\ &= \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x\in C_R} \lambda^w(\{x\}) N_{f_1}^w(T_{-x}\lambda)\dots N_{f_n}^w(T_{-x}\lambda)\\ &= \mathrm{Av}^w(N_{f_1}^w \dots N_{f_n}^w)(\lambda).\end{aligned}$$

We know that μ-almost surely this exists, it is independent of λ, and

$$\mathrm{Av}^w(N_{f_1}^w \dots N_{f_n}^w)(\lambda) = \dot{\mu}^w(N_{f_1}^w \dots N_{f_n}^w) =: \dot{\mu}_n^w((f_1, \dots, f_n)).$$

The measure defined on the right-hand side of this equation is the nth **weighted moment of the weighted Palm measure** at $(f_1, \dots, f_n)$, so we arrive at the useful fact which generalizes what we already know for the two-point correlation:

Proposition 12. *The $n+1$-point correlation measure exists almost everywhere on X and is given by $\dot{\mu}_n^w$.*

Of course, in the one-colour case where there are no weights (or if the weighting is trivial: $w = (1, \dots, 1)$), then these are ordinary moments.

Lemma 6. *Assume that the weights are all different and none of them is zero. Then the weighted intensity (and hence the first moment of μ) is determined by the $\dot{\mu}_n^w$, $n = 1, 2, \dots$.*

Proof. By Eq. (22) we need to know $\dot{\mu}^w(X)$. Now $\dot{\mu}^w$ is supported on the set X_0 of elements $\lambda \in X$ which have an atom at $\{0\}$ and we have

$$\begin{aligned}\dot{\mu}^w(X) &= \int_{X_0} N_{\mathbf{1}_{\{0\}}}(\lambda^{\downarrow}) \mathrm{d}\dot{\mu}^w(\lambda) = \int_{X_0} \sum_{i=1}^{m} N_{\mathbf{1}_{\{0\}}}(\lambda^i) \mathrm{d}\dot{\mu}^w(\lambda)\\ &= \sum_{i=1}^{m} \dot{\mu}^w(N_{\mathbf{1}_{\{0\}}} \circ \mathrm{res}^i).\end{aligned}$$

From $\dot{\mu}_1^w$ we have

$$\dot{\mu}_1^w(\{0\}) = \int_{X_0} N_{\mathbf{1}_{\{0\}}}^w(\lambda) \mathrm{d}\dot{\mu}^w(\lambda) = \sum_{i=1}^{m} w_i \dot{\mu}^w(N_{\mathbf{1}_{\{0\}}} \circ \mathrm{res}^i).$$

Similarly,

$$\dot{\mu}_2^w(\{(0,0)\}) = \int_{X_0} (N_{\mathbf{1}_{\{0\}}}^w)^2(\lambda) \mathrm{d}\dot{\mu}^w(\lambda) = \sum_{i=1}^{m} w_i^2 \dot{\mu}^w(N_{\mathbf{1}_{\{0\}}} \circ \mathrm{res}^i).$$

Continue this until we get to

$$\dot{\mu}_m^w(\{(0,\dots,0)\}) = \int_{X_0} (N_{\mathbf{1}_{\{0\}}}^w)^m(\lambda) \mathrm{d}\dot{\mu}^w(\lambda) = \sum_{i=1}^{m} w_i^m \dot{\mu}^w(N_{\mathbf{1}_{\{0\}}} \circ \mathrm{res}^i).$$

Using the same argument in Lemma 3, we can solve this system of equations for $\dot{\mu}^w(N_{\mathbf{1}_{\{0\}}} \circ \mathrm{res}^i)$ for $i = 1, \dots, m$ and hence determine the weighted intensity $\dot{\mu}^w(X)$. $\square$

Theorem 5. *Let (X, μ) be an m-coloured stationary uniformly discrete ergodic point process and w a system of weights in which w_i, $i = 1, \dots, m$, are all different from one another and none of them is equal to 0. Then the measure μ is completely determined by the weighted $n+1$-point correlations of μ-almost surely any $\lambda \in X$, $n = 1, 2, \dots$.*

The key to this is the known fact (in the non-weighted case) that the nth moment of the Palm measure, $n = 1, 2, \ldots$, is the same as the reduced $(n + 1)$st moment of the measure μ itself. Thus knowledge of the correlations gives us the moments $\dot{\mu}_n$ of the Palm measure, which in turn is the same as knowledge of the reduced moments of μ. These in turn determine the moments μ_{n+1}, $n = 1, 2, \ldots$ of μ. As for μ_1, we already know that it is just the intensity of the point process times the Lebesgue measure, and from Lemma 6, this is derivable from the moments.

First of all we give a short derivation of these facts in the unweighted $m = 1$ case, and then show how to augment these to the weighted case.

Let $g, h_1, \ldots h_n \in C_c(\mathbb{R}^d)$ be chosen freely. Let $G : \mathbb{R}^d \times X \longrightarrow \mathbb{C}$ be defined by

$$G(x, \lambda) = g(x) N_{T_x h_1}(\lambda) \ldots N_{T_x h_n}(\lambda).$$

We use the Campbell formula

$$\int_X \sum_{x \in \mathbb{R}^d} \lambda(\{x\}) G(x, \lambda) \mathrm{d}\mu(\lambda) = \int_{\mathbb{R}^d} \int_X G(x, T_x \lambda) \mathrm{d}\dot{\mu}(\lambda).$$

The left-hand side reads[8]

$$\begin{aligned} &\int_X \sum_{x \in \mathbb{R}^d} \lambda(\{x\}) g(x) N_{T_x h_1}(\lambda) \cdots N_{T_x h_n}(\lambda) \mathrm{d}\mu(\lambda) \\ &\quad = \int_X \lambda(g) \lambda(T_x h_1) \cdots \lambda(T_x h_n) \mathrm{d}\mu(\lambda) \\ &\quad = \mu_{n+1}(g(T_x h_1) \cdots (T_x h_n)) = \int_{\mathbb{R}^d} g(x) \mathrm{d}x \, \mu_{n+1}^{\mathrm{red}}(h_1 \cdots h_n), \end{aligned} \tag{35}$$

while the right-hand side reads

$$\int_{\mathbb{R}^d} \int_X g(x) N_{T_x h_1}(T_x \lambda) \cdots N_{T_x h_n}(T_x \lambda) \mathrm{d}\dot{\mu}(\lambda) \mathrm{d}x = \int_{\mathbb{R}^d} g(x) \mathrm{d}x \, \dot{\mu}_n(h_1 \cdots h_n),$$

since $N_{T_x h}(T_x \lambda) = N_h(\lambda)$. For the reduced moments see [6], Sec. 10.4, especially Lemma 10.4.III and Prop. 10.4.V. The point is that μ_{n+1} is invariant under simultaneous translation of its $n+1$ variables. This invariance can be factored out leading to the rewriting of μ_{n+1} as a product of the Lebesgue measure and another measure, which is, *by definition*, the reduced measure. This rewriting is exactly the last part of Eq. (35). Thus, $\mu_{n+1}^{\mathrm{red}} = \dot{\mu}_n$ and, using Proposition 11 and Proposition 12, Theorem 5 is proved in the one-coloured case.

To obtain the weighted version, we use now the functions

$$G^w(x, \lambda) = g(x) N^w_{T_x h_1}(\lambda) \ldots N^w_{T_x h_n}(\lambda)$$

and the weighted form of the Campbell formula. Then the same argument leads to $(\mu^w_{n+1})^{\mathrm{red}} = \dot{\mu}^w_n$, $n = 1, 2, \ldots$. Meanwhile, the first moment is determined by the weighted intensity, given in Lemma 6. The proof of Theorem 5 now follows as in the unweighted case. □

8. Examples

In this section we offer examples that show a variety of ways in which the image of the diffraction appears in the dynamics and in particular how the weighting system influences the outcome. We begin with a general construction of m-coloured uniformly discrete ergodic point processes from symbolic shift systems, which allows one to lift results from the theory of the discrete dynamics of sequences to our situation of continuous dynamics.

The first two of the examples come from well-known results about the Thue–Morse and Rudin–Shapiro sequences. Both sequences lead to dynamical systems $(X, \mathbb{R}, \mu)$ of point sets on the real line, which are uniquely ergodic and minimal but for which the mapping $\theta : L^2(\mathbb{R}, \widehat{I_q}) \to L^2(X, \mu)$ is not surjective. In both cases the diffraction and dynamical spectra are mixed (pure point + singular in the one case, pure point + absolutely continuous in the other).

---

[8] $g(T_x h_1) \ldots (T_x h_n)$ stands for the function whose value on $(x, (y_1, \ldots, y_n)) \in \mathbb{R}^d \times (\mathbb{R}^d)^n$ is $g(x)(T_x h_1)(y_1) \ldots (T_x h_n)(y_n)$.

However the mapping θ does not map the pure point diffraction surjectively to the pure point dynamical spectrum — in fact, it can miss entire spectral components — and this shows that θ itself is not in general surjective. This fact, that the diffraction does not convey full information on the dynamics was pointed out much earlier by van Enter and Miękisz [30].

We then look at extinctions in model sets and observe that even in these most well-behaved sets, the diffraction and dynamical spectra (both of which are pure point) need not match exactly. Finally we give an example to show the necessity of non-zero weights in Theorem 4.

We begin with a short general review of the discrete dynamics of the sequences and look at what happens when we move to the continuous setting by using suspensions. We have done this in slightly more generality than we need for the examples, but with a view to further applications [8].

8.1. The continuous dynamics of sequences on the real line

A good source for examples is to start with symbolic shifts. We start with a finite alphabet $\mathbf{m} = \{1, \ldots, m\}$ and then define $\mathbf{m}^{\mathbb{Z}}$ to be the set of all bi-infinite sequences $\zeta = \{z_i\}_{-\infty}^{\infty}$, which we supply with the product topology. Along with the usual shift action $(T(\zeta))_i = \zeta_{i+1}$ for all i, $\mathbf{m}^{\mathbb{Z}}$ becomes a dynamical system over the group $\mathbb{Z}$. We are interested in compact $\mathbb{Z}$-invariant subspaces $X_{\mathbb{Z}}$ of $(\mathbf{m}^{\mathbb{Z}}, \mathbb{Z})$. We will assume that $(X_{\mathbb{Z}}, \mathbb{Z})$ is equipped with an invariant and ergodic probability measure $\mu_{\mathbb{Z}}$. Such measures always exist. We define for all $\underline{i} = (i_0, \ldots, i_k) \subset \mathbf{m}^{k+1}$, $k = 0, 1 \ldots$, and $p \in \mathbb{Z}$,

$$X_{\mathbb{Z}}[\underline{i}; p] = \{\zeta \in X_{\mathbb{Z}} : z_{j+p} = i_j,\ j = 0, \ldots, k\}.$$

These cylinder sets form a set of entourages for the standard uniform topology on $X_{\mathbb{Z}}$ which defines the product topology. When $p = 0$, we usually leave it out and also leave off the parentheses; so, for example, $X_{\mathbb{Z}}[ij]$ means $X_{\mathbb{Z}}[(ij); 0]$.

We need to move from the discrete dynamics (action by $\mathbb{Z}$) of $(X_{\mathbb{Z}}, \mathbb{Z})$ to continuous dynamics with an $\mathbb{R}$-action. There is a standard way of doing this by creating the suspension flow of $(X_{\mathbb{Z}}, \mathbb{Z})$, and this new dynamical system has a natural invariant and ergodic measure and so satisfies our conditions PPI, PPII, PPIII. Basically each bi-infinite sequence ζ of $(X_{\mathbb{Z}}, \mathbb{Z})$ is converted into a bi-infinite sequence of coloured points on the real line with z_0 being located at 0. The most obvious thing is to space out the other points of the sequence on the integers, so that z_n ends up at position n. The result can be viewed as a tiling of the line with coloured tiles of length 1, the colour of a tile being the colour of the left end point that defines it. However, there are good reasons to allow different colours to have different tile lengths.[9]

For this purpose we take any set $\mathcal{L} = \{L_1, \ldots, L_m\}$ of positive numbers as the tile lengths, with an overall scaling so that

$$\sum_{j=1}^{m} L_j \mu_{\mathbb{Z}}(X_{\mathbb{Z}}[j]) = 1.$$

Let $r = \min\{L_1, \ldots, L_m\}/2$.

Given $\zeta = \{z_n\}_{-\infty}^{\infty} \in X_{\mathbb{Z}}$, define the sequence $S = S(\zeta) = \{S_n\}_{-\infty}^{\infty}$ by $S_0 = 0$, $S_n = \sum_{j=0}^{n-1} L_{z_j}$, if $n > 0$, $S_n = -\sum_{j=n}^{-1} L_{z_j}$ if $n < 0$.

Define

$$\begin{aligned} \pi^{\mathcal{L}} &: \mathbb{R} \times X_{\mathbb{Z}} \longrightarrow \mathcal{D}_r^{(m)}(\mathbb{R}) \\ &(t, \zeta) \mapsto \{(t + S_n, z_n)\}_{-\infty}^{\infty}, \end{aligned} \tag{36}$$

which "locates" the symbols of ζ along the line (including colour information) so that the nth symbol occurs at $t + S_n$. This simultaneously provides us with a tiling of the line by line segments of lengths $\{L_{z_n}\}$. We let $X_{\mathbb{R}}^{\mathcal{L}} := \pi^{\mathcal{L}}(X_{\mathbb{Z}}) \subset \mathcal{D}_r^{(m)}(\mathbb{R})$. Both $\mathbb{R} \times X_{\mathbb{Z}}$ and $\mathcal{D}_r^{(m)}(\mathbb{R})$ have natural $\mathbb{R}$-actions on them, and the mapping $\pi^{\mathcal{L}}$ is $\mathbb{R}$-invariant. It is easy to see that $\pi^{\mathcal{L}}$ is continuous.

[9] Readers interested only in the examples germane to this paper may ignore the introduction of different tile lengths that we introduce here.

Let R be the equivalence relation on $\mathbb{R} \times X_{\mathbb{Z}}$ defined by transitive, symmetric, and reflexive extension of $(t, \zeta) \equiv_R (t + L_{z_0}, T\zeta)$. Evidently pairs are R-equivalent if and only if they have the same image under $\pi^{\mathcal{L}}$. In fact, (t, ζ) is R-equivalent to a unique element of

$$F^{\mathcal{L}} := \bigcup_{i=1}^{m} (-L_i, 0] \times X_{\mathbb{Z}}[i]$$

and the mapping $\pi^{\mathcal{L}}$ is injective on this set. Since $\overline{F^{\mathcal{L}}} = \bigcup [-L_i, 0] \times X_{\mathbb{Z}}[i]$ is compact and $\pi^{\mathcal{L}}$ maps this set onto $X_{\mathbb{R}}^{\mathcal{L}}$, we see that $X_{\mathbb{R}}^{\mathcal{L}}$ is compact and hence $(X_{\mathbb{R}}^{\mathcal{L}}, \mathbb{R})$ is a topological dynamical system.

8.2. Measures on the suspension

We define a positive measure $\mu^{\mathcal{L}}$ on $X_{\mathbb{R}}^{\mathcal{L}}$ by

$$\mu^{\mathcal{L}}(B) := (\ell \otimes \mu_{\mathbb{Z}})((\pi^{\mathcal{L}})^{-1}(B) \cap F^{\mathcal{L}})$$

for all Borel subsets B of $X_{\mathbb{R}}^{\mathcal{L}}$. We observe that $\mu^{\mathcal{L}}$ is a probability measure since $\mu^{\mathcal{L}}(X_{\mathbb{R}}^{\mathcal{L}}) = (\ell \otimes \mu_{\mathbb{Z}})(F^{\mathcal{L}}) = \sum L_i \mu_{\mathbb{Z}}(X[i]) = 1$.

This is an $\mathbb{R}$-invariant measure on $X_{\mathbb{R}}^{\mathcal{L}}$. It suffices to show the shift invariance for sets of the form $J \times C$ where J is an interval in $(-L_i, 0]$ and C is a measurable subset of some $X_{\mathbb{Z}}[i]$, since these sets generate the σ-algebra of all Borel subsets of $F^{\mathcal{L}}$. We show that shifting of J by $s < 0$ leaves the measure invariant. It is sufficient to do this for $|s| < \min\{L_1, \ldots, L_m\}$, since we can repeat the process if necessary to account for larger s. If $s + J \subset (-L_i, 0]$, then the invariance of ℓ gives what we need immediately. If $s + J \not\subseteq (-L_i, 0]$ then we may break J into two parts; the part which remains in the interval and the part which moves out of it to the left. We can restrict our attention to the part that moves out and then assume that $(s + J) \cap (-L_i, 0] = \emptyset$. Then we bring $(s + J) \times C$ back into $F^{\mathcal{L}}$ by writing $C = \bigcup_{j=1}^{m} C \cap X_{\mathbb{Z}}[ij]$ so that

$$(s + J) \times C \equiv_R \bigcup_{j=1}^{m} (L_i + s + J) \times T(C \cap X_{\mathbb{Z}}[ij]).$$

The measure of this is $\sum_{j=1}^{m} \ell(J)\mu_{\mathbb{Z}}(T(C \cap X_{\mathbb{Z}}[ij])) = \ell(J)\mu_{\mathbb{Z}}(C) = (\ell \otimes \mu_{\mathbb{Z}})(J \times C)$, which is what we wished to show.

If the original measure $\mu_{\mathbb{Z}}$ on $X_{\mathbb{Z}}$ is ergodic, then so is the measure $\mu^{\mathcal{L}}$. One way to see this is to start with the case when $\mathcal{L} = \{1, \ldots, 1\}$. In this case we shall denote the objects that we have constructed above with a superscript $\mathbf{1}$ rather than $\mathcal{L}$. It is easy to see that $\mu^{\mathbf{1}}$ is an ergodic measure on $X_{\mathbb{R}}^{\mathbf{1}}$ since the latter can be thought of as $X_{\mathbb{Z}} \times U(1)$, where $U(1)$ is the unit circle in $\mathbb{C}$, with the action of $\mathbb{R}$ being such that going clockwise around the circle once returns one to the same sequence in $X_{\mathbb{Z}}$ except shifted once.

We can define a flow equivalence $\phi : X_{\mathbb{R}}^{\mathbf{1}} \longrightarrow X_{\mathbb{R}}^{\mathcal{L}}$ in the following way. For each $\zeta \in X_{\mathbb{Z}}$ define $f_\zeta : \mathbb{R} \longrightarrow \mathbb{R}$ by

$$f_\zeta(t) = \begin{cases} |S_{-(k-1)}| + L_{z_{-k}}(t - |S_{-(k-1)}|) & \text{if } t \geq 0, k - 1 \leq t < k \\ -S_{k-1} + L_{z_k}(t - S_{k-1}) & \text{if } t \leq 0, k - 1 \leq |t| < k. \end{cases}$$

This is a strictly monotonic piecewise linear continuous function which fixes 0. Its intent is clear: if (t, ζ) is understood to represent the sequence ζ placed down in equal step lengths of one unit starting with z_0 at t, then $(f_\zeta(t), \zeta)$ represents the same sequence, now scaled to the new colour lengths L_{z_j} where 0 is the fixed point.

Thus define a mapping $\mathbb{R} \times X_{\mathbb{Z}} \longrightarrow \mathbb{R} \times X_{\mathbb{Z}}$ by $(t, \zeta) \mapsto (f_\zeta(t), \zeta)$. This mapping factors through the equivalence relations that define $X_{\mathbb{R}}^{\mathbf{1}}$ and $X_{\mathbb{R}}^{\mathcal{L}}$ to give the mapping ϕ which is the flow equivalence that we have in mind. For $I \times X_{\mathbb{Z}}[\underline{u}]$, where $I \subset (-L_{u_0}, 0]$,

$$\phi^{-1}((I \times X_{\mathbb{Z}}[\underline{u}])^{\sim}) = \left(\frac{I}{L_{u_0}} \times X_{\mathbb{Z}}[\underline{u}]\right)^{\sim},$$

where the equivalence relations are taken for $\mathcal{L}$ and for $\mathbf{1}$ respectively. Furthermore, $\mu^{\mathcal{L}}((I \times X_{\mathbb{Z}}[\underline{u}])^{\sim}) = \ell(I)\mu_{\mathbb{Z}}(X_{\mathbb{Z}}[\underline{u}])$ and $\mu^{\mathbf{1}}((I/L_{u_0} \times X_{\mathbb{Z}}[\underline{u}])^{\sim}) = \ell(I/L_{u_0})\mu_{\mathbb{Z}}(X_{\mathbb{Z}}[\underline{u}])$.

Now, if B is an $\mathbb{R}$-invariant subset of $X_{\mathbb{R}}^{\mathcal{L}}$ then $\phi^{-1}(B)$ is an $\mathbb{R}$-invariant subset of $X_{\mathbb{R}}^{\mathbf{1}}$, and so, assuming that $\mu_{\mathbb{Z}}$ is ergodic, $\phi^{-1}(B)$ has measure 1 or 0. If the former, then for all $i \le m$, $\phi^{-1}(B) \cap ((-1, 0] \times X_{\mathbb{Z}}[i])$ has $\mu^{\mathbf{1}}$-measure $\mu_{\mathbb{Z}}(X_{\mathbb{Z}}[i])$ from which $B \cap ((-L_i, 0] \times X_{\mathbb{Z}}[i])$ has measure $L_i \mu_{\mathbb{Z}}(X_{\mathbb{Z}}[i])$, which shows that B is of full measure in $X_{\mathbb{R}}^{\mathcal{L}}$. A similar argument works for the measure 0 case. This shows that $\mu^{\mathcal{L}}$ is ergodic.

8.3. Spectral features of the suspension

At this point we have arrived at the setting of this paper: $(X_{\mathbb{R}}^{\mathcal{L}}, \mathbb{R}, \mu^{\mathcal{L}})$ is a dynamical system satisfying PPI, PPII, PPIII. Henceforth we shall assume that the set of lengths $\mathcal{L} = \{L_1, \dots, L_m\}$ is fixed, and drop them from the notation. We may weight the system by choosing any real vector $w = (w_1, \dots, w_m)$ of weights and assigning weight w_i to the colour a_i. According to Theorem 2, the weighted first moment $\dot{\mu}_1^w$ of the weighted Palm measure is almost everywhere the weighted autocorrelation of the point sets of $X_{\mathbb{R}}$, and this is everywhere true if the system is uniquely ergodic. We will use the symbol w to also denote the mapping $\mathbf{m} \longrightarrow \{w_1, \dots, w_m\}$, $w(i) = w_i$.

We now come to the autocorrelation. For the purposes of the examples, it is convenient to have all tile lengths equal to 1: $L_j = 1$ for all j, and we shall assume this for the remainder of this section.

Now let $\zeta = \{z_i\}_{-\infty}^{\infty} \in X_{\mathbb{Z}}$. Its autocorrelation, assuming that it exists, is

$$\gamma_{\zeta}^{w,\mathbb{Z}} = \sum_{k\in\mathbb{Z}} \eta^w(k)\delta_k^{\mathbb{Z}},$$

defined on $\mathbb{Z}$, where

$$\eta^w(k) := \lim_{N\to\infty} \frac{1}{2N+1} \sum_{i=-N}^{N} w(z_i)w(z_{i+k}).$$

Its autocorrelation $\gamma_{\zeta}^{w,\mathbb{R}}$ when thought of as an element of $X_{\mathbb{R}}$ is defined on $\mathbb{R}$ and is given by

$$\gamma_{\zeta}^{w,\mathbb{R}} = \sum_{k\in\mathbb{Z}} \eta^w(k)\delta_k^{\mathbb{R}},$$

with the same $\eta^w(k)$.

The difference is in the delta measures, which are defined on $\mathbb{Z}$ and $\mathbb{R}$ respectively. Thus $\widehat{\gamma_{\zeta}^{w,\mathbb{Z}}}$ is a measure on $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ while $\widehat{\gamma_{\zeta}^{w,\mathbb{R}}}$ is a measure on $\mathbb{R}$. The relationship between these two measures is simple: for $x \in \mathbb{R}$ and $\dot{x} := x \bmod \mathbb{Z}$,

$$\widehat{\delta_k^{\mathbb{Z}}}(\dot{x}) = \mathrm{e}^{-2\pi \mathrm{i}k.x}, \quad \widehat{\delta_k^{\mathbb{R}}}(x) = \mathrm{e}^{-2\pi \mathrm{i}k.x}.$$

Thus, for all $k \in \mathbb{Z}$, $\widehat{\delta_k^{w,\mathbb{R}}}$ is just the natural periodic extension of $\widehat{\delta_k^{w,\mathbb{Z}}}$ and $\widehat{\gamma_{\zeta}^{w,\mathbb{R}}}$ is the periodization of $\widehat{\gamma_{\zeta}^{w,\mathbb{Z}}}$:

$$\widehat{\gamma_{\zeta}^{w,\mathbb{R}}}(x) = \widehat{\gamma_{\zeta}^{w,\mathbb{Z}}}(\dot{x}).$$

The latter, and hence also the former, exists almost surely.

The pure point, singular continuous, and absolutely continuous parts are also periodized in this process and retain the same types. Thus if the pure point part of $\widehat{\gamma_{\zeta}^{w,\mathbb{Z}}}$ is $\sum_{\dot{k}\in S} a_{\dot{k}}\delta_{\dot{k}}^{\mathbb{Z}}$ then the pure point part of $\widehat{\gamma_{\zeta}^{w,\mathbb{R}}}$ is $\sum_{k\in\mathbb{R},\dot{k}\in S} a_{\dot{k}}\delta_k^{\mathbb{R}}$, where $\dot{k} = k \bmod \mathbb{Z}$.

When it comes to $L^2(X_{\mathbb{Z}}, \mu_{\mathbb{Z}})$ and $L^2(X_{\mathbb{R}}, \mu)$ we make the following observation. If $f_{\dot{k}}$ is an eigenfunction for the action of T on $L^2(X_{\mathbb{Z}}, \mu_{\mathbb{Z}})$ corresponding to the eigenvalue $\dot{k}$ — that is, $T^n f_{\dot{k}} = \exp(2\pi \mathrm{i}k.n) f_{\dot{k}}$ for some (or any) $k \in \mathbb{Z}$ with $\dot{k} = k \bmod \mathbb{Z}$, we can define a function f_k on $X_{\mathbb{R}}$ by

$$f_k(t+\zeta) = \exp(-2\pi \mathrm{i}k.t) f_{\dot{k}}(\zeta).$$

It is easy to see that this is well defined and is an eigenfunction for the $\mathbb{R}$-action on $X_{\mathbb{R}}$ with eigenvalue $-k$ on $\mathbb{R}$. (The change in sign results from the fact that T means shift left by 1, whereas T_t means shift right by t.) This way we see that we have eigenfunctions for $X_{\mathbb{R}}$ which are all the possible continuous lifts of the eigenfunctions on $\mathbb{R}/\mathbb{Z}$ to eigenfunctions on $\mathbb{R}$.

Unfortunately there does not seem to be any simple connection between the other spectral components of $(X_{\mathbb{Z}}, \mu_{\mathbb{Z}})$ and $(X_{\mathbb{R}}, \mu)$. Thus, for these components, we will be reduced to the consequences that come by the embedding of the diffraction into the dynamics.

8.4. The hull of a sequence

We start with an infinite sequence $\xi = (x_1, x_2, \ldots)$ of elements of our finite alphabet $\mathbf{m}$ and define $X_{\mathbb{Z}}(\xi)$ to be the set of all bi-infinite sequences $\zeta = \{z_i\}_{-\infty}^{\infty} \in \mathbf{m}^{\mathbb{Z}}$ with the property that every finite subsequence $\{z_n, z_{n+1}, \ldots, z_{n+k}\}$ (*word*) of ζ is also a word $\{x_p, x_{p+1}, \ldots, x_{p+k}\}$ of ξ. Then set $X_{\mathbb{Z}}(\xi)$ is a closed, and hence compact subset of $\mathbf{m}^{\mathbb{Z}}$, and $(X_{\mathbb{Z}}(\xi), \mathbb{Z})$ is a dynamical system, called the *dynamical hull* of ξ. $(X_{\mathbb{Z}}(\xi), \mathbb{Z})$ is minimal (every orbit is dense) if and only if ξ is repetitive (every word reoccurs with bounded gaps).

Given a word $s = \{x_p, x_{p+1}, \ldots, x_{p+k}\}$ of ζ, we can ask about the frequency of its appearance (up to translation) in ζ. Let $L(s, [M, N])$ be the number of occurrences of s in the interval $[M, N]$. The frequency of s (relative to $t \in \mathbb{Z}$) is $\lim_{N\to\infty} L(s, t + [-N, N])/2N$, if it exists. It is known that the system $X_{\mathbb{Z}}(\xi)$ is both minimal and uniquely ergodic (that is, *strictly ergodic*) if and only if for every $\zeta \in X_{\mathbb{Z}}(\xi)$ and every word s of ζ the frequency of s exists, the limit is approached uniformly for all in $t \in \mathbb{Z}$, and the frequency is positive. All of this is standard from the theory of sequences and symbolic dynamics [25], Cor. IV.12.

We can transform $X_{\mathbb{Z}}(\xi)$ into a flow over $\mathbb{R}$ by the technique discussed in the previous subsection and thus obtain $X_{\mathbb{R}}(\xi)$, which will be minimal (respectively ergodic, uniquely ergodic) according as $X_{\mathbb{Z}}(\xi)$ is.

In the next two subsections we consider situations which are derived from two famous sequences, the Thue–Morse and Rudin–Shapiro sequences.

8.5. Thue–Morse

The Thue–Morse sequence can be defined by iteration of the two-letter substitution (we use a, b instead of $1, 2$)

$$a \to ab; \quad b \to ba: \quad \xi = abbabaabbaababba \ldots.$$

based on the alphabet $A = \{a, b\}$ (we use $\{a, b\}$ instead of $\{1, 2\}$).

Since the substitution is primitive, it is known that the corresponding dynamical system $X_{\mathbb{Z}} = X_{\mathbb{Z}}(\xi)$, and hence also $X_{\mathbb{R}} = X_{\mathbb{R}}(\xi)$, is minimal and uniquely ergodic.

For an arbitrary weighting system $w = (w_a, w_b)$ we have the diffraction $w_a^2\gamma_{aa} + w_a w_b \gamma_{ab} + w_b w_a \gamma_{ba} + w_b^2 \gamma_{bb}$ where γ_{ij} is the correlation between points of types $i, j \in A$. The natural symmetry $a \leftrightarrow b$ of $X_{\mathbb{Z}}$ gives $\gamma_{aa} = \gamma_{bb}, \gamma_{ab} = \gamma_{ba}$.

Kakutani [15,16] has determined the diffraction for the weighting system $w = (1, 0)$ and it is

$$\frac{1}{4}\delta_0 + \mathrm{sc},$$

where sc is a non-trivial singular continuous measure on $\mathbb{Z}$. On the other hand, with the weighting $w = (1, 1)$ the elements of $X_{\mathbb{Z}}$ are all just the sequence $\mathbb{Z}$ as far as the autocorrelation is concerned, and the diffraction is $\delta_{\mathbb{Z}}$. From these it follows that the diffraction for a general weighting system is

$$\left(\frac{w_a + w_b}{2}\right)^2 \delta_0 + \left(\frac{w_a - w_b}{2}\right)^2 \mathrm{sc}.$$

In view of our remarks in Section 6.1, the diffraction for $X_{\mathbb{R}}$ is

$$\left(\frac{w_a + w_b}{2}\right)^2 \delta_{\mathbb{Z}} + \left(\frac{w_a - w_b}{2}\right)^2 \mathrm{scp}$$

where scp is the periodization of $\mathbb{R}$ of the measure sc on $\mathbb{T}$.

The dynamical system is also mixed, pure point plus singular continuous [18]. There is an obvious continuous involution $\sim$ on $X_{\mathbb{Z}}$ that interchanges the a and b symbols. $L^2(X_{\mathbb{Z}}, \mu_{\mathbb{Z}})$ splits into the ± 1-eigenspaces for $\sim$:

$L^2(X_{\mathbb{Z}}, \mu_{\mathbb{Z}}) = L^2_+(X_{\mathbb{Z}}) \bigoplus L^2_-(X_{\mathbb{Z}})$. $L^2_+(X_{\mathbb{Z}})$ is the pure point part of $L^2(X_{\mathbb{Z}}, \mu)$ and its eigenvalues are all the numbers of the form $k/2^n$, $n = 0, 1, \ldots; 0 \le k < 2^n$ (literally $\exp(2\pi i k/2^n)$). On the other hand $L^2_-(X_{\mathbb{Z}})$ is singular continuous.

When we move to the suspension of $X_{\mathbb{Z}}$ we obtain $L^2(X_{\mathbb{R}}, \mu)$ which we know certainly retains the eigenvalues of $L^2(X_{\mathbb{Z}}, \mu_{\mathbb{Z}})$ and, due to the embedding of $L^2(\mathbb{R}, \widehat{\dot{\mu}_1^{(1,0)}})$, also retains a singular continuous component.

The dynamical spectrum is, of course, independent of any particular assignments of weights to a and b. We can draw the following conclusions from this:

(i) $w_a = 1, w_b = 0$.

$$\widehat{\dot{\mu}_1^{(1,0)}} = \frac{1}{4}\delta_{\mathbb{Z}} + \text{scp}.$$

The eigenfunctions of $L^2(\mathbb{R}, \widehat{\dot{\mu}_1^{(1,0)}})$ are $\mathbf{1}_{\{k\}}, k \in \mathbb{Z}$. It follows that $\theta^w(\mathbf{1}_{\{-k\}})$ is an eigenfunction for eigenvalue k (Theorem 3, Corollary 2). Thus θ^w covers only the eigenvalues $k \in \mathbb{Z}$ of $L^2(X_{\mathbb{R}})$ and none of the fractional ones $k/2^n, n > 0$. This shows that θ^w is not surjective. Also θ^w embeds the singular continuous part of $L^2(\mathbb{R}, \widehat{\dot{\mu}_1^{(1,0)}})$ into $L^2_-(\mathbb{R})$, although we do not know the image.

(ii) $w_a = w_b = 1$. In this case the diffraction is $\delta_{\mathbb{Z}}$ (the Thue–Morse sequence with equal weights looks like $\mathbb{Z}$). Although $L^2(\mathbb{R}, \widehat{\dot{\mu}_1^{(1,1)}})$ is pure point, its image does not cover the pure point part of $L^2(X_{\mathbb{R}}, \mu)$, nor does it even generate it as an algebra. This shows that the requirement of unequal weights in Theorem 4 is necessary.

(iii) $w_a = 1, w_b = -1$. This time the diffraction is singular continuous and θ^w does not even cover anything of the pure point part of $L^2(X_{\mathbb{R}}, \mu)$.

Cases (ii) and (iii) show that the non-existence of a particular component in the diffraction spectrum implies nothing about its existence or non-existence in the dynamical spectrum.

8.6. Rudin–Shapiro

We define the Rudin–Shapiro sequence using the notation of [24]. Consider the substitution rule s defined on the alphabet $A' := \{1, \bar{1}, 2, \bar{2}\}$ as follows: $s(1) = 1\bar{2}$, $s(2) = \bar{1}\bar{2}$, $s(\bar{1}) = \bar{1}2$, $s(\bar{2}) = 12$. Let ξ be the s-invariant sequence that starts with the symbol 1. We can reduce this to a two-symbol sequence ξ' with alphabet $\{a, b\}$ by replacing the symbols with no overbar by the letter a and the others by the letter b. This two-symbol sequence is usually called the Rudin–Shapiro sequence [25], though Priebe–Frank uses this appellation for the original four-symbol sequence.

Let us start with the two-symbol sequence, which results in the two-coloured minimal and ergodic dynamical hull $(X_{\mathbb{Z}}(\xi'), \mathbb{Z})$, as developed above. There is a natural involution on the dynamical system that interchanges a and b. Once again we introduce a system of weights $w = (w_a, w_b)$.

Under the system of weights $(1, -1)$ it is well known that the diffraction measure of the elements of $X_{\mathbb{Z}}(\xi')$ is the normalized Haar measure on $\mathbb{R}/\mathbb{Z}$ [25], Cor. VIII.5. Thus $L^2(\mathbb{R}, \widehat{\dot{\mu}_1^{(1,-1)}}) = L^2(\mathbb{R}, \ell)$, where ℓ is the Lebesgue measure on $\mathbb{R}$.

On the other hand, the weighting system $(1, 1)$ reduces the elements of $X_{\mathbb{Z}}(\xi')$ to copies of the sequence $\mathbb{Z}$. So, just as in the case of the Thue–Morse sequence, we can deduce the general formula for the diffraction:

$$\left(\frac{w_a + w_b}{2}\right)^2 \delta_{\mathbb{Z}} + \left(\frac{w_a - w_b}{2}\right)^2 \ell.$$

The spectral decomposition of $L^2(X_{\mathbb{Z}}(\xi'))$ is of the form

$$L^2(X_{\mathbb{Z}}(\xi')) \simeq H \oplus Z(f)$$

where H is the pure point part with one simple eigenvalue $\exp(2\pi i q)$ for each dyadic rational number $q = a/2^n$, where $a \in \mathbb{Z}, n = 0, 1, 2, \ldots$ [7,22]; and $Z(f)$ is a cyclic subspace which is equivalent to $L^2(\mathbb{R}, \ell)$. In other words, the dynamical spectrum is mixed with a pure point and an absolutely continuous part.[10] In any case, we see that $L^2(X_{\mathbb{R}}(\xi'), \mu)$ contains a pure point part whose eigenvalues include all the dyadic rationals, and also an absolutely continuous part into which the absolutely continuous part of $L^2(\mathbb{R}, \widehat{\mu_1^w})$ must map by θ^w.

The analysis now proceeds exactly as in the case of the Thue–Morse sequence, with the same three types of possibilities except that now the singular continuous parts are replaced by absolutely continuous parts.

8.7. Regular model sets

In this example we see that even when everything is pure point and there is only one colour, still θ need not be surjective.

Let $(\mathbb{R}^d, \mathbb{R}^d, L)$ be a cut and project scheme with projection mappings π_i, $i = 1, 2$. Thus L is a lattice in $\mathbb{R}^d \times \mathbb{R}^d$, the projection π_1 to the first factor is one–one on L, and the projection $\pi_2(L)$ of L has dense image in the second factor. Let W be a non-empty compact subset which is the closure of its own interior and a subset of the second factor. We assume that the boundary of W has Lebesgue measure 0. The corresponding model set is

$$\Lambda(W) = \{\pi_1(t) : t \in L, \pi_2(t) \in W\}.$$

It is a subset of $\mathcal{D}_r$ for some $r > 0$ and it is pure point diffractive [13,28,3]. The orbit closure $X = \overline{\mathbb{R}^d + \Lambda(W)}$ is uniquely ergodic. Its autocorrelation γ, and hence its diffraction $\widehat{\gamma}$, is the same for all $\Gamma \in X$. Furthermore, the diffraction is explicitly known:

$$\widehat{\mu_1} = \widehat{\gamma} = \sum_{k \in L^0} a_k \delta_{\pi_1(k)}$$

where L^0 is the $\mathbb{Z}$-dual lattice of L with respect to the standard inner product on $\mathbb{R}^d \times \mathbb{R}^d \simeq \mathbb{R}^{2d}$ and

$$a_k = \left|\widehat{\mathbf{1}_W}(-\pi_2(k))\right|^2 .$$

For more on this see [13]. The main point is that $\widehat{\mu_1}(\pi_1(k)) = 0$ if and only if $a_k = 0$.

Likewise $L^2(X, \mu)$ is known and it is isometric in a totally natural way by an $\mathbb{R}^d$-map to $L^2(\mathbb{R}^{2d}/\mathbb{Z}^{2d}, \nu)$, where ν is the Haar measure on the torus. Thus the spectrum of X is pure point and the eigenvalues are precisely all the points of L^0. Thus the mapping θ embedding the diffraction into the dynamics will be surjective if and only if for all $k \in L^0$, $a_k \neq 0$.

Now it is easy to see that we can find model sets for our given cut and project scheme which fail to be surjective at any $k \in L^0$ that we wish, as long as $k \neq 0$. To do this take W to be something simple like a ball centred on 0 and for each scaling factor $s > 0$ let $\Lambda^{(s)} := \Lambda(sW)$. The intensities of the Bragg peaks become

$$a_k^{(s)} := s^2 \left|\widehat{\mathbf{1}_W}(-s\pi_2(k))\right|^2 .$$

Since $\widehat{\mathbf{1}_W}$ is continuous and takes positive and negative values on every ray through 0 in $\mathbb{R}^d$, but altogether takes the value 0 only on a meagre set, we see that by choosing s suitably we can arrange either that $a_k^{(s)}$ vanishes at any preassigned non-zero $k \in L^0$ (and θ is not surjective) or that alternatively $a_k^{(s)}$ vanishes nowhere on L^0 (and θ is a bijection).

8.8. The necessity of non-zero weights in Theorem 4

Let $\Lambda = (\Lambda_a, \Lambda_b)$ where

$$\Lambda_a = \{z \in \mathbb{Z} : z \equiv 0 \text{ or } 2 \mod 4\}, \quad \Lambda_b = \{z \in \mathbb{Z} : z \equiv 3 \mod 4\}.$$

[10] Explicitly f is the function on $X_{\mathbb{Z}}(\xi')$ which is defined by $f(\zeta) = 1 \text{or} - 1$ according as $\zeta(0)$ is a or b. This can be deduced from the main theorem of [24], where the equivalent result for the dynamical system arising from the four-symbol sequence gives two copies of $L^2(\mathbb{R}, \ell))$, by dropping to the factor.

Then Λ is periodic with period 4 and its hull — that is, the closure of its $\mathbb{R}$ translation orbit — is $X \simeq \mathbb{R}/4\mathbb{Z}$ (a conjugacy of dynamical systems with the standard action of $\mathbb{R}$ on $\mathbb{R}/4\mathbb{Z}$). Thus $L^2(X,\mu)$, where μ is the Haar measure on $\mathbb{R}/4\mathbb{Z}$, has pure point spectrum with eigenvalues $\frac{1}{4}\mathbb{Z}$.

Let (w_a, w_b) be a weighting system for Λ. The autocorrelation is everywhere the same and is easily seen to be

$$\dot{\mu}_1 = \frac{1}{2}w_a^2\delta_{2\mathbb{Z}} + \frac{1}{4}w_a w_b \delta_{1+4\mathbb{Z}} + \frac{1}{4}w_a w_b \delta_{-1+4\mathbb{Z}} + \frac{1}{4}w_b^2\delta_{4\mathbb{Z}}.$$

The Fourier transform, that is the diffraction, is then given by

$$\begin{aligned}\widehat{\dot{\mu}_1} &= \frac{1}{4}w_a^2\delta_{\frac{1}{2}\mathbb{Z}} + \frac{1}{16}w_a w_b \exp(-2\pi \mathrm{i}(\cdot))\delta_{\frac{1}{4}\mathbb{Z}} + \frac{1}{16}w_a w_b \exp(2\pi \mathrm{i}(\cdot))\delta_{\frac{1}{4}\mathbb{Z}} + \frac{1}{16}w_b^2\delta_{\frac{1}{4}\mathbb{Z}} \\ &= \frac{1}{4}\left\{\left(w_a^2 + \frac{1}{2}w_a w_b + \frac{1}{4}w_b^2\right)\delta_{\mathbb{Z}} + \left(w_a^2 - \frac{1}{2}w_a w_b + \frac{1}{4}w_b^2\right)\delta_{\frac{1}{2}+\mathbb{Z}}\right. \\ &\quad \left. + \frac{1}{4}w_b^2\delta_{\frac{1}{4}+\mathbb{Z}} + \frac{1}{4}w_b^2\delta_{-\frac{1}{4}+\mathbb{Z}}\right\}.\end{aligned}$$

Now it is clear that the image of θ can only generate eigenfunctions for the eigenvalues $\pm\frac{1}{4}+\mathbb{Z}$ if $w_b \neq 0$ (and then in fact it does so, independently of the value of w_a).

9. The square-mean Bombieri–Taylor conjecture

Theorem 6 (*The Square-Mean Bombieri–Taylor Conjecture*). *Let $(X, \mathbb{R}^d, \mu)$ be a uniformly discrete, multi-coloured stationary ergodic point process, and assume that w is a system of weights. Then the following are equivalent*[11]*:*

(i)

$$\frac{1}{\ell(C_R)}\sum_{x\in C_R}\lambda^w(\{x\})\mathrm{e}^{2\pi \mathrm{i}k.x} \nrightarrow 0 \quad \textit{as } R \to \infty;$$

(ii) $\widehat{\dot{\mu}_1^w}(\{k\}) \neq 0$;

(iii) *k is an eigenvalue of U.*

In the case that k is an eigenvalue, then

$$\frac{1}{\ell(C_R)}\sum_{x\in C_R}\lambda^w(\{x\})\mathrm{e}^{2\pi \mathrm{i}k.x} \to \theta^w(\mathbf{1}_k).$$

For notational simplicity we shall prove the two technical lemmas that precede the main proof in the one-dimensional case. However, it is easy to generalize the proof to any dimension d. Throughout, R is assumed to be a positive integer variable.

Lemma 7. *For all $\epsilon > 0$,*

$$\lim_{\epsilon\to 0}\widehat{\dot{\mu}_1^w}(\mathbf{1}_{[-\epsilon,\epsilon]}) = \widehat{\dot{\mu}_1^w}(\{0\}),$$

i.e. $\{\mathbf{1}_{[-\epsilon,\epsilon]}\}_{\epsilon\searrow 0} \longrightarrow \mathbf{1}_{\{0\}}$ *in* $L^2(\mathbb{R}, \widehat{\dot{\mu}_1^w})$.

Proof. Assume $\epsilon \to 0^+$. Let $F_\epsilon := \mathbf{1}_{[-\epsilon,\epsilon]} - \mathbf{1}_{\{0\}}$. Then for all $x \in \mathbb{R}$, $0 \le F_\epsilon(x) \le 1$ and $F_\epsilon(x) \searrow 0$ pointwise. Since $\widehat{\dot{\mu}_1^w}$ is a translation bounded positive measure, $\widehat{\dot{\mu}_1^w}(F_\epsilon) \searrow 0$. Now,

$$\int |\mathbf{1}_{[-\epsilon,\epsilon]} - \mathbf{1}_{\{0\}}|^2 \mathrm{d}\widehat{\dot{\mu}_1^w} = \int F_\epsilon^2 \mathrm{d}\widehat{\dot{\mu}_1^w} \le \int F_\epsilon \mathrm{d}\widehat{\dot{\mu}_1^w} \longrightarrow 0. \quad \square$$

[11] Limits here are taken in the L^2-norm on $(X, \mathbb{R}^d, \mu)$. Recently D. Lenz [20] has given a pointwise version of the result.

Lemma 8. *As functions of* $y \in \mathbb{R}$,

$$\frac{1}{2R}\int_{-R}^{R} \mathrm{e}^{2\pi \mathrm{i} y.x}\mathrm{d}x \longrightarrow \mathbf{1}_{\{0\}}(y)$$

in $L^2(\mathbb{R}, \widehat{\dot{\mu}_1^w})$ *as* $R \to \infty$.

Proof. Let

$$g_R(y) := \frac{1}{2R}\int_{-R}^{R} \mathrm{e}^{2\pi \mathrm{i} y.x}\mathrm{d}x = \frac{\sin(2\pi y R)}{2\pi y R}.$$

We need to show that $\int_{-\infty}^{\infty} |g_R(y) - \mathbf{1}_{\{0\}}(y)|^2 \mathrm{d}\widehat{\dot{\mu}_1^w}(y) \longrightarrow 0$. Since $|g_R(y) - \mathbf{1}_{\{0\}}(y)| \le F_\epsilon(y)$ for $-\epsilon \le y \le \epsilon$, we have $\int_{-\epsilon}^{\epsilon} |g_R(y) - \mathbf{1}_{\{0\}}(y)|^2 \mathrm{d}\widehat{\dot{\mu}_1^w}(y) \longrightarrow 0$ as $\epsilon \to 0$, and the convergence is uniform without reference to R.

For the remaining parts of the integral, we have (the part from $-\infty$ to $-\epsilon$ is the same)

$$\begin{aligned}
\int_{\epsilon}^{\infty} |g_R(y) - \mathbf{1}_{\{0\}}(y)|^2 \mathrm{d}\widehat{\dot{\mu}_1^w}(y) &= \int_{\epsilon}^{\infty} \frac{\sin^2(2\pi y R)}{(2\pi y R)^2}\mathrm{d}\widehat{\dot{\mu}_1^w}(y) \\
&\le \frac{1}{(2\pi R\epsilon)^2}\int_{\epsilon}^{\infty} \frac{\epsilon^2}{y^2}\mathrm{d}\widehat{\dot{\mu}_1^w}(y) \\
&\le \frac{1}{(2\pi R\epsilon)^2}\left\{\int_{\epsilon}^{\epsilon+1} \mathrm{d}\widehat{\dot{\mu}_1^w}(y) + \sum_{m=1}^{\infty}\frac{\epsilon^2}{m^2}\int_{\epsilon+m}^{\epsilon+m+1} \mathrm{d}\widehat{\dot{\mu}_1^w}(y)\right\}.
\end{aligned}$$

Since $\int_a^{a+1} \mathrm{d}\widehat{\dot{\mu}_1^w}(y)$ is uniformly bounded by some constant $C(1)$ (due to the translation boundedness of $\widehat{\dot{\mu}_1^w}$, Theorem 2) we see that

$$\int_{\epsilon}^{\infty} |g_R(y) - \mathbf{1}_{\{0\}}(y)|^2 \mathrm{d}\widehat{\dot{\mu}_1^w}(y) \longrightarrow 0$$

as long as $R\epsilon \to \infty$ as $R \to \infty$. Putting $\epsilon = R^{-1/2}$ gives the necessary convergence for both parts. □

Proof Theorem 6. (ii) ⇔ (iii): k is an eigenvalue if and only if $-k$ is an eigenvalue, $\widehat{\dot{\mu}_1^w}(\{k\}) = \widehat{\dot{\mu}_1^w}(\{-k\})$ for all k, and k is an eigenvalue if and only if $\widehat{\dot{\mu}_1^w}(\{k\}) \neq 0$.
(iii) ⇔ (i): Let $f_R := \frac{1}{2R}\chi_k \mathbf{1}_{[-R,R]}$ (see Eq. (19)). Then

$$\begin{aligned}
\widehat{f_R}(y) = &\ \frac{1}{2R}\int_{\mathbb{R}} \mathrm{e}^{-2\pi \mathrm{i} y.x}\chi_k(x)\mathbf{1}_{[-R,R]}(x)\mathrm{d}x \\
= &\ \frac{1}{2R}\int_{-R}^{R} \mathrm{e}^{2\pi \mathrm{i}(k-y).x}\mathrm{d}x \\
\longrightarrow &\ \mathbf{1}_{\{0\}}(k-y) = \mathbf{1}_{\{k\}}(y),
\end{aligned}$$

the convergence being as functions of y in $L^2(\mathbb{R}, \widehat{\dot{\mu}_1^w})$ as $R \to \infty$.

Let $\phi_{-k} = \theta^w(\mathbf{1}_{\{k\}})$. Thus $\widehat{f_R} \to \mathbf{1}_{\{k\}}$ implies that $\theta^w(\widehat{f_R}) \to \phi_{-k}$ in $L^2(X, \mu)$, so

$$\int_X |N^w_{f_R}(\lambda) - \phi_{-k}(\lambda)|^2 \mathrm{d}\mu(\lambda) \to 0,$$

which from (4) gives

$$\int_X \left|\frac{1}{2R}\sum_{x\in[-R,R]} \lambda^w(\{x\})\mathrm{e}^{2\pi \mathrm{i}k.x} - \phi_{-k}(\lambda)\right|^2 \mathrm{d}\mu(\lambda) \to 0.$$

Thus $\frac{1}{2R}\sum_{x\in[-R,R]} \lambda^w(\{x\})\mathrm{e}^{2\pi \mathrm{i}k.x}$ converges in square mean to ϕ_{-k}. Furthermore by Corollary 3, ϕ_{-k} is a χ_{-k}-eigenfunction for T_t if $\widehat{\dot{\mu}_1^w}(k) \neq 0$ and is 0 otherwise.

If $\phi_{-k} = 0$ then $\frac{1}{2R}\sum_{x\in[-R,R]} \lambda^w(\{x\})\mathrm{e}^{2\pi \mathrm{i}k.x} = 0$ μ-a.e. If $\phi_{-k} \neq 0$ then $\{\lambda : \phi_{-k}(\lambda) = 0\}$ is a measurable T-invariant subset of μ, since ϕ_{-k} is an eigenfunction, so by the ergodicity it is of measure 0 or 1. It must be the former. Now using the Fischer–Riesz theorem [9], there is a subsequence of $\{\frac{1}{2R}\sum_{x\in[-R,R]}\lambda^w(\{x\})e^{2\pi i k.x}\}_R$ which converges pointwise μ-a.e. to ϕ_{-k}. Since ϕ_{-k} is almost everywhere not zero, $\frac{1}{2R}\sum_{x\in[-R,R]}\lambda^w(\{x\})e^{2\pi i k.x} \nrightarrow 0$. $\square$

10. A strange inequality

Let $(X, \mathbb{R}^d, \mu)$ be a uniformly discrete stationary ergodic point process (no colour). Assume that the point sets of X have finite local complexity, μ-a.s. This implies that the autocorrelation measure $\dot{\mu}_1$ is supported on a closed discrete subset of $\Lambda - \Lambda$ for any Λ whose autocorrelation is $\dot{\mu}_1$. Thus for $\Lambda \in X$ we have, μ-almost surely,

$$\dot{\mu}_1(t) = \lim_{R\to\infty} \frac{1}{\ell(C_R)} \mathrm{card}((-t+\Lambda)\cap\Lambda\cap C_R).$$

Proposition 13. *For all $k, t \in \mathbb{R}^d$,*

$$|e^{2\pi i k\cdot t} - 1|\widehat{\dot{\mu}_1}^{1/2}(k) \le 2(\dot{\mu}_1(0) - \dot{\mu}_1(t)).$$

Proof. Let $k \in \mathbb{R}^d$. Then

$$\frac{1}{\ell(C_R)} \sum_{x\in\Lambda\cap C_R} e^{-2\pi i k\cdot x} \longrightarrow g_k$$

in the norm of $L^2(X, \mu)$, where g_k is an eigenfunction of T for the eigenvalue k if $\widehat{\dot{\mu}_1}(k) \neq 0$ and 0 otherwise (Theorem 6).

Suppose $\widehat{\dot{\mu}_1}(k) \neq 0$. Let $t \in \mathbb{R}^d$. Since

$$(T_t g_k)(\Lambda) = g_k(-t+\Lambda) = \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x\in(-t+\Lambda)\cap C_R} e^{-2\pi i k\cdot x}$$

for almost all $\Lambda \in X$,

$$\begin{aligned}((T_t - 1)g_k)(\Lambda) &= \lim_{R\to\infty} \frac{1}{\ell(C_R)} \left\{ \sum_{x\in(-t+\Lambda)\cap C_R} e^{-2\pi i k\cdot x} - \sum_{x\in\Lambda\cap C_R} e^{-2\pi i k\cdot x} \right\} \\ &= \lim_{R\to\infty} h_R(\Lambda),\end{aligned}$$

where

$$h_R(\Lambda) := \frac{1}{\ell(C_R)} \left\{ \sum_{x\in(-t+\Lambda)\setminus\Lambda\cap C_R} e^{-2\pi i k\cdot x} - \sum_{x\in\Lambda\setminus(-t+\Lambda)\cap C_R} e^{-2\pi i k\cdot x} \right\}.$$

Thus $h_R \to (e^{2\pi i k\cdot t} - 1)g_k$ in the L^2-norm on X.

Furthermore,

$$\begin{aligned}|h_R(\Lambda)| &\le \frac{1}{\ell(C_R)} \left(\sum_{x\in(-t+\Lambda)\setminus\Lambda\cap C_R} |e^{-2\pi i k\cdot x}| + \sum_{x\in\Lambda\setminus(-t+\Lambda)\cap C_R} |e^{-2\pi i k\cdot x}| \right) \\ &\le \frac{1}{\ell(C_R)} \left(\sum_{x\in(-t+\Lambda)\setminus\Lambda\cap C_R} 1 + \sum_{x\in\Lambda\setminus(-t+\Lambda)\cap C_R} 1 \right) \\ &= \frac{1}{\ell(C_R)} \sum_{x\in(-t+\Lambda)\triangle\Lambda\cap C_R} 1 = 2(\dot{\mu}_1(0) - \dot{\mu}_1(t)).\end{aligned}$$

Note that $\dot{\mu}_1(0) \ge \dot{\mu}_1(t)$.

With these preliminaries out of the way, the rest of the proof is straightforward. Since μ is a finite measure, $h_R \to (\mathrm{e}^{2\pi \mathrm{i} k\cdot t}-1)g_k$ in the L^1-norm also. Then there is a subsequence $\{h_{R_i}\}$ of $\{h_R\}$ which converges to $(\mathrm{e}^{2\pi \mathrm{i} k\cdot t}-1)g_k$ pointwise almost everywhere ([5], Sec. 3.1).

Using the dominated convergence theorem ($|h_R(\Lambda)| \le 2\dot{\mu}_1(0)$), we have

$$\begin{aligned}\int_X \left|\mathrm{e}^{2\pi \mathrm{i} k\cdot t} - 1\right|^2 |g_k(\Lambda)|^2 \,\mathrm{d}\mu(\Lambda) &= \lim_{R_i\to\infty} \int_X \left|h_{R_i}(\Lambda)\right|^2 \mathrm{d}\mu(\Lambda) \le \int_X |2(\dot{\mu}_1(0) - \dot{\mu}_1(t))|^2 \,\mathrm{d}\mu(\Lambda)\\ &= 4\,|(\dot{\mu}_1(0) - \dot{\mu}_1(t))|^2 .\end{aligned}$$

Meanwhile, from Theorem 3

$$\begin{aligned}\int_X \left|\mathrm{e}^{2\pi \mathrm{i} k\cdot t} - 1\right|^2 |g_k(\Lambda)|^2 \,\mathrm{d}\mu(\Lambda) &= |\mathrm{e}^{2\pi \mathrm{i} k\cdot t} - 1|^2 \int_{\mathbb{R}^d} \mathbf{1}_k^2 \mathrm{d}\widehat{\dot{\mu}_1}\\ &= |\mathrm{e}^{2\pi \mathrm{i} k\cdot t} - 1|^2 \widehat{\dot{\mu}_1}(k).\end{aligned}$$

So

$$\left|\mathrm{e}^{2\pi \mathrm{i} k\cdot t} - 1\right| \widehat{\dot{\mu}_1}^{\frac{1}{2}}(k) \le 2|(\dot{\mu}_1(0) - \dot{\mu}_1(t))| = 2(\dot{\mu}_1(0) - \dot{\mu}_1(t)). \quad \square$$

Remark 6. This result numerically links three interesting quantities. If $\Lambda \in X$ has autocorrelation $\dot{\mu}_1$ then for the set $P(\epsilon)$ of ϵ-statistical almost periods of Λ, i.e. t for which $\dot{\mu}_1(0) - \dot{\mu}_1(t) < \epsilon$, the Bragg peaks $I(a)$ of intensity greater than $a > 0$, i.e. k for which $\widehat{\dot{\mu}_1}(k) > a$, can occur only at points k which are $2\epsilon/\sqrt{a}$-dual to $P(\epsilon)$, i.e. k for which $\left|\mathrm{e}^{2\pi \mathrm{i} k\cdot t} - 1\right| < 2\epsilon/\sqrt{a}$ for all $t \in P(\epsilon)$. If this latter quantity is less than or equal to $1/2$ and either of $P(\epsilon)$ or $I(a)$ is relatively dense, then the other one is a Meyer set ([23] Thm. 9.1). Furthermore, Bragg peaks can occur only on the $\mathbb{Z}$-dual of the statistical periods (t for which $\dot{\mu}_1(t) = \dot{\mu}_1(0)$), a fact that is of course very familiar in the case of crystals.[12] We note that the inequality seems to be optimal. The maximum values of $\left|\mathrm{e}^{2\pi \mathrm{i} k\cdot t} - 1\right|$ and $\widehat{\dot{\mu}_1}^{\frac{1}{2}}(k)$ are 2 and $\dot{\mu}_1(0)$ respectively, whereas the minimum value of $\dot{\mu}_1(t)$ is 0.

11. Patterns and pattern frequencies

Let $(X, \mathbb{R}^d, \mu)$ be a multi-colour uniformly discrete stationary ergodic point process. It is of interest to define the frequency of finite colour patterns in X. This is made difficult because from the built in vagueness of the topology of X we know that we should not be looking for exact matches of some given colour pattern F of $\mathcal{D}_r^{(m)}$, but rather close approximations to it. In addition there is the problem of how to anchor F, in order to specify it exactly as we move it around. This leads us to always assume that F contains 0, and then to define a **pattern** in X as a pair (F, V) where $F = \cup_{i=1}^m (F_i, i)$ is a finite subset of $\mathcal{D}_r^{(m)}$ with $0 \in F^{\downarrow} := \cup F_i$ and V is a bounded measurable neighbourhood of 0 in $\mathbb{R}^d$. For a pattern (F, V) we then define the collection of elements of X that contain it as

$$X_{F,V} = X_{(F,V)} := \{\Lambda \in X : F \subset V + \Lambda\},$$

and write $\mathbf{1}_{F,V}$ for $\mathbf{1}_{X_{F,V}}$.

Throughout one should keep in mind that F and Λ are multi-colour sets; our conventions are that translations are by elements of $\mathbb{R}^d$ and are always on the left, and the inclusions take colour into account.

For any bounded region B define

$$L_{F,V}(\Lambda, B) := \mathrm{card}\{x \in \Lambda^{\downarrow} : F \subset -x + V + \Lambda, x - V + F \subset B\}.$$

An initial idea for the frequency of the pattern (F, V) in a set $\Lambda \in X$ might be

$$\mathrm{freq}(\Lambda, F, V) = \lim_{R\to\infty} \frac{1}{\ell(C_R)} L_{F,V}(\Lambda, C_R). \tag{37}$$

[12] We are grateful to Nicolae Strungaru for this last observation.

This definition is very sensitive to the boundary of V, as one can see from the simple example below. In general we do not know how to prove that this limit exists, even almost everywhere in X. However, we can prove that for V open or V closed, if the limit does exist then it is, almost surely, given by the Palm measure of $X_{F,V}$, and this, we know, does exist. Thus we are led to define:

The **frequency** of the pattern (F, V) in X is $\dot{\mu}(X_{F,V})$.

The connection with Palm measures comes because (as is easy to see from the van Hove property of expanding cubes)

$$\lim_{R\to\infty} \frac{1}{\ell(C_R)} L_{F,V}(\Lambda, C_R) = \lim_{R\to\infty} \frac{1}{\ell(C_R)} \sum_{x\in\Lambda^{\downarrow}\cap C_R} \mathbf{1}_{F,V}(-x+\Lambda).$$

The latter is the average over Λ (where the weighting system is all 1s) of $\mathbf{1}_{F,V}$, if it exists.

Proposition 14. *Let (F, V) be a pattern with V an open set. Then*

$$\underline{\mathrm{freq}}(\Lambda, F, V) = \dot{\mu}(X_{F,V})$$

μ-almost surely for $\Lambda \in X$, where $\underline{\mathrm{freq}}$ means that the lim inf *is taken in (37). Similarly, if (F, V) is a pattern with V a closed set, then*

$$\overline{\mathrm{freq}}(\Lambda, F, V) = \dot{\mu}(X_{F,V})$$

μ-almost surely for $\Lambda \in X$, where $\overline{\mathrm{freq}}$ means that the lim sup *is taken in (37).*

Lemma 9. *$X_{F,V}$ is open if V is bounded and open and closed if V is bounded and closed.*

Proof. Let V be open and let $\Lambda \in X_{F,V}$. Then $F \subset V + \Lambda$. Since V is open and F is finite, there is an $\epsilon > 0$ so that for each $f \in F$, with $f = v + x$, where $v \in V$, $x \in \Lambda$ (there may be choices, but fix one choice x for each f), $f + C_\epsilon \subset V + x$. Choose $R > 0$ so that $-V + F \subset C_R$. Let $\Lambda' \in U(\overline{C_R}, C_\epsilon)[\Lambda]$ and let $f = v + x \in F$, as above. Since $\overline{C_R} \cap \Lambda \subset C_\epsilon + \Lambda'$ and $x \in \overline{C_R} \cap \Lambda$, $x = c + x'$ where $x' \in \Lambda'$, $c \in C_\epsilon$. Then $c + f \in V + x$, so $f \in V + x'$.

Since $f \in F$ was arbitrary, $F \subset V + \Lambda'$ and $\Lambda' \in X_{F,V}$. Thus the open neighbourhood $U(\overline{C_R}, C_\epsilon)[\Lambda]$ of Λ lies in $X_{F,V}$.

The argument for V closed is similar. □

Proof of Proposition 14 (*Sketch*). Consider the case when V is open. Then $X_{F,V}$ is open and the value of any regular measure at $X_{F,V}$ can be approximated as closely as desired by a compact set $K \subset X_{F,V}$. For any such K we can find a continuous function f with $\mathbf{1}_K \le f \le \mathbf{1}_{X_{F,V}}$. Using Proposition 10, where all weights are assumed equal to 1, we obtain that $\dot{\mu}(f)$ is almost surely the average of f on Λ and, from the definition of f, that for any $\epsilon > 0$ and for large enough R,

$$\dot{\mu}(K) \le \dot{\mu}(f) \le \underline{\mathrm{freq}}(\Lambda, F, V) \le \frac{1}{\ell(C_R)} \sum_{x\in\Lambda\cap C_R} \mathbf{1}_{F,V}(-x+\Lambda) + \epsilon.$$

Integrating over X and using the Campbell formula we have, independently of R,

$$\dot{\mu}(K) \le \int_X \underline{\mathrm{freq}}(\Lambda, F, V)\mathrm{d}\mu \le \dot{\mu}(X_{F,V}).$$

Now since we can make $\dot{\mu}(K)$ as close as we wish to $\dot{\mu}(X_{F,V})$, we obtain both

$$\dot{\mu}(X_{F,V}) \le \underline{\mathrm{freq}}(\Lambda, F, V) \quad \text{and} \quad \int_X \underline{\mathrm{freq}}(\Lambda, F, V)\mathrm{d}\mu = \dot{\mu}(X_{F,V})$$

From this $\dot{\mu}(X_{F,V}) = \underline{\mathrm{freq}}(\Lambda, F, V)$, μ-almost everywhere.

The result for V closed is similar, this time approximating by open sets from above. □

Example. Consider the usual dynamical system based on $\mathbb{Z}$: $X(\mathbb{Z}) \simeq \mathbb{R}/\mathbb{Z}$. Let $F := \{0, 1/4\}$ and $V := (-1/4, 1/4)$. For any $\Lambda = t + \mathbb{Z}$, we have

$$\frac{1}{2n} \sum_{u\in(t+\mathbb{Z})\cap[-n,n]} \mathbf{1}_{F,V}(-u+t+\mathbb{Z}) = 0$$

while

$$\frac{1}{2n}\sum_{u\in(t+\mathbb{Z})\cap[-n,n]}\mathbf{1}_{F,\overline{V}}(-u+t+\mathbb{Z})\approx 1.$$

In this case we have $X_{F,V}=\emptyset$, $X_{F,\overline{V}}=X$ and

$$0=\mathrm{freq}(\Lambda,F,V)=\dot{\mu}(X_{F,V})<\dot{\mu}(X_{F,\overline{V}})=\mathrm{freq}(\Lambda,F,\overline{V})=1.$$

12. Final comments

After Theorem 5, it is natural to ask whether or not for an m-coloured stationary ergodic uniformly discrete point process (X,μ) *all* the correlations of μ are necessary in order to determine it. Is it possible that only a finite number of them will suffice? In [8] it is shown that given any $n\geq 2$ there are 1D examples based on multi-step Markov processes for which only the $2,3,\ldots,n$-point correlations are required to determine μ.

In [21] the pure point case is studied and Theorem 5 is used to relate the correlations to the extinctions (missing Bragg peaks) in the diffraction of (X,μ). For example, it is shown in the one-colour case that if there are no extinctions then the two- and three-point correlations determine the measure μ. This seems to be the generic situation for regular model sets based on real internal spaces.

Acknowledgments

The authors would like to thank Michael Baake and Daniel Lenz for their interest and their insightful comments on this work during its preparation. We also thank Haik Mashurian and Jeong-Yup Lee for their careful reading of the manuscript. Their contributions have substantially improved the paper. RVM gratefully acknowledges the support of this research by the Natural Sciences and Engineering Research Council of Canada.

References

[1] M. Baake, M. Birkner, R.V. Moody, Diffraction of stochastic point sets: exactly solvable models, preprint, 2007.
[2] M. Baake, D. Lenz, Dynamical systems on translation bounded measures: Pure point dynamical and diffraction spectra, Ergodic Theory & Dynam. Systems 24 (2004) 1867–1893. math. DS/0302231.
[3] M. Baake, R.V. Moody, Weighted Dirac combs with pure point diffraction, J. Reine Angew. Math. (Crelle) 573 (2004) 61–94. math.MG/0203030.
[4] C. Berg, G. Forst, Potential Theory on Locally Compact Abelian Groups, Springer, Berlin, 1975.
[5] Donald L. Cohn, Measure Theory, Birkhäuser, Boston, 1997.
[6] D.J. Daley, D. Vere-Jones, An Introduction to the Theory of Point Processes, in: Springer Series in Statistics, Springer, 1988; 2nd edition, vol. 1, 2002.
[7] F.M. Dekking, The spectrum of dynamical systems arising from substitutions of constant length, Z. Wahrscheinlichkeitstheorie verw. Geb. 41 (1978) 221–239.
[8] Xinghua Deng, How finite correlations determine uniformly discrete point processes: Examples (in preparation).
[9] J. Dieudonné, Treatise on Analysis, Vol. 2, Academic Press, London, 1970.
[10] S. Dworkin, Spectral theory and X-ray diffraction, J. Math. Phys. 34 (1993) 2965–2967.
[11] J.-B. Gouéré, Quasicrystals and almost periodicity, Comm. Math. Phys. 255 (3) (2005) 651–681. math-ph/0212012.
[12] A. Hof, On diffraction by aperiodic structures, Comm. Math. Phys. 169 (1995) 25–43.
[13] A. Hof, Diffraction by aperiodic structures, in: R.V. Moody (Ed.), The Mathematics of Long-Range Aperiodic Order, in: NATO-ASI Series C, vol. 489, Kluwer, Dordrecht, 1997, pp. 239–268.
[14] A. Hof, Diffraction by aperiodic structures at high temperatures, J. Phys. A: Math. Gen. 28 (1995) 57–62.
[15] S. Kakutani, Ergodic theory of shift transformations, in: Proc. Fifth Berkeley Sympos. Math Statist. Probability II, 1967, pp. 405–417.
[16] S. Kakutani, Strictly Ergodic Symbolic Dynamical Systems, in: Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 2, University of California Press, Berkeley, Los Angeles, 1972, pp. 319–326.
[17] A.F. Karr, Point Processes and their Statistical Inference, Marcel Dekker Inc., New York, Basel, Hong Kong, 1991.
[18] M. Keane, Generalized Morse sequences, Z. Wahrscheinlichkeitstheorie verw. Geb. 10 (1968) 335–353.
[19] J.-Y. Lee, R.V. Moody, B. Solomyak, Pure point dynamical and diffraction spectra, Ann. H. Poincaré 3 (2002) 1003–1018.
[20] D. Lenz, On a conjecture by Bombieri and Taylor, preprint, 2007.
[21] D. Lenz, R.V. Moody, Extinctions and correlations for uniformly discrete point processes with pure point dynamical spectra (in preparation).
[22] J.C. Martin, Minimal flows arising from substitutions of non-constant length, Math. Syst. Th. 7 (1973) 73–82.

[23] R.V. Moody, Model sets and their duals, in: R.V. Moody (Ed.), The Mathematics of Long-Range Aperiodic Order, in: NATO-ASI Series C, vol 489, Kluwer, Dordrecht, 1997, pp. 239–268.
[24] N. Priebe–Frank, Substitution sequences in Z^d with a nonsimple Lebesgue component in the spectrum, Ergodie Theory & Dynam. Systems 23 (2) (2003) 519–532.
[25] M. Queffélec, Substitution Dynamical Systems – Spectral Analysis, in: Lecture Notes in Mathematics, vol. 1294, Springer, 1987.
[26] C. Radin, M. Wolf, Geom. Dedicata 42 (1992) 355–360.
[27] W. Rudin, Fourier Analysis on Groups, Wiley, New York, 1962. Reprint (1990).
[28] M. Schlottmann, Generalized model sets and dynamical systems, in: M. Baake, R.V. Moody (Eds.), Directions in Mathematical Quasicrystals, in: CRM Monograph Series, vol. 13, AMS, Providence, RI, 2000, pp. 143–159.
[29] L. Schwartz, Théorie des Distributions, Hermann, Paris, 1966.
[30] A.C.D. van Enter, J. Międkisz, How should one define a (weak) crystal? J. Stat. Phys. 66 (1992).