Random walks in stochastic surroundings
Rolles, S.W.W.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 1

Introduction

This thesis contains work on reinforced random walks, the reconstruction of random
sceneries observed along a random walk path, and the length of a longest increasing
subsequence in a random permutation. In this introduction, I will survey some of the work
in the area and describe my results. Furthermore I will explain how all three subjects fit
into the framework of random walks in stochastic surroundings. Section 1 is dedicated to
reinforced random walks. Section 2 describes scenery reconstruction problems. Section
3 deals with random permutations and explains the connection with up-right paths in a
Poissonian field.

1 Reinforced random walks

1.1 A short history

Different surveys on random processes with reinforcement have been written by Pemantle
[47], Davis [7], and Benaïm [4]. My emphasis here is on work which has connections with
my own results.

Random walks with edge reinforcement

Reinforced random walks were invented by Coppersmith and Diaconis in 1987 (see [11]).
They introduced edge-reinforced random walk, a nearest-neighbor random walk on a lo-
cally finite graph, as follows: All edges are given strictly positive numbers as weights. In
each step, the random walker jumps to a nearest-neighbor vertex traversing an edge e
incident to her current location with probability proportional to the weight of e. Each
time an edge is traversed, its weight is increased by 1. The process remembers where it
has been before and prefers edges which have been traversed often in the past. Edge-
reinforced random walk can be considered as a simple model for a person exploring a
new city. First she traverses randomly the streets around her hotel. As a street becomes
familiar to her, she has a higher preference to traverse the street again in the future.

In a special case, edge-reinforced random walk is well-known. Consider the graph
which consists of two vertices u, v and two parallel edges e, f connecting them. The
sequence of edges traversed by edge-reinforced random walk on this graph is a Polya urn
process. Recall that in a Polya urn process balls are drawn from an urn containing balls
with label e and f; after each drawing the ball is returned with an additional ball with the
same label. The process was introduced by Eggenberger and Polya [14] in 1923. We will see below that some well-known properties for the Polya urn process can be generalized for edge-reinforced random walk on a general finite graph.

Coppersmith and Diaconis (see [11]) were interested in the asymptotic behavior of the local time for a finite graph. Generalizations of their model were introduced by Davis [6], among others reinforced random walks of sequence type. In this model, a sequence $\delta_k \geq 0$, $k \geq 1$ is given, and the weight of an edge is increased by $\delta_k$ after the $k$th traversal. If all edges have initially weight 1, the weight of an edge after $k$ traversals equals $D_k := 1 + \sum_{i=1}^{k} \delta_i$. Tóth [59] considered weakly reinforced random walk; he assumed $D_k \sim k^\rho$ for $k \to \infty$ and $\rho \in ]0, 1[$. In one dimension, he proved limit theorems for the local time process and the position of the random walker at late times. Recently Limic [28] showed that for strong reinforcement ($D_k = k^\rho$ with $\rho > 1$) the process will become “stuck” eventually, traversing the same edge back and forth.

Recurrence

Although people have been studying reinforced random walks for almost 15 years, many fundamental questions remain open, for instance the recurrence question. We call a random walk path recurrent if it visits all vertices infinitely often and transient if it visits all vertices at most finitely often. We call a random walk recurrent (transient) if almost all its paths are recurrent (transient). For simple random walk Polya [50] proved recurrence on $\mathbb{Z}$ and $\mathbb{Z}^2$ and transience on $\mathbb{Z}^d$ for $d \geq 3$. Kakutani is reported to have said: "A drunken man will eventually find his way home but a drunk bird may be lost forever." For reinforced random walks, usually much less is known.

Pemantle proved in [44] the existence of a phase transition for edge-reinforced random walk on an infinite binary tree. He assumed that the weight of an edge is increased by $\delta$ after each traversal, and showed that there exists $\delta_c \approx 4.29$ such that for $\delta \in ]0, \delta_c[$ the process is transient, whereas for $\delta > \delta_c$, it is positive recurrent. Knowing this result, one could conjecture that edge-reinforced random walk on $\mathbb{Z}^d$, $d \geq 2$, is recurrent if the increment $\delta$ is large enough. But there seems to be no result in this direction. Pemantle's proof for the binary tree uses that the process has the same distribution as a random walk in random environment. This equivalence is not available for $\mathbb{Z}^d$, $d \geq 2$. Even for edge-reinforced random walk on the ladder $\mathbb{Z} \times \{1, 2\}$ recurrence seems to be an open problem. In a sense this is quite surprising because the edge-reinforcement should push the random walk back to its starting point, and it is not hard to see that recurrence for edge-reinforced random walk is equivalent to infinitely many returns to the starting point. Recurrence for an “unbiased” reinforced random walk on trees was studied by Pemantle and Peres in [48].

For the integer line, Davis [6] proved a dichotomy for reinforced random walk of sequence type with all initial weights equal to 1. He showed that the reinforced random walk is recurrent if $\sum_{k=1}^{\infty} 1/D_k = \infty$ (recall $D_k = 1 + \sum_{i=1}^{k} \delta_i$), whereas the process visits only finitely many points if $\sum_{k=1}^{\infty} 1/D_k < \infty$. Takeshima, in [57] and [58], studied recurrence and transience on $\mathbb{Z}$ under less restrictive assumptions. For instance, he considered edge-reinforced random walk ($\delta_k = 1$ for all $k$) on $\mathbb{Z}$ with arbitrary initial weights and gave a necessary and sufficient condition for recurrence in terms of the initial weights. Sellke [54] studied recurrence of the coordinate processes of reinforced random walk of sequence type on $\mathbb{Z}^d$, $d \geq 2$. In particular he could prove that for edge-reinforced random walk...
walk on $\mathbb{Z}^2$, both coordinate processes visit 0 infinitely often.

A seemingly simple reinforced random walk is \textit{once-reinforced random walk}. In this model, the weight of an edge is increased by $\delta > 0$ after the first traversal, but from the second traversal on, the weight of an edge does not change. Sellke [53] showed that once-reinforced random walk on a multi-level ladder $\mathbb{Z} \times \{1, 2, \ldots, d\}$ is recurrent for all $\delta \in (0, 1/(d-2)]$. In his thesis [63] Vervoort extended Sellke's result for $\delta$ very large. The recurrence question seems to be open for ladders with at least 3 levels and intermediate values of $\delta$. Recently, Durrett, Kesten and Limic [13] proved transience for once-reinforced random walk on an infinite regular tree for all reinforcement parameters $\delta > 0$.

Many recurrence questions remain open. In particular for graphs with cycles, proving recurrence for a reinforced random walk seems to be a hard problem.

\textbf{Vertex-reinforced random walk}

Giving weights to the vertices instead of the edges makes a significant difference. Vertex-reinforced random walk was introduced by Pemantle [45], see also [46]. Each time a vertex is visited, its weight is increased by 1. In each step, the random walker jumps to a nearest neighbor $v$ with a probability proportional to the weight of $v$. Vertex-reinforced random walk localizes in a strong sense. In [49], Pemantle and Volkov proved that on $\mathbb{Z}$ the process visits almost surely only finitely many vertices; with positive probability the random walker visits only 5 points. Localization takes also place in higher dimensions. Volkov [64] proved that vertex-reinforced random walk on a class of general graphs gets stuck with positive probability.

\textbf{Applications}

Othmer and Stevens ([43], [56]) used reinforced random walks to model the motion of myxobacteria, a species which produces a slime trail and prefers to glide on the slime produced earlier. The simplest model of Othmer and Stevens is edge-reinforced random walk. They study also more complicated models taking into account the aggregation of bacteria and the decay of the slime in the course of time.

Mauldin, Monticino, and von Weizäcker [42] introduced directionally reinforced random walks as an elementary model for time and space correlations in ocean surface wave fields. In this model, the $2d$ directions given by the unit vectors in $\mathbb{Z}^d$, $d \geq 1$, are reinforced; the random walker prefers to keep the same direction. Limit theorems for this process were studied by Horváth and Shao [18].

\textbf{1.2 My own work}

\textbf{Asymptotic behavior of the local time for edge-reinforced random walk}

Chapter 2 contains the article [23]. There, edge-reinforced random walk on a finite graph with arbitrary initial values is studied. Recall that in this model each time an edge is traversed, its weight is increased by 1.

Let $G = (V, E)$ be a finite connected graph with vertex set $V$ and edge set $E$. We assume $G$ has no direct loops, i.e. each edge has two different endpoints. The graph may contain parallel edges; thus two edges may have the same pair of endpoints. In [11], Diaconis states without proof the following result of Coppersmith and Diaconis: For
Chapter 1. Introduction

$n \geq 0$ and $e \in E$, let $\alpha_n(e)$ denote the proportion of time the random walker spends traversing edge $e$ up to time $n$. The random vectors $\alpha_n := (\alpha_n(e); e \in E)$ converge almost surely as $n \to \infty$. The limit is random with a distribution which is absolutely continuous with respect to Lebesgue measure on the simplex $\Delta := \{(x_e; e \in E) : x_e \geq 0, \sum_{e \in E} x_e = 1\}$. The density $\Phi$ of the limiting distribution is given explicitly up to a normalizing constant in [11]; $\Phi$ is strictly positive on the interior of $\Delta$. According to Diaconis [personal communication], the proof announced by Coppersmith and Diaconis was never published.

Chapter 2 contains the proof of a refinement of the statement of Coppersmith and Diaconis. We were able to compute the normalizing constant of the density $\Phi$, which was not known before. Furthermore we proved another limit theorem: Let $\beta_n$ denote the normalized cycle numbers for a fundamental system of cycles of $G$ (the vector with components equal to the number of times the individual cycles of the graph are traversed up to time $n$ divided by $\sqrt{n}$; traversals of the same cycle in different directions are counted with different signs). Then $(\alpha_n, \beta_n)$ converges in distribution; the limiting distribution is absolutely continuous with respect to the product of Lebesgue measure on $\Delta$ and Lebesgue measure on $\mathbb{R}^d$ with $d$ equal to the number of fundamental cycles of $G$. The density $\varphi$ of the limiting distribution has the form

$$\varphi(x, y) = \Phi(x) \sqrt{\frac{\det[A(x)]}{(2\pi)^d}} e^{-\frac{1}{2} y A(x)^{-1} y'}, x \in \Delta, y \in \mathbb{R}^d$$

with $A(x)$ a $(d \times d)$-matrix defined in terms of the cycle structure of the graph; here $y'$ denotes the transpose of $y$.

Suppose $G$ is the graph consisting of two vertices with two parallel edges $e, f$ between them, and assume $e$ and $f$ have initially weights $a$ and $b$, respectively. As noted at the beginning of the introduction, the weight of edge $e$ after $n$ steps of the reinforced random walker is the same as the number of balls in a Polya urn containing initially $a$ balls with label $e$ and $b$ balls with label $f$. In this special case the above result for edge-reinforced random walk is well-known. It states that the proportion of balls with label $e$ converges as time goes to infinity to a random variable beta distributed with parameters $a$ and $b$.

Equivalence with random walk in random environment

An important property of edge-reinforced random walk is partial exchangeability. We introduce the concept for a possibly infinite graph $G$ as follows: We call two finite admissible paths in $G$ equivalent if they have the same starting point and the same number of edge traversals for all (non-directed) edges. We define a stochastic process to be partially exchangeable if any two equivalent paths have the same probability. This notion of partial exchangeability is different from the one introduced by Diaconis and Freedman in [12]. They call a process partially exchangeable if any two finite paths with the same starting point and the same number of transition counts for all directed edges have the same probability. Any process which is partially exchangeable in our sense is partially exchangeable in the sense of Diaconis and Freedman.

A de Finetti-type theorem due to Diaconis and Freedman [12] implies that any recurrent partially exchangeable process is a mixture of Markov chains. We prove in Chapter 3, that under a recurrence assumption, partial exchangeability in the stronger sense implies
1. Reinforced random walks

that the process is a unique mixture of reversible Markov chains. This result applies in particular to edge-reinforced random walk on a finite graph where the mixing measure can be described explicitly (see Chapter 3, Theorem 3.1). The transition probabilities of a nearest-neighbor random walk on $G$ which is a reversible Markov chain can be described by weights on the edges; transitions are made with probabilities proportional to the edge weights. Hence the mixing measure for edge-reinforced random walk can be described by a probability measure on edge weights, for instance a probability measure $\mu$ on $\Delta$. By Markov chain theory, $\mu$ is just the distribution of $\lim_{n \to \infty} (\alpha_n(e); e \in E)$. In other words, edge-reinforced random walk on a finite graph has the same distribution as a random walk in random environment, where the environment is given by random weights on the edges distributed according to $\mu$.

For the graph consisting of two vertices and two connecting edges, this result states the well-known fact that sampling from a Polya urn is the same as sampling from a beta distribution.

**Directed-edge-reinforced random walk**

In Chapter 4, we will prove equivalence in distribution between a related model, namely directed-edge-reinforced random walk and a random walk in random environment. Let $G$ be a locally finite graph with all edges directed. We introduce directed-edge-reinforced random walk as follows: All directed edges are given strictly positive weights. In each step the random walker traverses a directed edge pointing from her current location to a nearest neighbor; jumps are made with probabilities proportional to the edge weights. Each time a directed edge is traversed, its weight is increased by 1.

We will see that directed-edge-reinforced random walk on any graph has the same distribution as a random walk in random environment, where the transition probabilities to leave a vertex have a Dirichlet distribution. If all initial edge weights are equal to 1, the probability to leave vertex $v$ is uniformly distributed on the $d$-dimensional simplex with $d$ equal to the out-degree of $v$. Transition probabilities at different vertices are independent. A similar equivalence for edge-reinforced random walk on trees was proved by Penman [44]. For edge-reinforced random walk on finite graphs with cycles, the environment is given by dependent weights, as we will see in Chapter 3. An equivalence between edge-reinforced random walk and random walk in random environment for infinite graphs with cycles seems to be unknown. Such a result would probably help to answer questions about the qualitative behavior of edge-reinforced random walk.

Thanks to the equivalence between directed-edge-reinforced random walk and random walk in random environment, many results known about the latter model imply results for directed-edge-reinforced random walk. For instance, the zero-one-law of Zerner and Merkl [66] implies that directed-edge-reinforced random walk on $\mathbb{Z}^2$ with all initial values equal cannot drift off in a particular direction with positive probability.

**Tubular recurrence**

In Chapter 4, which contains the article [24], a recurrence criterion is proved for random walk in random environment on $\mathbb{Z} \times G$ with $G$ equal to a finite graph with all edges directed. Using the equivalence between directed-edge-reinforced random walk and random walk in random environment, we can conclude recurrence for directed-edge-reinforced random walk.
Chapter 1. Introduction

For a Markovian nearest-neighbor random walk on $\mathbb{Z} \times G$, one can write down harmonic equations for the escape probabilities. This allows us to characterize recurrence and transience for random walk in random environment in terms of the two middle Lyapunov exponents of certain i.i.d. random matrices. In fact, the sign of the sum of the two middle Lyapunov exponents determines whether the random walk in random environment is recurrent, tends to $+\infty$, or to $-\infty$. If the distribution of the environment has a certain symmetry property, we can conclude recurrence. In particular, recurrence for directed-edge-reinforced random walk with all initial weights equal follows from the equivalence with random walk in random environment.

A characterization of edge-reinforced random walk

Chapter 3 contains the article [51]. There, a characterization of edge-reinforced random walk in terms of certain partially exchangeable sequences is proved.

Suppose $G$ is 2-edge-connected, i.e. removing an edge makes $G$ disconnected. We consider the class $\mathcal{E}$ of all nearest-neighbor random walks on $G$ which are partially exchangeable (in our sense) and have the property that the transition probabilities to traverse edge $e$ in the next step given the past depend only on the current location, the edge $e$, the number of times the current location of the random walker has been visited and the number of times $e$ has been traversed in the past. If we impose in addition some mild technical conditions on the processes in $\mathcal{E}$, then $\mathcal{E}$ contains precisely edge-reinforced random walks with all possible initial weights and non-reinforced random walks where transition probabilities are given by edge weights which do not change.

If $G$ is not 2-edge-connected, then the graph decomposes into maximal 2-edge-connected components and bridges (edges which when removed make the graph disconnected). In this case, the transitions of a process in $\mathcal{E}$ on any 2-edge-connected component are basically made according to edge-reinforced random walk or non-reinforced random walk; different transitions are possible for different components. For a careful statement of the result we refer the reader to Chapter 3, Theorem 2.1.

In the special case in which $G$ consists of two vertices connected by $m \geq 3$ parallel edges, the above result has been observed by Johnson (see [65]) in the 1920s. In this case the result gives a characterization of Dirichlet distributions in terms of certain exchangeable sequences.

2 Scenery reconstruction

In Section 1, we have considered random walks in stochastic surroundings changing after each step; the stochastic surroundings determined the transition probabilities of the random walk. In this section, the random walker observes the stochastic surroundings, but the latter do not influence the transition probabilities. We consider the $d$-dimensional integer lattice $\mathbb{Z}^d$, $d \geq 1$. A scenery is a coloring of $\mathbb{Z}^d$ with colors from the finite set $\mathcal{C} := \{1, 2, \ldots, C\}$. Let $\xi$ be a scenery, and let $S := (S_k; k \geq 0)$ be a random walk on $\mathbb{Z}^d$. At time $k$, the random walker observes the color $\xi(S_k)$ at her location $S_k$.

Early work on random sceneries was done by Keane and den Hollander ([22] and [9]) who studied ergodic properties of color records observed along a random walk path. Their investigations were motivated among others by work of Kalikow [21] in ergodic theory. More recently, den Hollander and Steif [8] generalized Kalikow's results.
2. Scenery reconstruction

In the 80s, Benjamin and independently den Hollander and Keane (see Kesten [26]) asked how much information about the scenery $\xi$ can be retrieved from the color record $(\chi_k := \xi(S_k); k \geq 0)$ observed along the random walk path. Of course, we assume we have no information about the position of the random walker at times $k \geq 0$; we are given only the distribution of $S$. In this context, the following questions have acquired a certain degree of popularity:

2.1 The scenery distinguishing problem

Let $\xi$ and $\eta$ be two different sceneries, which are known to us. If we are given either $(\xi(S_k); k \geq 0)$ or $(\eta(S_k); k \geq 0)$, but we are not told which of them, can we decide (with zero probability of error) which of the two sequences was observed? More precisely, we denote by $Q^l_\xi$ the distribution of $(\xi(S_k); k \geq l)$ on $\mathbb{C}^\mathbb{N}$, and we ask whether $Q^l_\xi$ and $Q^l_\eta$ are mutually singular for all $l \geq 0$. If one requires only $Q^0_\xi$ and $Q^0_\eta$ to be mutually singular and $\xi$ and $\eta$ have a different color at the origin, then it is clear that we can distinguish $\xi$ and $\eta$ by looking only at the first observation of the random walker.

We call $\xi$ and $\eta$ equivalent if $\xi$ can be obtained from $\eta$ by a translation, reflection and/or a rotation. In general, one can hope to distinguish two sceneries only if they are not equivalent. If e.g. $S$ is a simple random walk on $\mathbb{Z}$, then one cannot distinguish $\xi$ from the reflection of $\xi$.

Let $\rho_C$ be the uniform measure on the set of colors $C = \{1, \ldots, C\}$. We denote by $\rho^{(C)} := \rho^{\mathbb{Z}^d}_C$ the distribution of i.i.d. colorings of $\mathbb{Z}^d$ with $C$ colors. Benjamin and Kesten [5] proved that many sceneries observed along a recurrent random walk path with bounded jumps can be distinguished on $\mathbb{Z}$ and $\mathbb{Z}^2$: Any fixed scenery $\xi$ can be distinguished from $\rho^{(C)}$-almost all sceneries $\eta$ for all $C \geq 2$. For simple random walk on $\mathbb{Z}^d$, $d \geq 3$, the result holds under the assumption that there are sufficiently many colors. On the other hand, if there are only two colors, distinction may typically not be possible in high dimensions: Benjamin and Kesten [5] proved that $\rho^{(2)} \times \rho^{(2)}$-almost all pairs $(\xi, \eta)$ of colorings of $\mathbb{Z}^d$ with two colors cannot be distinguished if they are observed along a random walk which chooses in each step uniformly one of the $d$ directions in the positive quadrant and the dimension $d$ is large enough. For simple random walk, the 2-color distinction problem in high dimensions seems to be open. In the case two sceneries differ in precisely one point, the distinguishing problem was examined by Kesten [25] and Howard [19].

Lindenstrauss [29] proved that on the integer line it is not possible to distinguish all pairs of non-equivalent sceneries. He showed the existence of uncountably many sceneries on $\mathbb{Z}$ which cannot be distinguished. His construction works for the simple random walk and 2 colors and even for infinitely many colors.

2.2 The scenery reconstruction problem

A more ambitious question is whether we can even reconstruct the scenery $\xi$ from the observations $\chi := (\xi(S_k); k \geq 0)$ along the random walk path. Again we can only hope to reconstruct the scenery up to equivalence. More formally, we ask the following question: Given a scenery $\xi$ and a recurrent random walk $S$, does there exist a function $A : C^\mathbb{N} \to \mathcal{C}^{\mathbb{Z}^d}$, measurable with respect to the canonical $\sigma$-algebras, such that $A(\chi)$ and $\xi$ are equivalent almost surely with respect to the distribution of the random walk $S$? Howard
proved in [19] that on $\mathbb{Z}$ all periodic sceneries observed along a simple random walk path can be reconstructed.

Since there exist sceneries which cannot be distinguished, these sceneries cannot be reconstructed. In the following, we assume that $\xi$ is random with distribution $\rho^{(C)}$, and $S$ is a recurrent random walk independent of $\xi$.

In his thesis [36], Matzinger (see also [38] and [37]) proved that $\rho^{(2)}$-almost all 2-color sceneries on $\mathbb{Z}$ can be almost surely reconstructed if they are observed along a simple random walk path with holding (i.e. the random walk may have a positive probability to stay at its current location instead of making a jump to the left or to the right). This result is quite surprising because a simple random walk with holding can read any finite string of 1's and 2's by just walking on the pattern 12 in the scenery. In [39], Matzinger showed that $\rho^{(2)}$-almost all 2-color sceneries observed along a simple random walk path with holding can be reconstructed in polynomial time, in the sense that a finite piece of scenery of length $l$ around the origin can be reconstructed with high probability from the first $p(l)$ observations with a polynomial $p$. This answered a question of Benjamini.

In [33], Löwe, Matzinger, and Merkl studied the scenery reconstruction problem for random walk with jumps of bounded size on $\mathbb{Z}$. Their work was motivated by Kesten, who noticed that earlier scenery reconstructions relied heavily on the skip-freeness of the random walk. Löwe, Matzinger, and Merkl assume the random walk is recurrent and can reach every integer with positive probability. If there is at least one color more than possible jumps for the random walk, then $\rho^{(C)}$-almost all sceneries can almost surely be reconstructed.

In [32], Löwe and Matzinger proved that almost all i.i.d. uniformly colored sceneries on $\mathbb{Z}^2$ can be reconstructed as long as there are many colors. They assume the scenery is observed along a simple random walk path. The 3-color reconstruction method invented by Matzinger in [37] for simple random walk on $\mathbb{Z}$ was further analyzed by Löwe and Matzinger in [31]. They proved that the reconstruction works even in cases where the scenery has some correlations.

### 2.3 Reconstruction in polynomial time

In Chapter 6, which contains the article [40], it is proved that the scenery reconstruction of Löwe, Matzinger, and Merkl in [33] can be done in polynomial time. The setup is as follows: Let $\xi$ be an i.i.d. uniformly colored scenery on $\mathbb{Z}$, and let $S$ be a random walk with i.i.d. increments independent of $\xi$. We assume the random walk is recurrent, reaches every integer with positive probability, and its increments have finite support. We require that there are strictly more colors than possible steps for the random walk. Löwe, Matzinger, and Merkl in [33] showed that under these assumptions almost all sceneries can be almost surely reconstructed up to equivalence. In Chapter 6 it is proved that a finite piece of scenery of length of order $l$ around the origin can be reconstructed with high probability (up to reflection or a "small" translation) from the first $p(l)$ observations with a polynomial $p$. The probability that the reconstruction is performed correctly converges to 1 as the length $l$ tends to infinity.
2.4 Errors in the observations

In Chapter 5, which contains the article [41], it is proved that typical sceneries can be reconstructed even if there are some errors in the observations. We make the same assumptions on scenery and random walk as in the preceding reconstruction problem. In order to keep the exposition as simple as possible, we assume that for the random walk the maximal jump length to the left and to the right agree. We believe, our result is valid without this assumption. At time $k$ the random walker observes the color $\xi(S_k)$ at her current location with probability $1 - \delta$, whereas she observes a random error $Y_k$ with probability $\delta$. We assume that the occurrences of the errors are i.i.d., $Y := (Y_0, Y_1, \ldots)$ is stationary, ergodic, and independent of $\xi$ and $S$. Our main result states that almost all sceneries $\xi$ can be reconstructed almost surely up to equivalence for all sufficiently small error probabilities $\delta$.

2.5 Related coin tossing problems

Harris and Keane [17] and Levin, Pemantle, and Peres [27] studied closely related questions about random coin tossing. Let $c_{\text{fair}}$ be a fair coin and $c_{\text{biased}}$ a biased coin. Suppose we are given a coin tossing record $\chi' := (\chi'_k; k \geq 0)$ which is obtained a) from i.i.d. tosses with $c_{\text{fair}}$ or b) at renewal times of a renewal process $c_{\text{biased}}$ is tossed and at all other times $c_{\text{fair}}$. Can we almost surely distinguish whether $\chi'$ originates from a) or b)? In [17], Harris and Keane show that if $u_n$ denotes the probability of a renewal at time $n$ and $\sum u_n^2 = \infty$, then one can a.s. distinguish between a) and b) whereas distinction is not possible if $\sum u_n^2 < \infty$ and the bias is small. Levin, Pemantle, and Peres [27] noticed that the size of the bias $\theta$ can be crucial. They proved that for some renewal sequences there is a phase transition: There exists a $\theta_c \in (0, 1]$ such that for $|\theta| > \theta_c$ distinction is possible, whereas this is not the case for $|\theta| < \theta_c$.

The problem of reconstructing sceneries observed with errors as described above covers also the following coin tossing problem: Suppose we have "coins" $\gamma_1, \gamma_2, \ldots, \gamma_C$, each one with $C$ different faces $1, 2, \ldots, C$. The probability of tossing $i$ with coin $\gamma_i$ equals $1 - \delta + \delta/C$, whereas the probability of tossing any $j \neq i$ equals $\delta/C$; here $\delta \in [0, 1]$ is a fixed but sufficiently small parameter. Thus coin $\gamma_i$ will typically show face $i$, but with a small probability we will see something else. For all $z \in \mathbb{Z}$, we choose i.i.d. uniformly a coin $\zeta(z)$ from $\gamma_i$, $1 \leq i \leq C$. Let $S$ be a recurrent random walk with bounded jumps on $\mathbb{Z}$, independent of $\zeta$, and satisfying the assumptions made in the scenery reconstruction problem with errors. We produce a coin tossing record $\chi' := (\chi'_k; k \geq 0)$ tossing at time $k$ the coin $\zeta(S_k)$ at the location $S_k$ of the random walker. The scenery reconstruction result in Chapter 5 [41] implies that for $\rho(C)$-almost all $\zeta$, we can reconstruct up to equivalence the "scenery of coins" $\zeta$ from the coin tossing record $\chi'$ for almost all realizations of the random walk.

3 Random permutations

3.1 Up-right paths

In the preceding sections, we considered random walks in changing environments and in a random scenery. In this section, we study a different process which still fits in the
framework of random walks in stochastic surroundings. The stochastic surroundings are given by a Poisson process with intensity one in the plane. Given a realization $\omega$ of the Poisson process, we look at up-right paths. These are polygonal paths from $(0,0)$ to $(1,1)$ connecting piecewise linearly points of $\omega$ in such a way that the path only moves upwards and to the right. Once the Poisson point configuration is given, the paths are determined. The quantity of interest is the maximal number of points occurring in an up-right path.

The problem is connected with permutations: We order the $x$- and $y$-coordinates of the points in $\omega$ in increasing order. The point with the $i^{th}$ $x$-coordinate induces a pair $(i, \pi(i))$ where $\pi(i)$ indicates the order number of its $y$-coordinate. If there are $n$ points in the unit square $[0,1]^2$, then $\omega$ induces a permutation $\pi$ on $\{1,2,\ldots,n\}$. Conditioned on having $n$ points in $[0,1]^2$, $\pi$ is chosen with uniform probability $1/n!$ from the permutation group $S_n$. The points on an up-right path induce an increasing subsequence in $\pi$, namely a subsequence $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ with $\pi(i_1) < \pi(i_2) < \ldots < \pi(i_k)$. An up-right path containing a maximal number of points corresponds to a longest increasing subsequence. We denote the length of a longest increasing subsequence of $\pi$ by $L_n(\pi)$. In general, there may be more than one increasing subsequence of $\pi$ having the same length $L_n(\pi)$. At first sight, it is not obvious how the length of a longest increasing subsequence of a permutation $\pi \in S_n$ can be determined, in particular if $n$ is large. Schensted [52] found a bijection (the so-called Schensted correspondence) between permutations $\pi \in S_n$ and pairs of $n$-Young tableaux with the same shape having both the first row of length $L_n(\pi)$. The Schensted algorithm (see also [2]) provides an efficient way to find the length of a longest increasing subsequence.

### 3.2 Asymptotic behavior of the length of a longest increasing subsequence

In the last five years, considerable interest has been shown in the asymptotic behavior of $L_n(\pi)$ as $n$ goes to infinity if $\pi$ is chosen uniformly at random from $S_n$. People rediscovered the subject forty years after Ulam’s question about the “typical” asymptotic behavior of $L_n$ [61]. Based on Monte Carlo simulations, Ulam conjectured that $c := \lim_{n \to \infty} E_n[L_n]/\sqrt{n}$ exists; here $E_n$ denotes the expectation with respect to uniform measure on $S_n$. Ulam’s conjecture was proved in 1972 by Hammersley [16] using the subadditive ergodic theorem. The correct value $c = 2$ was identified independently by Logan and Shepp [30] and Vershik and Kerov [62] in 1977. Alternative proofs showing $c = 2$ were provided by Aldous and Diaconis [1], Seppäläinen [55], Johansson [20] and Groeneboom [15].

Until the mid to late 1990s there were only conjectures about the asymptotic behavior of the variance of $L_n$. It was believed that $\text{Var}(L_n) \sim n^\alpha$ as $n \to \infty$ for some $\alpha$. There were various conjectures for the values of $\alpha$, among them $\alpha = 1/3$ given by Kesten. In 1999, Baik, Deift, and Johansson [3] proved a non-standard central limit theorem using the theory of integrable systems and noncommutative matrix-valued Riemann-Hilbert problems. They proved that $(L_n(\pi) - 2\sqrt{n})/n^{1/6}$ converges in distribution as $n \to \infty$ to the Tracy-Widom distribution introduced by Tracy and Widom in [60] and that all moments converge to the corresponding moments of the Tracy-Widom distribution. Their result confirms in particular Kesten’s conjecture for the asymptotic behavior of $\text{Var}(L_n)$.

The large deviation principle to the law of large numbers was derived by Seppäläinen [55] and Deuschel and Zeitouni [10].
3.3 Moderate deviations

The upper tail moderate deviations were obtained by Löwe and Merkl [34]. This concerns the regime between the upper tail large deviation regime and the central limit regime. Their proof uses a formula to describe the relevant probabilities in terms of a solution of a rank 2 Riemann-Hilbert problem; this formula was invented by Baik, Deift, and Johansson in [3].

Chapter 7 contains the article [35]. There, the moderate deviation picture is completed: A moderate deviation principle for the lower tail probabilities of $L_n$ is proved. This refers to the regime between the lower tail large deviation regime and the central limit regime. Using an estimate in Baik, Deift, and Johansson’s article [3], we first derive lower tail moderate deviations for the Poissonized version of the problem. The Poissonized version naturally arises in the terminology of up-right paths when we do not condition on the number of particles in a Poissonian cloud. A de-Poissonization procedure yields the desired result.

References


Chapter 1. Introduction


References


