RECONSTRUCTION OF PERIODIC SCENERIES SEEN ALONG A RANDOM WALK WITH JUMPS

Heinrich Matzinger, Jüri Lember[1]

Abstract. This article was motivated by a question of Kesten. Kesten asked for which random walks periodic sceneries can be reconstructed. Among others, he asked the question for random walks which at each step can move by one or two units to the right. Previously, Howard [How97] proved that all periodic sceneries can be reconstructed provided they are observed along the path of a simple random walk.

We prove that for a large class of random walks it is possible to reconstruct periodic sceneries. Among others, we provide a positive answer in the case, when the random walk can only move by one or two units to the right at each step, with a two unit step being less likely, than a one unit step.

1 Introduction

A (2-color) scenery ξ is a coloring of the integers $\mathbb{Z}$ with 2 colors. (Hence, a (2-color) scenery is a map $\xi : \mathbb{Z} \to \{0, 1\}$.) Two sceneries ξ, ξ' are called equivalent, $\xi \approx \xi'$, if one of them is obtained from the other by a translation or reflection. Let $(S(t))_{t\geq 0}$ be a recurrent random walk on the integers. Observing the scenery ξ along the path of this random walk, one sees the color $\xi(S(t))$ at time t. The *scenery reconstruction problem* is to determine the scenery ξ if we are only given one realization of the observations $\chi := (\xi(S(t)))_{t\geq 0}$. In general, reconstructing sceneries works at best up to equivalence. For an overview on scenery reconstruction and related problems, we refer the reader to the excellent review paper [Kes98] by Kesten.

The reconstructability strongly depends on the properties of the random walk. In [How97], Howard proves that all periodic sceneries can be reconstructed up to equivalence if the random walk is a simple random walk. Since, in general, a scenery can be reconstructed only up to the equivalence, the starting point of the random walk plays no role.

Let us briefly describe the origins of the scenery reconstruction problem. One could describe the general area as the study of the ergodic properties of the observations χ. Among others, the ergodic properties of observations were studied by den Hollander and Keane in [KdH86], Heicklen and Hoffman and Rudolph in [HHR00]. In [dH88], den Hollander studies the mixing properties of observations and in [dHS97], den Hollander and Steif study the mixing property of the generalized T, T^{-1}-process. The T, T^{-1} problem motivated den Hollander, Keane and Benjamini to ask if any pair of non-equivalent sceneries could be distinguished. The scenery distinguishing problem can be described as follows:

Given a set of two sceneries $\{\xi_1, \xi_2\}$ (known to us), is it possible from one realization of χ alone to determine a.s. if the observations where produced on ξ_1 or ξ_2? If the answer is yes, we say that ξ_1 and ξ_2 are distinguishable.

Lindenstrauss [Lin99] showed that there exist pairs of non-equivalent sceneries which can not be distinguished. On the other hand, most typical pairs are distinguishable, as was proven by

---

[1] Author is supported by Estonian Science Foundation Grant nr. 5694

Kesten and Benjamini [BK96]. They show that almost all pairs of independent 2-color sceneries can be distinguished. For this they assume the sceneries to be generated by a random process. (The sceneries are taken i.i.d.).

A related problem is the distinguishing of sceneries which differ only in one point. This problem is called the "detection of single defects problem". Kesten [Kes96] showed that single defects can be detected in almost all 5-color sceneries. Previously, Howard [How96] had shown that single defects can always be detected in periodic sceneries observed along a simple random walk.

The scenery reconstruction techniques differ very much when the sceneries are periodic. The main result of this paper is that for a large class of random walks it is possible to reconstruct periodic sceneries. Howard [How97] had already proved that all periodic sceneries observed along a simple random walk path, can be reconstructed. This lead Kesten to ask what happens when the random walk is not simple.

A problem closely related to the reconstruction of periodic sceneries, is the reconstruction of sceneries with a finite number of ones. This problem was solved in a more general form by Levin and Peres [LP02]. These authors proved that *stochastic sceneries* can be reconstructed. A stochastic scenery is a map $\xi : \mathbb{Z} \to I$, where I denotes a set of distributions. At time t, one observes the random variable $\chi(t)$, drawn according to the distribution $\xi(S_t) \in I$. Given S and ξ, the observations $\chi(t)$ for different t's are independent of each other.

The "finite stochastic scenery problem" is a generalization of the Keane-Harris coin tossing problem. In this problem, one considers a binary sequence χ which can be generated in two manners:

1. Throw an unbiased coin independently. Hence $\chi = (\chi(0), \chi(1), ...)$ is a sequence of i.i.d. Bernoulli variables with parameter 0.5.
2. Modify the observations χ in the following way: at each time $t \in T$ replace the fair coin by a coin with bias $\theta > 0$. The set T is a a random set of renewal times which is not observable.

The Harris-Keane coin tossing problem is to determine from one set of observations alone if we are in case 1 or in case 2. Harris and Keane [HK97] showed the existence of a critical phenomena: Depending on the finiteness of the first moment of the inter-arrival times, one can a.s. distinguish the two cases or not. This problem was carried further by Peres, Pemantle and Levin [LPP01].

2 Notations and assumptions

We define the main concepts used in this paper: scenery, periodic scenery, random walk and observations.

- A *scenery* is a map $\xi : \mathbb{Z} \to \{0,1\}$. We will also view sceneries as elements of $\{0,1\}^{\mathbb{Z}}$.
- Two sceneries $\xi, \xi' \in \{0,1\}^{\mathbb{Z}}$ are called *equivalent* if ξ is obtained by some translation and reflection of ξ'. This means that for some $a \in \{-1,+1\}$, $b \in \mathbb{Z}$, we have:

  $$\xi(z) = \xi'(az+b), \quad \forall z \in \mathbb{Z}.$$

  When ξ and ξ' are equivalent, we write: $\xi \approx \xi'$.

- Let PER_l denote the set of all sceneries with period l, i.e.

  $$\mathrm{PER}_l = \{\xi \in \{0,1\}^{\mathbb{Z}} \mid \xi(z) = \xi(z+l), \quad \forall z \in \mathbb{Z}\}.$$

  Let PER be the set of all periodic sceneries, i.e. $\mathrm{PER} := \cup_{l=1}^{\infty} \mathrm{PER}_l$.

- We shall denote by $\mathbb{Z}_l$ the quotient ring $\mathbb{Z}/l\mathbb{Z}$ and we identify $\mathbb{Z}_l$ with the set $\{0,\dots,l-1\}$. Every scenery $\xi \in \mathrm{PER}_l$ is determined by its values on $\{0,\dots,l-1\}$. We denote the vector made of these values by ξ_l, i.e.

  $$\xi_l := (\xi(0),\dots,\xi(l-1)).$$

- In this paper, $S = \{S(t)\}_{t\in\mathbb{N}}$ is a random walk with initial distribution π. Let $p = (p(z))_{z\in\mathbb{Z}}$ be the distribution of the increments of the random walk S, that is

  $$p(z) := \mathbf{P}(S(1) - S(0) = z).$$

  Throughout this paper, we assume the distribution p to be known.

- Let S_l be the random walk on $\mathbb{Z}_l$ induced by the random walk S:

  $$S_l(t) := \ (S(t) \mod l)$$

  for all $t \in \mathbb{N}$. If p_l designates the distribution of the increments of S_l, then we have that:

  $$p_l(i) := \sum_{j:j \bmod l = i} p(j).$$

  We view p_l as a l-dimensional vector:

  $$p_l := (p_l(0),\dots,p_l(l-1)).$$

  Clearly, S_l is a stationary Markov Chain with state space $\mathbb{Z}_l$. The initial distribution of S_l depends on π. The stationary distribution is uniform.

- We denote by χ the *observations* :

  $$\chi := (\xi\left(S\left(0\right)\right), \xi\left(S\left(1\right)\right), \xi\left(S\left(2\right)\right), \dots).$$

Let us formalize the periodic scenery reconstruction problem treated in this paper:

Definition 2.1 *Let S be a random walk with given initial distribution π and given distribution of increments p. We say that periodic sceneries can be reconstructed when observed along the path of S, if there exists a mapping depending on the distribution of S*

$$\mathcal{A} : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{Z}},$$

such that for all $\xi \in \mathrm{PER}$,

$$\mathbf{P}\big(\mathcal{A}(\xi \circ S) \approx \xi\big) = 1.$$

In this paper, we describe classes of random walks for which the reconstruction of periodic sceneries is possible.

3 Reconstruction of a periodic scenery with known period

3.1 The D-function

Throughout this section, we fix l and a scenery $\xi \in \mathrm{PER}_l$. As usual, we identify $\mathbb{Z}_l$ with $\{0, \dots, l-1\}$. Let $\mathbb{Z}_l^l$ denote the set of l-tuples of elements of $\mathbb{Z}_l$:

$$\mathbb{Z}_l^l := \{ \ (z_1, \dots, z_l) \ \mid \ z_1, \dots, z_l \in \mathbb{Z}_l \ \}.$$

For every $j \in \mathbb{Z}_l$, we define the function

$$D_j : \ \mathbb{Z}_l^l \to \{0, 1\}, \ \ (z_1, \dots, z_l) \mapsto D_j(z_1, \dots, z_l),$$

where

$$D_j(z_1, \dots, z_l) := \xi(j) \cdot \prod_{k=1}^{l} \xi(j + \sum_{i=1}^{k} z_i).$$

With this definition, $D_j(z_1, \dots, z_l) = 1$ if and only if

$$\xi(j) = \xi(j + v_1) = \xi(j + v_2) = \cdots = \xi(j + v_l) = 1,$$

where

$$v_k = \sum_{i=1}^{k} z_i.$$

Otherwise, $D_j(z_1, \dots, z_l) = 0$.
We define the *D-function of* ξ:

$$D : \mathbb{Z}_l^l \to [0, 1], \quad D := \frac{1}{l} \sum_{i=1}^{l} D_i.$$

Note that D depends on ξ. As we prove in the next lemma, the D-function uniquely determines ξ up to equivalence (see also the numerical example right after the proof of the lemma).

Lemma 3.1 *Let $\xi, \xi' \in \mathrm{PER}_l$ with the corresponding D-functions D and D'. If $\xi \not\approx \xi'$, then $D \neq D'$.*

Proof. Let $\prec$ be the lexicographic ordering of the set $\{1, \dots, l\}^l$, i.e. for every pair $z = (z_1, \dots, z_l), y = (y_1, \dots, y_l) \in \{1, \dots, l\}^l$, $z \prec y$ if and only if there exists $k \in \{1, \dots, l\}$ so that $z_i = y_i$ $i = 1, \dots, k-1$ and $z_k < y_k$. We now use the ordering $\prec$ in the set $\{0, \dots, l-1\}^l$, where 0 is identified with l. Formally, we define a bijection

$$T : \{0, \dots, l-1\} \to \{1, \dots, l\}, \quad T0 = l, Ti = i, \quad i \neq 0$$

and we define the ordering $\prec$ in $\{0, \dots, l-1\}^l$ by $(x_1, \dots, x_l) \prec (y_1, \dots, y_l)$ if and only if

$$(Tx_1, \dots, Tx_l) \prec (Ty_1, \dots, Ty_l).$$

Let $\xi \in \mathrm{PER}_l$ and let D be the D-function of ξ. Let

$$V = \{z \in \mathbb{Z}_l^l : D(z) > 0\}.$$

Let $\bar{z} = (\bar{z}_1, \ldots, \bar{z}_l)$ be the minimal element of V corresponding to the ordering $\prec$. Let

$$\phi^o = \Big(\phi^o(0), \phi^o(1), \phi^o(2), \ldots, \phi^o(\bar{z}_1 + \cdots + \bar{z}_l)\Big) \in \{0,1\}^{\bar{z}_1 + \cdots + \bar{z}_l + 1}$$

be defined as follows:

$$\phi^o(i) = 1 \text{ if and only if } i \in \{0, \bar{z}_1, \bar{z}_1 + \bar{z}_2, \ldots, \bar{z}_1 + \cdots + \bar{z}_l\}. \tag{3.1}$$

The length of ϕ^o is at least $l+1$. Let $\phi \in \mathrm{PER}_l$ be a periodic sequence so that

$$\phi_l = \big(\phi^o(0), \ldots, \phi^o(l-1)\big).$$

So, D uniquely determines ϕ^o, and ϕ^o fully determines ϕ.
Finally note that $\phi \approx \xi$. By definition of $\bar{z}$, there exists $j \in \{0, \ldots, l-1\}$ so that

$$\xi(j) = \xi(j + v_1) = \xi(j + v_2) = \xi(j + v_3) = \cdots = \xi(j + v_l) = 1, \tag{3.2}$$

with $v_k = \sum_{i=1}^k \bar{z}_i$. Since $\bar{z}$ is the smallest element in V with respect to the lexicographic order, we have that

$$\xi(j) = 0, \quad \text{if } i \notin \{j, j + v_1, j + v_2, j + v_3, \ldots, j + v_l\}. \tag{3.3}$$

Indeed: If (3.3) fails, then there would be a $k \in \{0, \ldots, l\}$ and $m \in \{j, \ldots, j + v_l\}$ such that

$$j + v_k < m < j + v_{k+1}, \qquad \text{and } \xi(m) = 1.$$

Here $v_0 \equiv 0$. But in this case, there would be an $y \in V$ such that $y_1 = \bar{z}_1, y_2 = \bar{z}_2, \ldots, y_k = \bar{z}_k, y_{k+1} = m < \bar{z}_{k+1}$, implying that $y \prec \bar{z}$. This contradicts the assumption of $\bar{z}$ being the smallest in the lexicographic order. The relations (3.2) and (3.3) together with (3.1) imply that

$$\xi(j + i) = \phi^o(i), \quad i = 0, \ldots l.$$

By definition of ϕ, $\xi \approx \phi$. ■

Let us look at a numerical example:

Take $\xi \in \mathrm{PER}_5$:

$\xi(z)$	$\ldots$	0	0	1	0	1	0	0	1	0	1	0	$\ldots$
z	$\ldots$	0	1	2	3	4	5	6	7	8	9	10	$\ldots$

So $\xi_5 = (0,0,1,0,1)$. In this case there are two ones per period, located in the points 2 and 4.
Let us walk on $\mathbb{Z}$ in the following way: Start at the point 2, from there move two units to the right, then three units to the right before moving two units to the right again .... When we follow this path we only visit points z where $\xi(z) = 1$. Hence $D_2(2,3,2,3,2) = 1$ and it follows that $D(2,3,2,3,2) = \frac{1}{5} > 0$.
Another possibility for observing only ones is the following: Start at the point 4, then move three units to the right, then two units, then three again.... This yields $D_4(3,2,3,2,3) = 1$ and $D(3,2,3,2,3) = \frac{1}{5} > 0$.
Of course, a possibility to observe only the ones is to start from 2, move 2 and then not move at all. This yields $D_2(2,0,0,0,0) = 1$.
Since 0 is identified with l, the lexicographical smallest element for which D is different from zero, is

$(2,3,2,3,2)$. If we are only given D, we can use this to reconstruct ξ up to equivalence: since $D(2,3,2,3,2) > 0$, we must have a one in ξ followed by another one two units to the right, from where we have another one located three units to the right... This yields the following reconstruction:

$$\begin{array}{ccccccccccccc} \ldots & 1 & . & 1 & . & . & 1 & . & 1 & . & . & 1 & \ldots \\ \hline \ldots & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & \ldots \end{array}$$

of locations of ones (up to shift and translation). Since $(2,3,2,3,2)$ is the lexicographically smallest element on which $D > 0$, it follows that between the reconstructed positions of ones, there can only be zeros. Hence we obtain as a reconstruction (up to equivalence) for the scenery:

$$\begin{array}{ccccccccccccc} \ldots & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & \ldots \\ \hline \ldots & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & \ldots \end{array}$$

We implicitly assumed the period of the scenery ξ to be known .

3.2 Reconstructing the D-function

Lemma 3.1 shows that from the D-function of ξ, one can easily reconstruct ξ: Take the lexicographically minimal element x of V and define ϕ^o as in (3.1). In the present subsection, we prove that under some conditions on the distribution of the increments of the random walk, one can reconstruct the D-function of ξ from the observations.
For this, we view D as a l^l-dimensional vector. The entries of this vector are indexed by the set $\mathbb{Z}_l^l$. The entry corresponding to $(z_1,\ldots,z_l) \in \mathbb{Z}_l^l$ is $D(z_1,\ldots,z_l)$, so

$$D := \left(D(z_1,\ldots,z_l)\right)_{(z_1,\ldots,z_l)\in\mathbb{Z}_l^l}.$$

Recall that $p_l \in \mathbb{R}^l$ is the distribution of the increments of S_l. We view p_l as the vector of $\mathbb{R}^l$

$$p_l := (p_l(0),\ldots,p_l(l-1)).$$

Let p_l^{*k} denote the distribution p_l convoluted k times with itself on $\mathbb{Z}_l$. We also view p_l^{*k} as a vector of $\mathbb{R}^l$

$$p_l^{*k} = (p_l^{*k}(0),\ldots,p_l^{*k}(l-1)).$$

We will use direct products of l distributions from the set

$$\left\{p_l^{*k} | k \in \mathbb{N}\right\}.$$

(Probabilist call "direct product" what in other areas of mathematics is called "tensor product"). A direct product of l such distributions, can be viewed as a vector with l^l entries indexed by the set $\mathbb{Z}_l^l$. In this way, the entry corresponding to $(z_1,\ldots,z_l)$ of the "vector" $p_l^{*t_1}\otimes\ldots\otimes p_l^{*t_l}$ equals

$$p_l^{*t_1}\otimes\ldots\otimes p_l^{*t_l}(z_1,\ldots,z_l) := p_1^{*t_1}(z_1)\cdot\ldots\cdot p_l^{*t_l}(z_l).$$

We can now form the scalar product between the vector D and the vector $p_l^{*t_1}\otimes\ldots\otimes p_l^{*t_l}$:

$$D \times \ (p_l^{*t_1}\otimes\ldots\otimes p_l^{*t_l}) := \sum_{(z_1,\ldots,z_l)\in\mathbb{Z}_l^l} D(z_1,\ldots,z_l)\cdot\left(p_l^{*t_1}\otimes\ldots\otimes p_l^{*t_l}(z_1,\ldots,z_l)\right).$$

The next lemma shows that there is a simple probabilistic interpretation to this scalar product. It is equal to the probability that $\chi(t) = 1$ for every t in the set

$$T := \{0, t_1, t_1+t_2, \ldots, t_1+t_2+\ldots+t_l\}.$$

Lemma 3.2 *Let the initial distribution of S_l be uniform. Then*

$$D \times \ (p_l^{*t_1} \otimes \ldots \otimes p_l^{*t_l}) = q(t_1, \ldots, t_l),$$

where

$$q(t_1, \ldots, t_l) := \ \mathbf{P}\,(\ \chi(t) = 1\ ,\ \forall t \in T\ )\,.$$

Proof. Let A be the event

$$A := \{\chi(t) = 1\ ,\ \forall t \in T\ \}\,.$$

Let, for every $i \in \mathbb{Z}_l$ and $(z_1, \ldots, z_l) \in \mathbb{Z}_l^l$, $B(i, z_1, \ldots, z_l)$ be the event

$$B(i, z_1, \ldots, z_l) := \{S_l(0) = i\}\ \cap\ \left(\bigcap_{j=1}^{l} \{S_l(t_j + s_{j-1}) - S_l(s_{j-1}) = z_j\} \right),$$

where $s_j := t_1 + t_2 + \ldots + t_j$ for all $j = 1, 2, \ldots, l$ and $s_0 := 0$.
By law of total probability,

$$\mathbf{P}(A) = \sum_{i, z_1, \ldots, z_l \in \mathbb{Z}_l} \mathbf{P}\,(A \mid B(i, z_1, \ldots, z_l))\ \cdot\ \mathbf{P}(B(i, z_1, \ldots, z_l)). \tag{3.4}$$

We assume that the random walk S_l starts in its stationary distribution. Hence

$$\mathbf{P}(S_l(0) = i) = \frac{1}{l}.$$

Using this and the fact that random walk has independent increments,

$$\mathbf{P}(B(i, z_1, \ldots, z_l)) = \frac{1}{l} \cdot \prod_{j=1}^{l} \mathbf{P}(S_l(t_j + s_{j-1}) - S_l(s_{j-1}) = z_j). \tag{3.5}$$

Hence,

$$\mathbf{P}(B(i, z_1, \ldots, z_l)) = \frac{1}{l} \cdot \prod_{j=1}^{l}\ p_l^{*t_j}(z_j). \tag{3.6}$$

Conditional on $B(i, z_1, \ldots, z_l)$, the probability of the event A is either zero or one. If the scenery ξ is equal to one on all the points of the set

$$\{i, i + z_1, i + z_1 + z_2, \ldots, i + z_1 + \ldots + z_l\},$$

then that conditional probability is one, otherwise it is zero. Hence,

$$\mathbf{P}(A | B(i, z_1, \ldots, z_l)) = D_i(z_1, \ldots, z_l). \tag{3.7}$$

Together, (3.4), (3.6) and (3.7) imply that

$$\mathbf{P}(A) = \sum_{i, z_1, \ldots, z_l \in \mathbb{Z}_l} \frac{1}{l} D_i(z_1, \ldots, z_l)\ \left(\prod_{j=1}^{l} p_l^{*t_j}(z_j) \right). \tag{3.8}$$

Since $D = \frac{1}{l}\sum_{i\in\mathbb{Z}_l} D_i$, (3.8) yields:

$$\mathbf{P}(A) = \sum_{z_1,\ldots,z_l\in\mathbb{Z}_l} D(z_1,\ldots,z_l)\left(\prod_{j=1}^{l} p_l^{*t_j}(z_j)\right). \tag{3.9}$$

Equivalently, the last equation can be written as

$$q(t_1,\ldots,t_l) = \mathbf{P}(A) = \sum_{z_1,\ldots,z_l\in\mathbb{Z}_l} D(z_1,\ldots,z_l)\cdot\left(p_l^{*t_1}\otimes\ldots\otimes p_l^{*t_1}(z_1,\ldots,z_j)\right) = D\times(p_l^{*t_1}\otimes\ldots\otimes p_l^{*t_l}).$$

■

Next we show that periodic sceneries can be reconstructed when we know their period l. For this, we also assume that the convolutions p_l^{*k} for $k = 1,\ldots,l$ are linearly independent in $\mathbb{R}^l$.

Theorem 3.1 *Let $\xi \in \mathrm{PER}_l$. Suppose the vectors $p_l, p_l^{*2},\ldots,p_l^{*l}$ are linearly independent. Then there exists a map*

$$\mathcal{D}_l : \{0,1\}^{\mathbb{N}} \to [0,1]^{l^l},$$

such that $\mathcal{D}_l(\xi\circ S)$ is the D-function of ξ, a.s.. In other words, one realization of the observations χ a.s. determines D.

Proof. The tensor product of a basis is again a basis. Since the distributions $p_l, p_l^{*2},\ldots,p_l^{*l}$ are linearly independent of each other, they form a basis of $\mathbb{R}^l$. It follows, that

$$\left\{p_l^{*t_1}\otimes\ldots\otimes p_l^{*t_l} \mid t_1,\ldots,t_l \in \mathbb{Z}^l\right\} \tag{3.10}$$

forms a basis of $\mathbb{R}^{(l^l)}$.
Knowing the scalar product of a vector with each element of a basis uniquely determines that vector. Apply this to the vector $D \in \mathbb{R}^{l^l}$ and to the basis (3.10). According to Lemma 3.2, the scalar products are then given by the numbers $q(t_1,\ldots,t_l)$, where $(t_1,\ldots,t_l) \in \mathbb{Z}_l^l$. Hence, the set of scalars

$$\left\{q(t_1,\ldots,t_l) \mid (t_1,\ldots,t_l) \in \mathbb{Z}_l^l\right\}$$

uniquely determines the vector D. Recall the set T from Lemma 3.2. Let

$$U_t(t_1,\ldots,t_l) = \begin{cases} 1 & \text{if } \chi(s) = 1,\ s \in t+T, \\ 0 & \text{otherwise.} \end{cases}$$

Recall that S_l is a random walk on $\mathbb{Z}_l$. Since S_l is finite irreducible Markov Chain with uniform stationary distribution, by ergodic theorem,

$$q(t_1,\ldots,t_l) = \lim_{n\to\infty}\frac{\sum_{t=0}^{n} U_t(t_1,\ldots,t_l)}{n}.$$

■

To summarize: One realization of χ a.s. uniquely determines the coefficients $q(t_1,\ldots,t_l)$. These coefficients in turn, uniquely determine D.

3.3 Discrete Fourier Transform

Let $\mathcal{F} : L_2(\mathbb{Z}_l, \mathbb{R}) \to L_2(\mathbb{Z}_l, \mathbb{C})$ be the discrete Fourier transform, i.e. for $f \in L_2(\mathbb{Z}_l, \mathbb{R})$

$$\hat{f}(u) := \mathcal{F}f(u) = \sum_{y \in \mathbb{Z}_l} f(y) \exp[-\frac{(2\pi i u)y}{l}].$$

Clearly $\mathcal{F}$ is a linear map. Since $L_2(\mathbb{Z}_l, \mathbb{C})$ is isomorphic with $\mathbb{C}^l$, the mapping $\mathcal{F}$ is a complex-valued $l \times l$-matrix. Since $\mathcal{F}$ is also 1-1 (see, e.g.[Ter99]), the matrix has full rank. Hence, to prove that vectors $f_1, \ldots, f_l \in \mathbb{C}^l$ are linearly independent, it suffices to show that the corresponding Fourier transforms $\hat{f}_1, \ldots, \hat{f}_l$ are linearly independent. This is the content of the next lemma.

Let $\hat{p}_l$ designate the Fourier transform of p_l. Thus:

$$\hat{p}_l(u) := \sum_{y=0}^{l-1} p_l(y) \exp[-\frac{(2\pi i u)y}{l}] = \sum_{y \in \mathbb{Z}} p(y) \exp[-\frac{(2\pi i u)(y \mod l)}{l}].$$

Hence,

$$\big(\hat{p}_l(0), \ldots, \hat{p}_l(l-1)\big) = \mathcal{F}p_l.$$

Lemma 3.3 *If* $\forall u, v \in \mathbb{Z}_l$, *with* $u \neq v$, *we have:*

$$\hat{p}_l(u) \neq \hat{p}_l(v), \tag{3.11}$$

then the vectors $p_l, p_l^{*2}, \ldots, p_l^{*l}$ *are linearly independent of each other.*

Proof. To simplify notation in this proof, we write $\hat{p}(u)$ instead of $\hat{p}_l(u)$. Also, we write $\hat{p}^j(u)$ for $(\hat{p}_l(u))^j$. Since the Fourier transform is a 1-1 linear map, the vectors $p_l, p_l^{*2}, \ldots, p_l^{*l}$ are linearly independent if and only if the corresponding Fourier transforms $\mathcal{F}p, \mathcal{F}p_l^{*2}, \ldots, \mathcal{F}p_l^{*l}$ are linearly independent. The Fourier transformation of a convolution is the pointwise product of the Fourier transformations:

$$\mathcal{F}p_l^{*j} = (\hat{p}^j(0), \hat{p}^j(1), \ldots, \hat{p}^j(l-1)).$$

Hence, $p_l, p_l^{*2}, \ldots, p_l^{*l}$ are linearly independent if the vectors

$$\big(\hat{p}(0), \ldots, \hat{p}(l-1)\big), \big(\hat{p}^2(0), \ldots, \hat{p}^2(l-1)\big), \ldots, \big(\hat{p}^l(0), \ldots, \hat{p}^l(l)\big)$$

are linearly independent. These vectors are linearly independent if the vectors

$$(1, \ldots, 1), \big(\hat{p}(0), \ldots, \hat{p}(l-1)\big), \big(\hat{p}^2(0), \ldots, \hat{p}^2(l-1)\big), \ldots, \big(\hat{p}^{l-1}(0), \ldots, \hat{p}^{l-1}(l-1)\big). \tag{3.12}$$

are linearly independent. The vectors (3.12) are linearly independent, if the Vandermonde determinant

$$\begin{vmatrix} 1 & \hat{p}(0) & \hat{p}^2(0) & \cdots & \hat{p}^{l-1}(0) \\ 1 & \hat{p}(1) & \hat{p}^2(1) & \cdots & \hat{p}^{l-1}(1) \\ 1 & \hat{p}(2) & \hat{p}^2(2) & \cdots & \hat{p}^{l-1}(2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \hat{p}(l-1) & \hat{p}^2(l-1) & \cdots & \hat{p}^{l-1}(l-1) \end{vmatrix} = \prod_{u<v} (\hat{p}(v) - \hat{p}(u)). \tag{3.13}$$

is different from zero. This finishes the proof. ■

The next theorem says that all periodic sceneries of known period l can be reconstructed up to equivalence if condition (3.11) holds.

Theorem 3.2 *Let $\xi \in \mathrm{PER}_l$. Assume p to be such that $\hat{p}_l(u) \neq \hat{p}_l(v)$ for all $u, v \in \mathbb{Z}_l$ with $u \neq v$. Then, there exists a map*

$$\mathcal{A}_l : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{Z}},$$

such that for all $\xi \in \mathrm{PER}_l$,

$$\mathbf{P}\big(\mathcal{A}_l(\xi \circ S) \approx \xi\big) = 1.$$

Proof. By Lemma 3.3 and Theorem 3.1, there exists a map $\mathcal{D}_l$, depending on l, such that $\mathcal{D}_l(\xi \circ S)$ is the D-function of ξ. By Lemma 3.1, the D-function of ξ uniquely determines ξ up to equivalence. ■

3.4 Counterexamples

Theorem 3.2 states that for a large class of random walks, every periodic scenery $\xi \in \mathrm{PER}_l$ can be reconstructed. However, there exist random walks that do not allow any reconstruction. Let us present an example: Let S be such that

$$p(i) = \frac{1}{l}, \quad i = 0, \dots, l-1.$$

Then, $p_l(i) = p(i)$ and for any two $\xi_1, \xi_2 \in \mathrm{PER}_l$, the observations

$$\xi^1(S(1)), \xi^1(S(2)), \xi^1(S(3)), \dots \text{ and } \xi^2(S(1)), \xi^2(S(2)), \xi^2(S(3)), \dots$$

have the same distribution, provided ξ^1 and ξ^2 have the same number of ones. In fact, $\chi(0), \chi(1), \dots$ is then an i.i.d. sequence. In this case, it is not possible to reconstruct the observed scenery. Clearly, for such a distribution, condition (3.11) is not fulfilled:

$$\hat{p}(1) = \frac{1}{l}\Big(\sum_{y=0}^{l-1} \cos 2\pi \frac{y}{l} - i \sum_{y=0}^{l-1} \sin 2\pi \frac{y}{l}\Big) = \frac{1}{l}\Big(\sum_{y=0}^{l-1} \cos 2\pi \frac{y(l-1)}{l} - i \sum_{y=0}^{l-1} \sin 2\pi \frac{y(l-1)}{l}\Big) = \hat{p}(l-1).$$

3.5 Random walk with 2 steps

In the present subsection, we consider a random walk S with distribution p such that $p(1)+p(2) = 1$. (Such a random walk can at each time only move by one unit or two to the right.) Kesten asked whether the reconstruction of a periodic scenery is possible when it is observed along such a random walk. The following corollary of Lemma 3.3 partly answers the question.

Corollary 3.1 *Let p be such that $p(1) + p(2) = 1$ and $p(1) > 2p(2)$. Then, for every l, the vectors $p_l, p_l^{*2}, \dots, p_l^{*l}$ are linearly independent.*

Proof. Let $l = 2$. In this case, $p_l(0) = p(2)$, $p_l(1) = p(1)$ and

$$\hat{p}(0) = p_l(0) + p_l(1) = 1, \quad \hat{p}(1) = p_l(0) + p_l(1)\exp[-\pi i] = p(2) + p(1)\exp[-\pi i] = p(2) - p(1).$$

Since $p(2) < p(1)$, the vectors $(p(2), p(1)), (1, p(2) - p(1))$ are independent.
Let $l > 2$. In this case, $p_l = p$ and

$$\hat{p}(u) = p(1)\exp[-\frac{2\pi iu}{l}] + p(2)\exp[-\frac{2\pi i2u}{l}].$$

Let $u, v \in \{0, \ldots, l-1\}$, $u \neq v$,

$$a := \exp[-\frac{2\pi iu}{l}], \quad b := \exp[-\frac{2\pi iv}{l}]. \tag{3.14}$$

With this notation, $\hat{p}(u) = p(1)a + p(2)a^2$, $\quad \hat{p}(v) = p(1)b + p(2)b^2$. If $\hat{p}(u) = \hat{p}(v)$, then

$$p(1)(a - b) = p(2)(b^2 - a^2) = p(2)(b - a)(b + a). \tag{3.15}$$

Since $a \neq b$, (3.15) implies $b + a = -\frac{p(1)}{p(2)}$ and

$$|a + b| = \frac{p(1)}{p(2)}. \tag{3.16}$$

But a, b are complex numbers with modulo 1. Hence $|a + b| \leq 2$. However, by assumption, $\frac{p(1)}{p(2)} > 2$. This contradicts (3.16). Therefore, $\hat{p}(u) \neq \hat{p}(v)$ and Lemma 3.3 finishes the proof. ■

3.5.1 Equal probabilities

Let us briefly analyze the case

$$p(1) = p(2) = 1/2.$$

Let $u, v \in \mathbb{Z}_l$, $u \neq v$, and let a, b be as in (3.14). Then $\hat{p}(u) = \hat{p}(v)$ is equivalent to $b + a = -1$. Since

$$a = \cos[2\pi\frac{u}{l}] - i\sin[2\pi\frac{u}{l}], \quad b = \cos[2\pi\frac{v}{l}] - i\sin[2\pi\frac{v}{l}],$$

the condition $a + b = -1$ implies that the following conditions both hold

$$\sin[2\pi\frac{u}{l}] = -\sin[2\pi\frac{v}{l}] \tag{3.17}$$

$$\cos[2\pi\frac{u}{l}] + \cos[2\pi\frac{v}{l}] = -1. \tag{3.18}$$

For (3.17) to hold, u, v must satisfy one of the following conditions:

$$\frac{2\pi v}{l} = 2\pi - \frac{2\pi u}{l} = 2\pi(\frac{l - u}{l}) \tag{3.19}$$

or, with $u < v$,

$$\frac{2\pi v}{l} = \Big(\pi + \frac{2\pi u}{l}\Big). \tag{3.20}$$

Suppose (3.20) holds. Then

$$\cos[2\pi\frac{v}{l}] = \cos[\pi + \frac{2\pi u}{l}] = -\cos[\frac{2\pi u}{l}],$$

and (3.18) does not hold.
Suppose (3.19) holds. Then (3.18) is

$$\cos[2\pi\frac{v}{l}] + \cos[2\pi\frac{u}{l}] = \cos[2\pi - \frac{2\pi u}{l}] + \cos[2\pi\frac{u}{l}] = 2\cos[\frac{2\pi u}{l}] = -1.$$

Hence

$$\cos[\frac{2\pi u}{l}] = -\frac{1}{2}. \tag{3.21}$$

Assuming $u < v$, (3.21) implies that

$$\frac{2\pi u}{l} = \frac{2\pi}{3} \text{ and } \frac{2\pi v}{l} = \frac{4\pi}{3}.$$

Example 3.1 Let $l = 6$. Then $\hat{p}(2) = \hat{p}(4)$ and the vectors $p_6, p_6^{*2}, \dots, p_6^{*6}$ are not linearly independent. Indeed, the following matrix is singular

$$\begin{pmatrix} p_6^{*6}(0) & p_6^{*5}(0) & \cdots & p_6(0) \\ p_6^{*6}(1) & p_6^{*5}(1) & \cdots & p_6(1) \\ \cdots & \cdots & \cdots & \cdots \\ p_6^{*6}(5) & p_6^{*5}(5) & \cdots & p_6(5) \end{pmatrix} = \frac{1}{64}\begin{pmatrix} 2 & 10 & 24 & 8 & 0 & 0 \\ 6 & 20 & 16 & 0 & 0 & 32 \\ 15 & 20 & 4 & 0 & 16 & 32 \\ 20 & 10 & 0 & 8 & 32 & 0 \\ 15 & 2 & 4 & 24 & 16 & 0 \\ 6 & 2 & 16 & 24 & 0 & 0 \end{pmatrix}.$$

However, every scenery $\xi \in \mathrm{PER}_6$ can still be reconstructed. (For this we assume that we are given the period.) The algorithm for reconstruction of $\xi \in \mathrm{PER}_6$ is the following:

- Let $r := \sum_{i=0}^{l-1} \xi(i)$ be the number of ones in ξ_l. Find $\lim_{n\to\infty} \frac{\sum_{t=0}^{n} \chi(t)}{n+1}$. This limes is equal to $\mathbf{P}(\chi_6(0) = 1) = \frac{r}{6}$ a.s. and the number of ones, r, can be determined.
- Determine the numbers $q(1)$ and $q(2)$. According to our notation:

  $$q(1) := P(\chi_6(0) = \chi_6(1) = 1), \quad q(2) := P(\chi_6(0) = \chi_6(2) = 1).$$

  In order to determine $q(1)$ and $q(2)$, use the fact that by the ergodic theorem, we have

  $$\lim_{n\to\infty} \frac{\sum_{t=0}^{n} U_t(1)}{n+1} = \mathbf{P}(\chi_6(1) = \chi_6(0) = 1), \quad \text{a.s.}$$
  $$\lim_{n\to\infty} \frac{\sum_{t=0}^{n} U_t(2)}{n+1} = \mathbf{P}(\chi_6(2) = \chi_6(0) = 1), \quad \text{a.s..}$$

- Let $r = 2$. There exists, up to equivalence, only three sceneries in PER_6 with $r = 2$. Hence, when $r = 2$, we have that ξ_6 is equivalent to ϕ_1, ϕ_2 or ϕ_3, where:

  $$\phi_1 := (1,1,0,0,0,0), \phi_2 := (1,0,1,0,0,0), \phi_3 := (1,0,0,1,0,0).$$

  Now take

  $$\xi_6 \approx \begin{cases} \phi_3 & \text{if } q(1) = 0, \\ \phi_1 & \text{if } q(1) > 0 \text{ and } q(2) = 0 \\ \phi_2 & \text{if } q(1) > 0 \text{ and } q(2) > 0 \end{cases}$$

- Let $r = 3$. There exist three different equivalence classes of sceneries in PER_6 for which $r = 3$:
$$\phi_4 := (1,0,1,0,1,0), \phi_5 := (1,1,0,1,0,0), \phi_6 := (1,1,1,0,0,0).$$
Now, we have that
$$\xi_6 \approx \begin{cases} \phi_5 & \text{if } q(1) = \frac{1}{6}, \\ \phi_4 & \text{if } q(1) = \frac{1}{4}, q(2) = \frac{1}{4} \\ \phi_6 & \text{if } q(1) = \frac{1}{4}, q(2) = \frac{1}{12}. \end{cases}$$
- Let $r = 4$. Then change the roles of 0 and 1, and use the rule for $r = 2$.

Hence, knowing the period $l = 6$ and observing one realization of χ, we can a.s. determine $\xi \in \mathrm{PER}_6$ up to equivalence. (For this we assumed that $p(1) = p(2) = 0.5$). Here, reconstruction is possible, although the assumptions of Theorem 3.2 do not hold. So, the assumptions of Theorem 3.2 are sufficient but not necessary for the reconstructability of a scenery $\xi \in \mathrm{PER}_l$.

4 Reconstruction Algorithm

4.1 Unknown period

Theorem 3.2 states that under condition (3.11), there exists a map $\mathcal{A}_l$ that reconstructs any periodic scenery $\xi \in \mathrm{PER}_l$. The algorithm $\mathcal{A}_l$ (as well as condition (3.11)) depends on l. Applying $\mathcal{A}_l$ to a scenery ξ that does not belong to PER_l might give a wrong result. Hence, Theorem 3.2 gives a sufficient condition for the reconstructability of a scenery with known period.

Our ultimate goal, however, is to prove that periodic sceneries can be reconstructed without knowledge of the period (at least in the cases when the distribution of S is well behaved). The following theorem asserts that it is possible to reconstruct periodic sceneries when condition (3.11) holds for every l. It is not assumed that the period of the scenery is known.

Theorem 4.1 *Suppose S is such that* (3.11) *holds for every $l \in \mathbb{N}$. Then, there exists a map $\mathcal{A} : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{Z}}$ such that for every $\xi \in \mathrm{PER}$,*
$$\mathbf{P}\big(\mathcal{A}(\xi \circ S) \approx \xi\big) = 1.$$

Proof. By Theorem 3.2, for every j there exists $\mathcal{A}_j$ that a.s. reconstructs every scenery with period j. That is, for every $\xi \in \mathrm{PER}_j$, $\mathbf{P}\big(\mathcal{A}_j(\xi \circ S) \approx \xi\big) = 1$. Let l be the period of the scenery ξ. For any $k = 1, 2, \ldots$, kl is also a period of ξ. Hence, with probability one,
$$\mathcal{A}_{kl}(\xi \circ S) \approx \xi, \quad \forall k. \tag{4.1}$$
So, there exists at least one natural number so that (4.1) holds. Let l' be a natural number as a candidate for the unknown period. Apply the algorithms $\mathcal{A}_{kl'}$, $k = 1, 2, \ldots$. If they all give the same scenery (up to the equivalence), then the solution is ξ, a.s. Indeed, by (4.1), $\mathcal{A}_{l'l}(\xi \circ S) \approx \xi$. If there exists a $k \geq 2$ so that $\mathcal{A}_{kl'}(\xi \circ S) \not\approx \mathcal{A}_{l'}(\xi \circ S)$, the number l' is not the period, and the next candidate should be taken. Formally, let
$$l(\chi) = \min\{l' : \mathcal{A}_{kl'}(\xi \circ S) \approx \mathcal{A}_{l'}(\xi \circ S), \quad \forall k = 1, 2, \ldots\}, \quad \mathcal{A}(\xi \circ S) := \mathcal{A}_{l(\chi)}(\xi \circ S).$$
■

4.2 The algorithm

Let us summarize how to reconstruct periodic sceneries when (3.11) holds for every $l \in \mathbb{N}$. For this we assume that we are only given the observations $\chi = (\chi(0), \chi(1), \ldots)$ and the distribution of the random walk S. We do not assume that the period of the scenery is known. When (3.11) holds for every $l \in \mathbb{N}$, our method allows to construct a.s. a scenery equivalent to ξ.

Algorithm 4.1 *1. Determine a natural number $l > 0$, such*

$$\mathcal{A}_{kl}(\chi) \approx \mathcal{A}_l(\chi), \quad \forall k = 2, 3, \ldots.$$

Here, $\mathcal{A}_j(\chi)$ denotes the scenery obtained by applying the reconstruction algorithm for period j to the observations χ.

2. Let the output $\mathcal{A}(\chi)$ of this reconstruction algorithm be

$$\mathcal{A}(\chi) := \mathcal{A}_l(\chi).$$

Let us next describe the reconstruction algorithm $\mathcal{A}_l$. This algorithm is used when the period l of the scenery ξ is known. Recall that $U_t(t_1, \ldots, t_l) = 1$ if and only if

$$\chi(t) = \chi(t + t_1) = \chi(t + t_1 + t_2) = \cdots = \chi(t + t_1 + t_2 + \ldots + t_l) = 1.$$

Otherwise, $U_t(x_1, \ldots, x_l) = 0$.

Algorithm 4.2 *1. For all $(t_1, \ldots, t_l) \in \mathbb{Z}_l^l$, determine the coefficients $q_l(t_1, \ldots, t_l)$ by*

$$q_l(t_1, \ldots, t_l) = \lim_{n \to \infty} \frac{\sum_{t=0}^{n} U_t(t_1, \ldots, t_l)}{n + 1}$$

2. Use the coefficients $q(t_1, \ldots, t_l)$ to determine D. To this end, apply the set of equalities

$$D \times (p_l^{*t_1} \otimes \ldots \otimes p_l^{*t_l}) = q(t_1, \ldots, t_l)$$

from lemma 3.2 and the fact that

$$\left\{ p_l^{*t_1} \otimes \ldots \otimes p_l^{*t_l} \mid (t_1, \ldots, t_l) \in \mathbb{Z}_l^l \right\}$$

is a basis of $\mathbb{R}^{l^l}$ (when condition (3.11) holds).

3. Determine the scenery ξ from D. For this take the lexicographical smallest (with 0 identified as l) element in $(z_1, \ldots, z_l) \in \{1, \ldots, l\}^l$, such that $D(z_1, \ldots, z_l) \neq 0$. Then, determine a l-periodic scenery $\hat{\xi}$, such that the sequence of distances between successive ones in $\hat{\xi}$ is

$$z_1, z_2, \ldots, z_{l-1}, z_l, z_1, z_2, \ldots.$$

The scenery $\hat{\xi}$ is our reconstructed scenery, i.e.

$$\mathcal{A}_l(\chi) := \hat{\xi}.$$

The algorithm we describe here is theoretical: It uses an infinite amount of observations. One could also build a practical algorithm by approximating the coefficients $q(t_1, \ldots, t_l)$, instead of calculating them exactly. For this, we would rely on a finite number of observations only. The output of the reconstruction algorithm, when we take only a finite number of observations, is not guaranteed. But as we increase the number of observations, the probability to reconstruct ξ correctly (up to equivalence) goes to one.

References

[BK96] I. Benjamini and H. Kesten. Distinguishing sceneries by observing the scenery along a random walk path. *J. Anal. Math.*, 69:97–135, 1996.

[dH88] W. Th. F. den Hollander. Mixing properties for random walk in random scenery. *Ann. Probab.*, 16(4):1788–1802, 1988.

[dHS97] F. den Hollander and J. E. Steif. Mixing properties of the generalized T, T^{-1}-process. *J. Anal. Math.*, 72:165–202, 1997.

[HHR00] D. Heicklen, C. Hoffman, and D. J. Rudolph. Entropy and dyadic equivalence of random walks on a random scenery. *Adv. Math.*, 156(2):157–179, 2000.

[HK97] M. Harris and M. Keane. Random coin tossing. *Probab. Theory Related Fields*, 109(1):27–37, 1997.

[How96] C. D. Howard. Detecting defects in periodic scenery by random walks on $\mathbb{Z}$. *Random Structures Algorithms*, 8(1):59–74, 1996.

[How97] C. D. Howard. Distinguishing certain random sceneries on $\mathbb{Z}$ via random walks. *Statist. Probab. Lett.*, 34(2):123–132, 1997.

[KdH86] M. Keane and W. Th. F. den Hollander. Ergodic properties of color records. *Phys. A*, 138(1-2):183–193, 1986.

[Kes96] H. Kesten. Detecting a single defect in a scenery by observing the scenery along a random walk path. In *Itô's stochastic calculus and probability theory*, pages 171–183. Springer, Tokyo, 1996.

[Kes98] H. Kesten. Distinguishing and reconstructing sceneries from observations along random walk paths. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, pages 75–83. Amer. Math. Soc., Providence, RI, 1998.

[Lin99] E. Lindenstrauss. Indistinguishable sceneries. *Random Structures Algorithms*, 14(1):71–86, 1999.

[LP02] D. Levin and Y. Peres. Random walks in stochastic scenery on $\mathbb{Z}$. Preprint, 2002.

[LPP01] D. A. Levin, R. Pemantle, and Y. Peres. A phase transition in random coin tossing. *Ann. Probab.*, 29(4):1637–1669, 2001.

[Ter99] A. Terras. *Fourier Analysis on Finite Groups and Applications.* Cambridge University Press, Cambrige, 1999.