Title
Instruction sets for walks and the quantile path transformation

Author
Forman, Noah Mills

Publication Date
2013

Peer reviewed|Thesis/dissertation

Instruction sets for walks and the quantile path transformation

by

Noah Mills Forman

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jim Pitman, Chair
Professor Steven N. Evans
Professor Rhonda Righter

Fall 2013

Instruction sets for walks and the quantile path transformation

Abstract

Instruction sets for walks and the quantile path transformation

by

Noah Mills Forman

Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Jim Pitman, Chair

This thesis examines two objects: the stacked-instructions representation of a walk on a general state space, and the novel *quantile* path transformation for real-valued walks.

Instead of representing a walk by a chronological sequence of states visited, we may represent the walk by a collection of lists of instructions located at each state. On successive visits to a state, the walker reads and follows successive instructions from the list. However, there are some collections of finite lists for which there is no walk that follows all listed instructions; given such instructions, a walker would eventually arrive at some state at which it had already exhausted all instructions, and become stuck, prior to having read all instructions at other states. This thesis characterizes the instruction sets that can be exhausted by the walker before they become stuck.

The quantile transform is a novel path transformation on real-valued walks and Brownian motions of finite duration. This transformation relates to identities in fluctuation theory due to Wendel, Port, Dassios and others, and to discrete and Brownian versions of Tanaka's formula. For an n-step random walk, the quantile transform reorders increments according to the value of the walk at the start of each increment. We describe the distribution of the quantile transform of a simple random walk of n steps, using a bijection to characterize the number of pre-images of each possible transformed path. We deduce, both for simple random walks and for Brownian motion, that the quantile transform has the same distribution as Vervaat's transform. For Brownian motion, the quantile transforms of the embedded simple random walks converge to a time change of the local time profile. We characterize the distribution of the local time profile, giving rise to an identity that generalizes Jeulin's description of the local time profile of a Brownian bridge or excursion.

Contents

List of Figures

Acknowledgments

I am indebted to my coauthors on the quantile project, Sami Assaf and Jim Pitman, for the loan of their insights, their mathematical and academic wisdom, and their editing advice on the article that forms the spine of this thesis. I also thank Wenpin Tang for his work and our discussions on the quantile transform. And I again thank my advisor, Jim Pitman, for his supervision and support of me over the past four years.

Introduction

This thesis is themed around reorganizing information about the increments of a Markov chain on a finite time scale.

In Chapter 1 we consider the stacked-instructions representation of walks on general state spaces. Instead of representing a walk by a chronological sequence of states visited, we represent the walk by a collection of lists of instructions located at each state. On successive visits to a state, the walker reads and follows successive instructions from the list. The main result of this chapter characterizes collections of finite stacks of instructions that correspond, in a certain sense, to walks of finite length.

Chapter 2 studies a novel path transformation, called the *quantile transform*, on walks of finite length on $\mathbb{R}$; in particular, the chapter focuses on walks with ± 1 increments. This transformation reorders increments based on the value of the walk prior to each increment. If a given increment emerges from a low value visited by the walk, then it appears early in the quantile transformed walk. In some sense, increments from low values tend to go up, and increments from high values tend to go down. Therefore, the quantile transformed walk tends to rise early and fall late. This chapter presents joint work with Assaf and Pitman, currently being submitted for publication as [AFP13].

Chapter 3 continues the work of the previous chapter, now in the Brownian setting. We use limit results to describe the quantile transform of Brownian motion, which is an expression involving Brownian local times. The basic result of this chapter is an identity in distribution between the quantile transform and the Vervaat transform[Ver79] of Brownian motion; this result, too, appears in [AFP13]. The chapter then summarizes further descriptions of these transformed processes due to Lupus, Pitman, and Tang[LPT13].

Finally, chapter 4 discusses enumerative identities related to the instruction sets of chapter 1 and the quantile transform of chapter 2.

Figure 0.1 depicts the dependency relations between sections of the thesis. Note that several sections of chapter 4 require the definitions in the introduction to chapter 2 but none of the other content of that chapter.

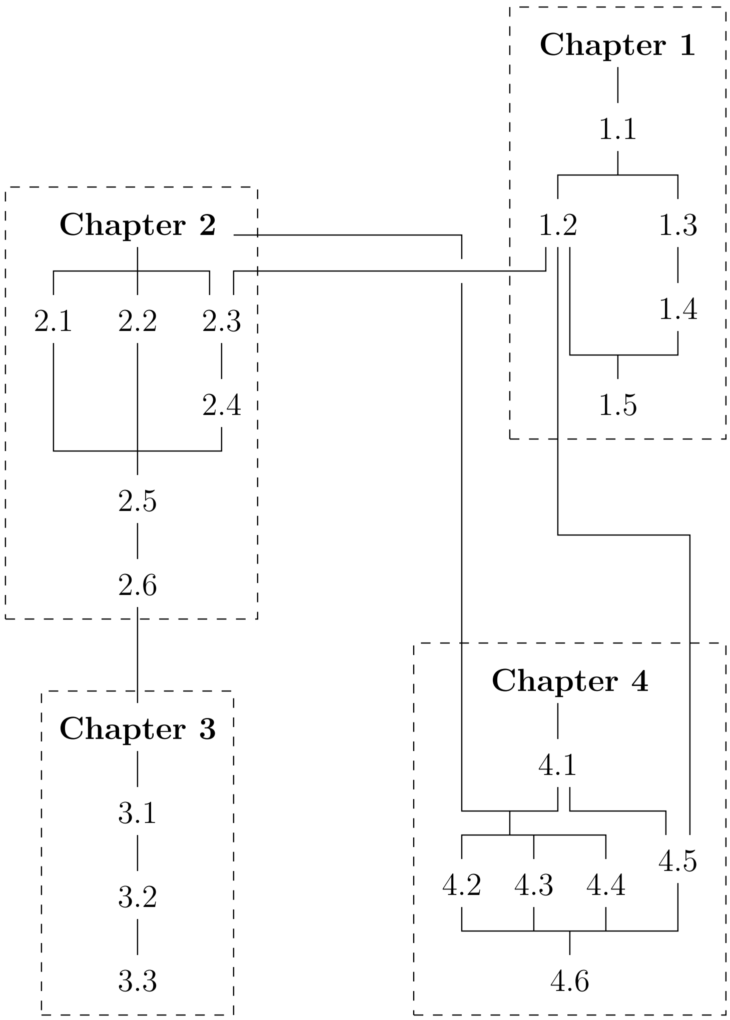


Figure 0.1: Dependence between sections of the thesis.

Chapter 1

Instruction sets

A path on a state space is typically represented by a chronological sequence of states visited. In this chapter we consider an alternative representation: a collection of lists of instructions located at each state. On successive visits to a state, the walker reads and follows successive instructions from the list.

Definition 1.1. An *instruction set* on a finite state space S is a pair $(\mathbf{x}, s)$ where $\mathbf{x} = (\underline{x}_a)_{a \in S}$ is a collection of finite sequences of states and s is a single *start state*. Given a finite undirected (or directed) graph $G = (S, E)$, we say that $(\mathbf{x}, s)$ is an *instruction set on* G if, for every $a \in S$ and every $b \in \underline{x}_a$, the (directed) edge (a, b) appears in E.

Instruction sets are a deterministic generalization of the stack model of Markov processes, discussed by Diaconis and Fulton[DF91, p. 4] in terms of stacks of cards and by Propp and Wilson[PW98, p. 205] in connection with the cycle-popping algorithm for generating a random spanning tree of an edge-weighted digraph. The stack model represents a Markov chain via a random instruction set, with each stack $\underline{x}_a$ comprising an infinite i.i.d. sequence of states drawn from the distribution for transitions of the chain away from a.

The instructions perspective also arises in Diaconis and Freedman[DF80]; see their Theorem (7). They use it to discuss a version of de Finetti's theorem, but relating the Markov property rather than independence.

The aim of this chapter is to give conditions for whether an instruction set minimally describes a path.

Definition 1.2. We say that a sequence $(b_0, b_1, \cdots, b_n)$ in S – also called a *path* – is *described* by, or *corresponds* to, an instruction set $(\mathbf{x}, s)$ if $b_0 = s$ and for every $a \in V$, the sequence $\underline{x}_a$ equals the subsequence $(b_j \ : \ b_{j-1} = a)$. An instruction set is called *valid* if it corresponds to some path.

Figure 1.2 states an algorithm that takes in an instruction set and outputs the path that results from "following instructions."

Some instruction sets are not valid. For example, given $S = \{u, v\}$, the instruction set with $\underline{x}_u = (v)$, $\underline{x}_v = ()$, and $s = v$ is invalid. In such cases, `instr_tracker` terminates and

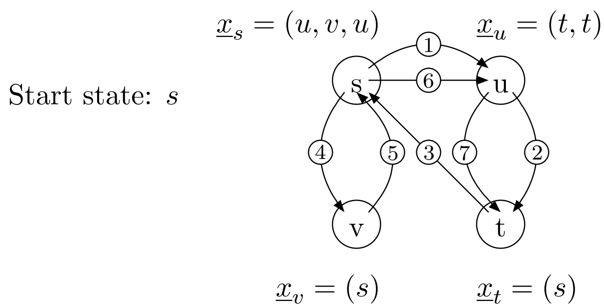


Figure 1.1: A valid instruction set and the corresponding path: s, u, t, s, v, s, u, t.

```
instr_tracker(x,s):
  ## Takes an instruction set on a state space S, returns a path.
  ## Treats each list x[a] as a queue.

  B[0] = s
  n = 0
  while x[B[n]] is not empty
    B[n+1] = x[B[n]].popleft()
    n = n + 1

  output B
```

Figure 1.2: Algorithm builds a path from an instruction set.

outputs a path, but this path does not incorporate all of the instructions listed in $\mathbf{x}$. On the other hand, if the algorithm incorporates every instruction from $\mathbf{x}$ into a path before terminating, then clearly this output path is described by the input instruction set.

Section 1.1 is devoted to proving Theorem 1.7, which characterizes the valid instruction sets. The remainder of the chapter explores variations on and extensions of this theorem. Section 1.2 presents alternative formulations of several of the definitions and results of section 1.1, first generally, and then in the setting of walks on groups. Section 1.3 discusses examples of instruction sets, the stack model, and related objects in the literature. Since several of these examples involve multiple walkers sharing an instruction set, in section 1.4 we extend Theorem 1.7 to instruction sets for multiple walkers. Finally, in section 1.5 we discuss possible directions for further research in this area.

The author realized belatedly that a version of the main result, Theorem 1.7, was previously demonstrated by van Aardenne-Ehrenfest and de Bruijn[vAEdB51] in the setting of Eulerian circuits. Nonetheless, it is useful to pass their result into the setting of stacked instructions and explore variations on the result as it applies to particular state spaces.

1.1 Characterization of valid instruction sets

We will say "multiset" to mean an unordered collection of objects that is allowed to have multiple copies of the same object. For example, as multisets,

$$\{a, a, b, c, c, c\} = \{c, a, c, a, c, b\} \neq \{a, b, c\}.$$

For a sequence or multiset X, we will write $\#\{x \in X\}$ to denote the multiplicity with which x occurs in X. Let $|X|$ denote the length or cardinality of X.

Definition 1.3. Given an instruction set $(\mathbf{x}, s)$ on S and a distinguished state $t \in S$, the *last-exit graph* with terminal state t is the directed graph $G_{\mathbf{x},t} = (S_{\mathbf{x}}, E_{\mathbf{x},t})$ with

$$\begin{aligned} S_{\mathbf{x}} &:= \{a \in S \ : \ |\underline{x}_a| > 0 \text{ or } a \in \underline{x}_b \text{ for some } b\} \text{ and} \\ E_{\mathbf{x},t} &:= \{(a, b) \in S^2 \ : \ b \text{ is the final entry of } \underline{x}_a,\ a \neq t)\}. \end{aligned} \tag{1.1}$$

Proposition 1.4. *If $(\mathbf{x}, s)$ is an instruction set on a directed (respectively, undirected simple) graph $G = (S, E)$ and $t \in S$ then (the simple graph underlying) $G_{\mathbf{x},t}$ is a subgraph of G.*

This proposition follows immediately from definitions.

Definition 1.5. A directed graph $T = (V, E)$ with $r \in V$ is a *directed tree rooted at r* if there exists a unique path in G from each other vertex to r.

Note that such a graph T cannot be a directed tree with respect to multiple different roots. For example, the root is the unique vertex in T with out-degree zero. If any edge were to leave r then paths to r would not be unique, since they could cycle out from r and back any number of times. Note also that the undirected multigraph underlying T is a tree. By "multigraph" we mean a graph whose edge set is a multiset – i.e. a graph that may have multiple edges between a single pair of vertices.

Definition 1.6. An instruction set $(\mathbf{x}, s)$ on a state space S has the *Degree property* with *terminal state* $t \in S$ if, for every $a \in S$,

$$|\underline{x}_a| = \mathbf{1}\{a = s\} - \mathbf{1}\{a = t\} + \sum_{b \in S} \#\{a \in \underline{x}_b\}. \tag{1.2}$$

Note that if $(\mathbf{x}, s)$ satisfies the Degree property for some t, then that t is uniquely specified. We say that $(\mathbf{x}, s)$ has the *Tree property* with terminal state t if its last-exit graph $G_{\mathbf{x},t}$ is a directed tree rooted at t.

The following is our main result.

Theorem 1.7 (Characterization of valid instruction sets)**.** *An instruction set $(\mathbf{x}, s)$ on a state space S is valid if and only if it has the Degree and Tree properties with some terminal state t, and then the corresponding path is unique and terminates at t.*

We prove this result below, after a proposition and a lemma relating instruction sets to Eulerian paths in directed graphs.

Definition 1.8. An *Eulerian path* on a graph (directed or undirected, with or without multi-edges) is a path that traverses each edge of the graph exactly once.

As we noted earlier, we realized belatedly that a version of Theorem 1.7 as it applies to Eulerian circuits – i.e. Eulerian paths that end back at their start vertices – was proved by van Aardenne-Ehrenfest and de Bruijn[vAEdB51], via the same approach that we take here. See in particular Theorem 5b in that paper. Indeed, our proof of necessity goes by way of a connection between Eulerian paths and instruction sets, but our proof of sufficiency appears below it in the language of instruction sets.

Proposition 1.9. *(i) Let $G = (V, E)$ be a connected directed multigraph and $s, t \in V$. Then G admits an Eulerian path from s to t if and only if, for every $a \in V$,*

$$\textit{out-degree}(a) = \textit{in-degree}(a) + \mathbf{1}\{a = s\} - \mathbf{1}\{a = t\}. \tag{1.3}$$

(ii) Let P denote an Eulerian path terminating at t on a directed multigraph $G = (V, E)$ with no isolated vertices. Let G' denote the subgraph of G whose edge set is the set of edges traversed by P upon its last exit from each vertex other than t. Then G' is a directed tree rooted at t.

Both of these properties are simple to prove. Property (i) dates back to Euler's study of the Seven Bridges of Königsberg; property (ii) was proved by [vAEdB51]. A discussion of these and related results appears in Bollobás[Bol98, p. 18]. Broder[Bro89] and Aldous[Ald90] use the tree property in an algorithm to generate random spanning trees with a certain probability distribution. The cycle-popping algorithm of Propp and Wilson [PW98], which we discuss in section 1.3, is a better-optimized version of the Broder-Aldous algorithm.

The following lemma asserts the necessity of the Degree and Tree properties in Theorem 1.7.

Lemma 1.10. *Let $(\mathbf{x}, s)$ be the instruction set corresponding to a path $(b_0 = s, b_1, \cdots, b_n)$ on a state space S. Then $(\mathbf{x}, s)$ satisfies the Degree and Tree properties with $t = b_n$.*

Proof. We define a directed multigraph $G_b = (B, E)$ with vertex set $B = \{b_0, \cdots, b_n\} \subseteq S$ (ignoring multiplicity) and edge multiset$\{(b_0, b_1), (b_1, b_2), \cdots, (b_{n-1}, b_n))\}$. Then $(b_0, \cdots, b_n)$ is the vertex sequence of an Eulerian path on G_b, so Proposition 1.9 applies. From the second part of that proposition we find that, since $(\mathbf{x}, s)$ corresponds to the path, it satisfies the tree property with terminal vertex b_n.

We now demonstrate the Degree property.

$$\text{out-degree}(b) = |\underline{x}_b|, \quad \text{and} \quad \text{in-degree}(b) = \sum_{a \in B} \#\{b \in \underline{x}_a\}.$$

Thus, the Degree property of $(\mathbf{x}, s)$ is equivalent to equation (1.3). $\square$

We are now ready to prove the theorem.

Proof of Theorem 1.7. Lemma 1.10 asserts that if an instruction set $(\mathbf{x}, s)$ corresponds to a path on a state space S terminating at t, then it satisfies the Degree and Tree properties with terminal state t. We now prove the sufficiency of these conditions for the existence of a corresponding path. If such a path exists, then its uniqueness is clear from the `instr_tracker` algorithm, stated in Figure 1.2.

Let $(\mathbf{x}, s)$ be an instruction set on S that satisfies the Degree and Tree properties with respect to some terminal state t. Suppose that we run `instr_tracker(x,s)`. We will prove that by the time the algorithm terminates, all queues $\underline{x}_a$ will be empty. Therefore the algorithm will output a path corresponding to $(\mathbf{x}, s)$.

The `instr_tracker` function terminates after appending some state a to the end of B, and then finding that the queue $\underline{x}_a$ has already been emptied. First we will show that this final state a must equal the terminal state t. Then we will prove by induction that by the time the algorithm terminates, all of the queues $\underline{x}_b$ have been emptied out.

Consider a vertex $a \neq t$. The Degree property indicates that, by the time the Instruction Tracking algorithm has incorporated a in B a total of $|\underline{x}_a|$ times, it will have read every instance of a in $\mathbf{x}$. Therefore, after it pops the final instruction from $\underline{x}_a$, it will never visit a again. Consequently, the algorithm cannot terminate by observing $\underline{x}_a$ to be empty. So by elimination, it must terminate by observing $\underline{x}_t$ to be empty.

We now proceed to our induction to show that, by the time `instr_tracker` terminates, all queues $\underline{x}_a$ are empty. We induct down generations of the last-exit tree $G_{\mathbf{x},t} = (V_{\mathbf{x}}, E_{\mathbf{x},t})$, beginning with the root.

Base case. We have shown that the algorithm must terminate upon observing $\underline{x}_t$ to be empty, so clearly $\underline{x}_t$ is empty when the algorithm terminates.

Inductive step. Suppose that when the algorithm terminates, the queue $\underline{x}_b$ is empty, for some $b \in V_{\mathbf{x}}$. Let a be a child of b in $G_{\mathbf{x},t}$ and we will show that the algorithm terminates with $\underline{x}_a$ empty as well.

In order to read all $|\underline{x}_b|$ instructions in that queue, the algorithm must visit vertex b a total of $|\underline{x}_b|$ times. If $b \neq t$ then, by the Degree property, this means reading every instance of b in $\mathbf{x}$. By definition of $G_{\mathbf{x},t}$ and by our choice of a, the last entry in $\underline{x}_a$ is a b, thus the algorithm must also read every instruction in $\underline{x}_a$ and terminate with that queue empty. If instead $b = t$ then, as we have observed, the algorithm will visit $\underline{x}_t$ a total of $|\underline{x}_t| + 1$ times. Again, the Degree property indicates that this requires reading every instance of t in $\mathbf{x}$, and by the same reasoning, we conclude that $\underline{x}_a$ must be empty by the end.

We conclude by induction down the generations of the last-exit tree that the algorithm terminates with every queue $\underline{x}_a$ empty. $\square$

1.2 Crossings and increment arrays

The Degree property may be restated in terms of sets of states, rather than individual states.

Definition 1.11. An instruction set $(\mathbf{x}, s)$ on state space S has the *Strong Crossings property* with terminal state t if, for every partition $\{S_1, S_2\}$ of S,

$$\sum_{a \in S_1,\ b \in S_2} \#\{b \in \underline{x}_a\} = \sum_{a \in S_2,\ b \in S_1} \#\{b \in \underline{x}_a\} + \mathbf{1}\{s \in S_1\} - \mathbf{1}\{t \in S_1\}. \tag{1.4}$$

The instruction set has the *Crossings property* with terminal state t, with respect to a total ordering $\preceq$ on S, if for every $a \in S$, equation (1.4) holds with

$$S_1 = \{b \in S \ : \ b \preceq a\} \ \text{ and } S_2 = \{b \in S \ : \ a \prec b\}.$$

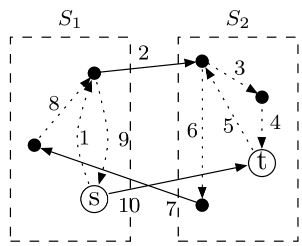


Figure 1.3: Equation (1.4) indicates that, in this case, one more step should pass from S_1 into S_2 than the reverse.

We illustrate the assertion of equation (1.4) in Figure 1.3. The edges in that figure are labeled with numbers to indicate the order in which steps are taken by the walk. Steps that occur within one of the partition blocks, which do not figure into the Crossings property, are dotted lines; steps from one block to the other are solid lines. Note that $s \in S_1$ and $t \in S_2$, so equation (1.4) indicates that there should be exactly one more step going from S_1 into S_2 than the reverse; and indeed there are two steps, steps 2 and 10, from S_1 into S_2, whereas only one step, number 7, traverses the other way.

Proposition 1.12. *Fix any total ordering $\preceq$ of S. The Degree, Strong Crossings, and Crossings (with respect to $\preceq$) properties for instruction sets on S are equivalent.*

Proof. The Strong Crossings property trivially implies the Crossings property with respect to any total ordering of S. We will show that the Degree property implies the Strong Crossings, and then that the Crossings property with respect to $\preceq$ implies the Degree property. Let $(\mathbf{x}, s)$ be an instruction set on S and $t \in S$. For this proof, we adopt the notation

$$f(S_1, s, t) := \mathbf{1}\{s \in S_1\} - \mathbf{1}\{t \in S_1\}.$$

Independently of the Degree or Crossings properties, the lengths of the $\underline{x}_a$ may be computed in terms of the $\#\{b \in \underline{x}_a\}$:

$$|\underline{x}_a| = \sum_{b \in S} \#\{b \in \underline{x}_a\}, \text{ and so} \tag{1.5}$$

$$\sum_{a \in S_1} |\underline{x}_a| = \sum_{a \in S_1,\ b \in S_1} \#\{b \in \underline{x}_a\} + \sum_{a \in S_1,\ b \in S_2} \#\{b \in \underline{x}_a\}. \tag{1.6}$$

In light of this identity, equation (1.4) is equivalent to

$$\sum_{a \in S_1} |\underline{x}_a| = \sum_{b \in S,\ a \in S_1} \#\{a \in \underline{x}_b\} + f(S_1, s, t). \tag{1.7}$$

This is also the identity that we obtain if we substitute in for $|\underline{x}_a|$ with equation (1.2); thus, the Strong Crossings property is implied by the Degree property.

Now suppose that $(\mathbf{x}, s)$ has the Crossings property with terminal state t with respect to $\preceq$. Let $a_1, a_2, \cdots, a_n$ be the elements of $S_{\mathbf{x}}$, as in (1.1), listed in ascending $\preceq$ order. We will derive the Degree property, equation (1.2), with $a = a_j$ first for $j = 1$, then for $j > 1$.

We appeal to equation (1.7) with $S_1 = \{a_1\}$.

$$\begin{aligned} |\underline{x}_{a_1}| &= \sum_{b \in S} \#\{a_1 \in \underline{x}_b\} + f(\{a_1\}, s, t) \\ &= \sum_{b \in S} \#\{a_1 \in \underline{x}_b\} + \mathbf{1}\{a_1 = s\} - \mathbf{1}\{a_1 = t\}. \end{aligned}$$

This is exactly equation (1.2) with $a = a_1$.

Now fix $j \leq n$. We appeal twice to equation (1.7), with $S_1 = \{a_1, \cdots, a_j\}$ and with $S_1' = \{a_1, \cdots, a_{j-1}\}$, and take the differences of corresponding sides to find

$$\begin{aligned} |\underline{x}_{a_j}| &= \sum_{b \in S} \#\{a_j \in \underline{x}_b\} + f(S_1, s, t) - f(S_1', s, t) \\ &= \sum_{b \in S} \#\{a_j \in \underline{x}_b\} + \mathbf{1}\{a_j = s\} - \mathbf{1}\{a_j = t\}. \end{aligned}$$

Thus, $(\mathbf{x}, s)$ satisfies the Degree property. □

In the case of a walk on a group, we may restate our definitions and results in terms of increments.

Definition 1.13. An *increment array* on a group $(S, *)$ is formally an instruction set $(\mathbf{x}, s)$ on S, but with a different notion of correspondence.

A sequence (i.e. path) $(b_0, b_1, \cdots, b_n)$ in S *corresponds* to or is *described* by $(\mathbf{x}, s)$ *as an increment array* if $b_0 = s$ and for every $a \in V$, the sequence $\underline{x}_a$ equals the subsequence $(b_j * b_{j-1}^{-1} \ : \ b_{j-1} = a)$.

Given a (directed) graph $G = (S, E)$ on the group S, we say that $(\mathbf{x}, s)$ is an *increment array on* G if, for every $a \in S$ and each $b \in \underline{x}_a$, the (directed) edge $(a, b * a)$ appears in E.

The elements of the sequences $\underline{x}_a$ are increments in this sense: from state a, if the walker reads the increment b from the list $\underline{x}_a$, then the walker proceeds to state $b * a$. This is described in Figure 1.4, which states the algorithm for building a path from an increment array.

```
incr_tracker(x,s):
  ## Takes an increment array on a group (S,*), returns a path.
  ## Treats each list x[a] as a queue.

  B[0] = s
  n = 0
  while x[B[n]] is not empty
    B[n+1] = x[B[n]].popleft() * B[n]
    n = n + 1

  output B
```

Figure 1.4: Algorithm builds a path from an increment array.

Definition 1.14. Let $(\mathbf{x}, s)$ be an increment array and $(\mathbf{y}, s)$ an instruction set on a group $(S, *)$. We say that $(\mathbf{x}, s)$ *corresponds* to $(\mathbf{y}, s)$ (as an increment array to an instruction set) if for each $a \in S$ we find $|\underline{x}_a| = |\underline{y}_a|$ and for each $n \in [1, |\underline{x}_a|]$ we get $x^a_n * a = y^a_n$, where x^a_n and y^a_n denote the n^{th} elements of $\underline{x}_a$ and $\underline{y}_a$ respectively.

We say that the increment array $(\mathbf{x}, s)$ has the Degree, Weak or Strong Crossings, or Tree property if and only if its corresponding instruction set $(\mathbf{y}, s)$ has that property.

Proposition 1.15. *If an increment array $(\mathbf{x}, s)$ on a group $(S, *)$ corresponds to an instruction set $(\mathbf{y}, s)$, then $(\mathbf{x}, s)$ describes a path if and only if $(\mathbf{y}, s)$ describes that same path.*

This proposition follows easily from the definitions.

For the purposes of the following chapters, we are particularly interested in certain classes of paths on $\mathbb{Z}$. In particular, we are interested in *simple walks*: processes that start at 0 and take steps of ± 1. We will also consider *skip-free (left) walks*: these begin at 0 as well and take steps of integer length no less than -1. In these contexts we obtain simpler statements to replace the Degree and Tree properties.

Lemma 1.16 (Characterization of valid increment arrays for simple walks)**.** *Let $(\mathbf{x}, s)$ be an increment array on $(\mathbb{Z}, +)$ with each increment equaling ± 1 or 0. For each $i \in \mathbb{Z}$ let u_i denote the number of $+1$s in $\underline{x}_i$ and d_i the number of -1s. Let x^*_i denote the final entry in $\underline{x}_i$. Then $(\mathbf{x}, s)$ has the* Crossings property *with terminal value t if and only if for each $i \in \mathbb{Z}$,*

$$u_{i-1} - d_i = \mathbf{1}\{i \leq t\} - \mathbf{1}\{i \leq s\}. \tag{1.8}$$

Furthermore, $(\mathbf{x}, s)$ has the Tree property *with terminal value t if and only if for every $i > t$ with $\underline{x}_i$ non-empty, the increment $x^*_i = -1$, and for every $i < t$ with $\underline{x}_i$ non-empty, the increment $x^*_i = 1$.*

Lemma 1.17 (Characterization of valid increment arrays for skip-free walks)**.** *Let* $(\mathbf{x}, s)$ *be an increment array on* $(\mathbb{Z}, +)$ *with each increment greater than or equal to* -1*. For each* $i \in \mathbb{Z}$ *let* u_i^n *denote the number increments of at least* n *in* $\underline{x}_i$ *and let* d_i *denote the number of* -1*s in* $\underline{x}_i$*. Let* x_i* *denote the final entry in* $\underline{x}_i$*. Then* $(\mathbf{x}, s)$ *has the* Crossings property *with terminal value* t *if and only if for each* $i \in \mathbb{Z}$*,*

$$\sum_{j<i} u_j^{i-j} - d_i = \mathbf{1}\{i \leq t\} - \mathbf{1}\{i \leq s\}. \tag{1.9}$$

Furthermore, $(\mathbf{x}, s)$ *has the* Tree property *with terminal value* t *if and only if the following conditions are met.*

(i) For $i \neq t$*, the increment* x_i^* *cannot equal* 0*.*

(ii) For $i > t$ *with* $\underline{x}_i$ *non-empty, the increment* $x_i^* = -1$*.*

(iii) For $i < t$*, if* $x_i^* = -1$ *then there is some* $j < i$ *with* $x_j^* > 0$*, and the greatest such* j *satisfies* $j + x_j^* > i$*.*

The proofs of these restatements are straightforward and are left to the reader.

1.3 Related topics in the literature

Here are several examples of the stack model and other objects resembling instruction sets in the literature. Several of these examples rely on multiple walkers sharing a single instruction set; we discuss this generalization in section 1.4.

Partial exchangeability. Given two paths $b = (b_i)_{i=0}^n$ and $c = (c_i)_{i=0}^n$ on a state space S, we will say $b \sim c$ if $b_0 = c_0$ and for each $x, y \in S$,

$$\#\{i \in [1, n] \ : \ b_{i-1} = x, \ b_i = y\} = \#\{i \in [1, n] \ : \ c_{i-1} = x, \ c_i = y\}.$$

I.e. this is the equivalence relation that equates paths that make the same numbers of transitions between the same pairs of states. A stochastic process $(X_i, \ i \in \mathbb{N})$ on S is called *partially exchangeable* if, for every finite n, the probability distribution on $(X_i, \ i \in [0, n])$ is uniform within each $\sim$-equivalence class. This is a weakening of both the Markov property and the exchangeability of the X_i. This property was studied by Diaconis and Freedman – see [DF84], and especially [DF80]. It has a simple interpretation if we represent X as a random valid instruction set: permuting the instructions within each stack does not change the probability of a given outcome, unless the resulting instruction set fails the Tree property, in which case it has probability 0.

The cycle-popping algorithm. The purpose of this algorithm is to provide a random directed spanning tree for a graph that selects edges biased by edge weight. This is achieved by giving mutually independent, i.i.d. sequences of instructions located at each vertex other than the intended terminal vertex t, with the distribution of each instruction biased by edge weight. If the front instructions – the first-exit graph – contains at least one cycle, then the algorithm selects one such cycle arbitrarily and deletes all instructions that contribute to it. This process is repeated until no cycles remain. In Theorem 16 of their paper, Propp and Wilson[PW98] make the interesting observation that it doesn't matter which cycle is deleted or in what order: in the end, the same tree is left. For more discussion of this and related algorithms see Lyons and Peres[LP13, Ch. 4].

Stacked instructions have also arisen in the area of chip-firing games. The general scheme of such games is this. We are given a graph – typically a simple graph – and a certain finite number of chips are initially placed on each vertex. The chips are then *fired* from vertices to their neighbors according to a given set of rules. Two of the better-known games are the *rotor-router* model of random walk and the game of *sandpile*. Holroyd et. al.[HLM$^+$08] offer a lengthy and well-referenced introduction to these two games. Eriksson[Eri96] discusses chip-firing on a mutating directed multigraph; this game specializes to both sandpile and rotor-router, among others. We briefly describe the two games below.

Rotor-routers. The rotor-router model is a deterministic model of a simple random walk on a simple graph in which each vertex has finite degree. Each vertex is equipped with an infinite sequence of instructions that cycles through a list of all of its neighbors. Initially some number of chips are distributed to various vertices in the graph. Then the chips advance according to their given instructions, one at a time. If an instruction is used to advance some chip from a vertex v, then the next chip to visit that vertex will read the next instruction down the list; that is, each instruction is read only once overall, rather than once by each visiting chip. Holroyd and Propp[HP10] demonstrate that some of the asymptotic features of random walks may also be observed with faster rates of convergence in certain natural rotor-router models.

Sandpile. In this game, a number of chips are spread among the vertices of a connected simple graph in which each vertex has finite degree. At any time, if a vertex has at least as many chips as it has neighbors, then that vertex may be fired. When a vertex is fired, one of its chips is shifted to each of its neighbors. A configuration of chips is *stable* if no vertex has enough chips to fire. If a stable configuration is reachable, then that is the only stable configuration reachable, and it will be reached regardless of the order in which vertices are fired. This independence of order is called the *Abelian* property of the game. See [LP10] for brisk and compelling introduction to this game; see [BS13] for a discussion of connections between sandpile and potential theory on graphs, as well as references discussing applications of sandpile in various areas of physics and mathematics.

Sandpile isn't explicitly described in terms of stacked instructions, but we could view it as a version of the rotor-router model in which a chips may only be fired from vertex v in multiples of degree(v).

In order to relate our instructions sets to chip-firing games and the Rotor-router model, we should study a multiple-walker version of instruction sets.

1.4 Multiple walkers

Definition 1.18. A *multiple walker instruction set*, or *m-w instruction set* $(\mathbf{x}, I)$ on a state space S is the same as an instruction set, but instead of including a single start state $s \in S$, its *initial configuration* is described by a non-empty finite sequence I of states.

Having multiple walkers necessitates a more complicated notion of correspondence between instructions and path.

Definition 1.19. Given a list $P = (p^1, p^2, \cdots)$ of sequences $p^i = (p^i_0, p^i_1, \cdots)$, a *monotone merge* of the collection is a sequence formed by merging the sequences together in a way that preserves order within each sequence and ultimately incorporates every element of all of the sequences. More formally, it is a sequence $p^* = (p^*_0, p^*_1, \cdots)$ such that there exists some *reference string* $(n_0, n_1, \cdots)$ of numbers with the property that for each n, the subsequence $(p^*_i \ : \ n_i = n)$ equals p^n.

A m-w instruction set $(\mathbf{x}, I)$ on a state space S *corresponds* to a list of paths $(p^1, p^2, \cdots)$ if $I = (p^1_0, p^2_0, \cdots)$ and for every $a \in S$ the sequence $\underline{x}_a$ is a monotone merge of the subsequences $(p^n_i \ : \ p^n_{i-1} = a)$ of the p^ns.

Unlike in the case of a single walker, we will not say that the instructions *describe* the collection of paths because this correspondence is not unique. Indeed, this is a weak notion of correspondence; there is no rule enforcing an order in which walkers take their steps. We have adopted this sense of correspondence because it is sufficient to classify which instruction sets describe collections of walks: for this purpose, it doesn't matter in which order walkers move. But it may be of interest to study stronger notions.

Proposition 1.20. *The relation between collections of paths given by corresponding to a common m-w instruction set is not an equivalence relation.*

Proof. Consider the following m-w instruction sets on $\{u, v, w\}$:

$$\underline{x}_u = (v, w),\ \underline{x}_v = (u),\ \underline{x}_w = (u),\ \text{with } I = (u, u)\text{; and}$$
$$\underline{y}_u = (w, v),\ \underline{y}_v = (u),\ \underline{y}_w = (u),\ \text{with } J = (u, u).$$

Compare these to the collections of paths:

$$((u, v, u, w, u),\ (u)),\ ((u, v, u),\ (u, w, u)),\ \text{and } ((u),\ (u, w, u, v, u)).$$

The first and second pairs of paths both correspond to the first m-w instruction set, and the second and third pairs of paths both correspond to the second m-w instruction set. There is no other instruction set that corresponds to any of these collections of paths. Therefore, the relation on collections of paths given by corresponding to a common m-w instruction set is not transitive. □

Despite the lack of specificity in our notion of correspondence, we offer a recursive algorithm to produce some collection of paths from instructions. This algorithm, presented in Figure 1.5, works by advancing a single chip (or walker) as many times as possible until it gets stuck at a vertex with no remaining instructions, and then doing the same with a second chip, and so on. As with `instr_tracker` in Figure 1.2, if this algorithm exhausts all queues `x[a]` before terminating then it outputs a corresponding collection of paths. Otherwise, the algorithm still outputs a collection of paths, but they will not correspond to the input.

```
m_w_instr_tracker(x,I)
  ## Takes a m-w instruction set on space S, returns a collection
  ## of paths. Treats each list x[a] as a queue.

  B[0] = I[0]  ## This part of the algorithm mimics instr_tracker
  n = 0
  while x[B[n]] is not empty
    B[n+1] = x[B[n]].popleft()
    n = n + 1

  if len(I) > 1
    B[1:len(I)] = m_w_instr_tracker(x,I[1:len(I)])

  output B
```

Figure 1.5: Algorithm builds a collection of paths from an m-w instruction set.

Definition 1.21. Given a m-w instruction set $(\mathbf{x}, I)$ on S and a set $T \subseteq S$, the *last-exit graph* of $(\mathbf{x}, I)$ with *terminal set* T is the directed graph $G_{\mathbf{x},I,T} = (S_{\mathbf{x},I}, E_{\mathbf{x},T})$ with

$$\begin{aligned} S_{\mathbf{x},I} &:= I \cup \{a \in S \ : \ |\underline{x}_a| > 0 \text{ or } a \in \underline{x}_b \text{ for some } b\} \text{ and} \\ E_{\mathbf{x},T} &:= \{(a,b) \in S^2 \ : \ b \text{ is the final entry of } \underline{x}_a, \ a \notin T)\}. \end{aligned} \tag{1.10}$$

Definition 1.22. A m-w instruction set $(\mathbf{x}, I)$ on a state space S has the *Degree property* with *terminal multiset* T if, for every $a \in S$,

$$|\underline{x}_a| = \#\{a \in I\} - \#\{a \in T\} + \sum_{b \in S} \#\{a \in \underline{x}_b\}. \tag{1.11}$$

We say that $(\mathbf{x}, I)$ has the *Forest property* with terminal multiset T if its last-exit graph $G_{\mathbf{x},I,T}$ is a directed forest rooted at the vertices of T. For the purposes of the last-exit graph, we ignore the multiplicities of elements of T.

Note that, as in the single-walker case, if an m-w instruction set satisfies the Degree property with some terminal multiset, then it cannot satisfy the property with any other. However, due to the non-uniqueness of corresponding collections of paths, the collection of terminal states cannot be arranged uniquely into a sequence to correspond to the sequence of initial states.

Theorem 1.23 (Characterization of valid m-w instruction sets)**.** *A m-w instruction set $(\mathbf{x}, I)$ on a state space S corresponds to some collection of paths on S with terminal multiset T if and only if it satisfies the Degree and Forest properties. Moreover, if $(\mathbf{x}, I)$ satisfies these two properties then the algorithm* `m_w_instr_tracker(x,I)` *outputs such a corresponding collection.*

Note that unlike Theorem 1.7, this does not assert the uniqueness of the corresponding collection of paths; Proposition 1.20 has precluded this. To prove this theorem we require a further definition.

Definition 1.24. Given two instruction sets $(\mathbf{x}, s)$ and $(\mathbf{y}, s)$ on a state space S, we say that the latter is an *initial segment* of the former if for every $a \in S$, the sequence $\underline{y}_a$ is an initial subsequence of $\underline{x}_a$. If $(\mathbf{y}, s)$ is an initial segment of $(\mathbf{x}, s)$, let the difference $\mathbf{x} - \mathbf{y}$ denote the collection of sequences that remains after deleting the initial subsequence $\underline{y}_a$ from each $\underline{x}_a$.

Given two m-w instruction sets $(\mathbf{x}, I)$ and $(\mathbf{y}, J)$ on S, we say that the latter is an *initial segment* of the former if the above condition holds, and if furthermore $J \subseteq I$ when I and J are viewed as multisets.

Proof of Theorem 1.23. We omit the proof of the necessity of the Degree and Forest properties as this follows along the same lines as the argument for Lemma 1.10. We prove sufficiency by proving that if $(\mathbf{x}, I)$ has the Degree and Forest properties then the function `m_w_instr_tracker(x,I)`, defined in Figure 1.5, outputs a corresponding collection of paths. We induct on the size of I.

Base case: $|I| = 1$. Say $I = \{s\}$. Then the Degree and Forest properties for $(\mathbf{x}, I)$ are equivalent to the Degree and Tree properties for $(\mathbf{x}, s)$, and `m_w_instr_tracker` behaves like `instr_tracker`. Therefore, our claim follows from Theorem 1.7.

Inductive step. Suppose that `m_w_instr_tracker(x,I)` outputs a corresponding collection of paths in the case $|I| = n$ and let $(\mathbf{x}, I)$ be a m-w instruction set with $|I| = n + 1$ that satisfies the Degree and Forest properties with some terminal multiset T. Let s denote the first entry in I. The first path, `B[0]`, in the sequence of paths output by `m_w_instr_tracker` corresponds to an instruction set $(\mathbf{y}, (s))$ that is an initial segment of $(\mathbf{x}, I)$. Since $(\mathbf{y}, (s))$ corresponds to a (collection of one) path, it satisfies the Degree and Forest properties with terminal set $\{t\}$.

Let $\mathbf{z} = \mathbf{x} - \mathbf{y}$ and let J denote the sequence I with its first element, s, removed. After producing `B[0]`, the function `m_w_instr_tracker(x,I)` recursively calls itself on input $(\mathbf{z}, J)$. By our inductive hypothesis, it suffices to prove that $t \in T$ and that the m-w instruction set $(\mathbf{z}, J)$ has the Degree and Forest properties with terminal multiset $U = T \backslash \{t\}$.

First, we demonstrate the Degree property for $(\mathbf{z}, J)$. By the Degree properties of $(\mathbf{x}, T)$ and $(\mathbf{y}, t)$, for $a \neq t$,

$$
\begin{aligned}
|\underline{z}_a| &= |\underline{x}_a| - |y_a| \\
&= \#\{a \in I\} - \mathbf{1}\{a = s\} - \#\{a \in T\} + \sum_{b \in S} (\#\{a \in \underline{x}_b\} - \#\{a \in \underline{y}_b\}) \\
&= \#\{a \in J\} - \#\{a \in U\} + \sum_{b \in S} \#\{a \in \underline{z}_b\}.
\end{aligned}
$$

Here, we are using the property $\#\{a \in T\} = \#\{a \in U\}$, which follows from our assumption $a \neq t$. In the case $a = t$, by definition of `instr_tracker` we have $|\underline{z}_a| = 0$. Thus the Degree property applied to $\underline{z}_a$ gives

$$
0 = \#\{t \in J\} - \#\{t \in T\} + 1 + \sum_{b \in S} \#\{t \in \underline{z}_b\}.
$$

Rearranging terms,

$$
\#\{t \in T\} = 1 + \#\{t \in J\} + \sum_{b \in S} \#\{t \in \underline{z}_b\} \geq 1.
$$

Therefore t must belong to T, and plugging back in to the previous equation,

$$
|\underline{z}_a| = \#\{t \in J\} - \#\{t \in U\} + \sum_{b \in S} \#\{t \in \underline{z}_b\}.
$$

We conclude that $(\mathbf{z}, J)$ has the Degree property with terminal multiset U.

We now show that $(\mathbf{z}, J)$ has the Forest property. As we noted above, the algorithm `m_w_instr_tracker` exhausts all instructions in $\underline{x}_t$ while generating `B[0]`. Therefore, regardless of whether $t \in U$, the last-exit graph $G_{\mathbf{z},J,U}$ has no edge leading away from t. So the edge set $E_{\mathbf{z},U} \subseteq E_{\mathbf{x},T}$. Consequently, our assumption that $G_{\mathbf{x},I,T}$ is a directed forest implies that $G_{\mathbf{z},J,U}$ is as well; and the definition ensures that every vertex of U is a root. So it suffices to prove that no vertex other than those in U is a root. By definition of the last-exit graph, it suffices to check that for every state $a \in S \setminus U$, if a appears in $\underline{x}_b$ for some $b \in S$ or if $a \in I$ then $\underline{x}_a$ is not empty. This is guaranteed by the Degree property.

The theorem follows by induction. $\square$

1.5 Further problems

We have not discussed infinite stacks of instructions or countably infinite state spaces. It should be possible to extend our main results to this class of instruction sets if we separately treat recurrent versus transient states.

We also have not attempted to apply our main results to study chip-firing games. It would be interesting to see to what extent this is possible. This would mean considering different rules for the order in which different walkers advance on a m-w instruction set.

Chapter 2

The quantile transform of a walk

Given a finite-duration simple walk with increments of ± 1 one observes that, except in the case of a first-passage bridge, the step immediately following the maximum value attained must be a down step, and the step immediately following the minimum value must be an up step. More generally, at a given value, the subsequent step is more likely to be an up step the closer the value is to the minimum and more likely to be a down step the closer the value is to the maximum. To study this phenomenon more precisely, one can form a two-line array with the increments of the walk and the value of the walk, and then sort the array with respect to the values line and consider the walk defined by the correspondingly re-ordered increments. It is this transformation, which we term the *quantile transform*, that we study here. Our study yields new distributional identities describing features of random walks and Brownian motion.

More formally, we will say *walk* to mean a discrete-time process $w : [0, n] \to \mathbb{R}$ with $w(0) = 0$. We will call a walk *simple* if its increments are all ± 1.

Definition 2.1. The *quantile permutation* corresponding to a walk w of length n, denoted ϕ_w, is the unique permutation of $[1, n]$ with the property that

$$(w(\phi_w(1) - 1), \phi_w(1) - 1);\ (w(\phi_w(2) - 1), \phi_w(2) - 1);\ \cdots;\ (w(\phi_w(n) - 1), \phi_w(n) - 1)$$

is the increasing lexicographic reordering of the sequence

$$(w(0), 0);\ (w(1), 1);\ \cdots;\ (w(n-1), n-1).$$

The *quantile path transform* sends w to the walk $Q(w)$ that starts at 0 and satisfies

$$Q(w)(j) = \sum_{i=1}^{j} x_{\phi_w(i)} \text{ for } j \in [1, n]. \tag{2.1}$$

In Figure 2.1 we show an example of a simple walk and its quantile transform; for each j the j^{th} increment of w is labeled with $\phi_w^{-1}(j)$. Observe that for a walk w of length n, we have $Q(w)(n) = w(n)$. As j increases, the process $Q(w)(j)$ incorporates increments which arise at higher values within w. Consider the example in Figure 2.1. The first increment of $Q(w)$ correspond to the only increment in w that originates at the value -2, the first four increments of $Q(w)$ correspond to those that originate at or below the value -1, and so on.

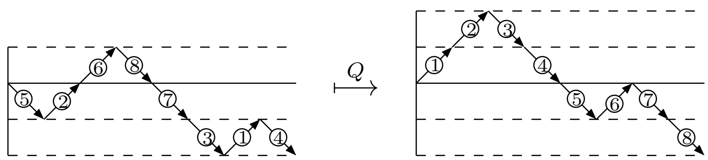


Figure 2.1: A walk and its quantile transform.

Note that the quantile permutation does not depend on the final increment x_n of w. A variant that does account for this final increment was previously considered by Wendel[Wen60] and Port[Por63], among others; this and several related path transformations and identities are discussed in section 4.2.

In this chapter we characterize the image of the quantile transform on simple walks.

Definition 2.2. A *quantile pair* is a pair (v, k) where v is a simple walk of length n and k is a nonnegative integer such that $v(j) \geq 0$ for $j \in [0, k)$ and $v(j) > v(n)$ for $j \in [k, n)$. For brevity, we refer to the *length* of the pair to mean the length of v.

The following is the main result of this chapter; we discuss the organization of our lengthy proof of the theorem below.

Theorem 2.3 (Quantile bijection)**.** *The map $w \mapsto (Q(w), \phi_w^{-1}(n))$ is a bijection between the set of simple walks of length n and the set of quantile pairs (v, k) of length n.*

In this statement, $\phi_w^{-1}(n)$ serves as a *helper variable*, distinguishing between multiple walks with the same Q-image. Figure 2.2 illustrates which indices k may appear as helper variables alongside a particular quantile walk v, depending on the sign of $v(n)$. If $v(n) < 0$ then its helper k may be any time from 1 up to the hitting time of -1. If $v(n) \geq 0$ and v ends in a down-step then k may be any time in the final excursion above the value $v(n)$, including time n. In the special case where $v(n) \geq 0$ and v ends with an up-step, k can only equal n.

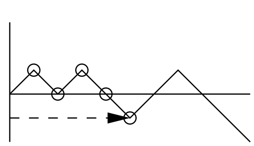


Figure 2.2: The allowed times for the helper variable (circled).

Figure 2.3 presents the five walks that have the same Q-image as the walk in Figure 2.1, corresponding to the five allowed helper times k. Note that in Figure 2.1, the final increment of the walk on the left is numbered '4' and arises fourth in the quantile transformed walk. This walk therefore appears fourth in our list of preimage walks in Figure 2.3. The other four preimage walks have final increments that arise 1^{rmst}, 2^{rmnd}, 3^{rmrd}, and 5^{rmth} when they are each quantile transformed. The procedure for inverting the quantile transform, as we have done in the figure, is not obvious; over the course of our proof of Theorem 2.3, we uncover such an inversion algorithm. The three stages of this algorithm are stated in Figure 2.16, equation (2.23), and Figure 2.9.

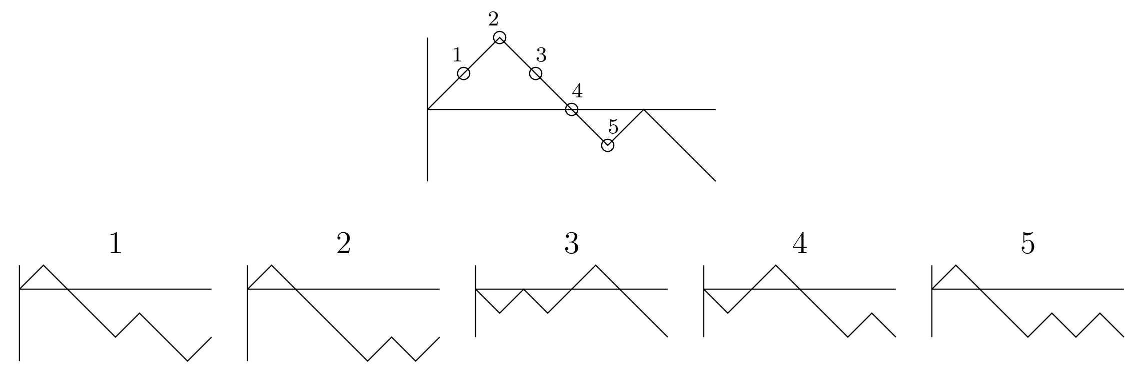


Figure 2.3: The five Q-preimages of a walk, corresponding to the five allowed helper times.

We prove Theorem 2.3 by proving three assertions: (i) for every simple walk w, the pair $(Q(w), \phi_w^{-1}(n))$ is a quantile pair; (ii) the number of quantile pairs (v, k) of length n is 2^n; and (iii) the map $w \mapsto (Q(w), \phi_w^{-1}(n))$ is injective. The last of these is the most difficult; we accomplish this by decomposing the quantile transform into three maps:

$$(Q(w), \phi_w^{-1}(n)) = \gamma \circ \beta \circ \alpha(w). \tag{2.2}$$

In the middle stages of our sequence of maps we obtain combinatorial objects which we call *marked (increment) arrays* and *partitioned walks.*

$$\text{walk} \xmapsto{\alpha} \text{marked array} \xmapsto{\beta} \text{partitioned walk} \xmapsto{\gamma} \text{walk-index pair.} \tag{2.3}$$

The three maps α, β, and γ are discussed in sections 2.3, 2.4, and 2.5 respectively.

Section 2.1 offers an image-but-no-multiplicities version of the Quantile bijection theorem for general (non-simple) walks. For simple walks w, this means that $(Q(w), \phi_w^{-1}(n))$ is a quantile pair.

In section 2.2 we relate Theorem 2.3 to a very similar theorem for a simpler path transformation, called the Vervaat transform. We use this to prove that the number of quantile pairs (v, k) of length n equals 2^n.

In section 2.3 we modify the terminology and reframe the results of chapter 1 to better suit our present purpose. As mentioned above, this section deals with the map α and *marked increment arrays.* We show that α is injective, and we describe its image.

In section 2.4 we introduce *partitioned walks* and the map β. This map is trivially a bijection; the main result of this section is Theorem 2.31, which describes the image of $\beta \circ \alpha$. Equation (2.21) in this chapter is a discrete version of Tanaka's formula; this formula has previously been studied in [Kud82, CR85a, Sza90, SS09], among other articles, and it plays a key role both in this section and in the continuous setting in chapter 3.

Finally, in section 2.5 we prove that γ acts injectively on the image of $\beta \circ \alpha$, thereby completing our proof of Theorem 2.3.

The bulk of this and the following chapter appear in the my pre-print with Sami Assaf and Jim Pitman[AFP13]. The major differences between that article and the presentation here are as follows.

- In this thesis we use the connection to Vervaat to enumerate the quantile pairs. The pre-print [AFP13] does this more directly, using well-known combinatorial identities. The derivation in that article has been postponed until section 4.3 in this thesis, so as to consolidate enumerative results in one chapter.
- The main result of section 2.3 is now a corollary, Corollary 2.21, to the results of chapter 1. The proof provided in that earlier chapter is substantially different from that offered in [AFP13].
- This thesis includes an algorithm, in Figure 2.16, to invert the quantile transform on a quantile pair.

2.1 The quantile transform of a non-simple walk

It is relatively easy to describe the image of the quantile transform; the difficulty lies in enumerating the preimages of a given image walk. In this section we do the easy work, offering in Theorem 2.7 a weak version of Theorem 2.3 in the more general setting of non-simple walks.

In discussing the proof and consequences of Theorem 2.3 it is helpful to refer to several special classes of walks.

Definition 2.4. We have the following special classes of walks:

- A *bridge to value* a is walk w of length n with $w(n) = a$; when $a = 0$, the walk w is simply called a *bridge*.
- A *non-negative walk* is a walk of finite length that is non-negative at all times.
- A *first-passage bridge* of length n is a walk w that does not reach $w(n)$ prior to time n.
- A *Dyck path* is a non-negative bridge (to value 0).

- An *excursion* is a Dyck path that does not visit the value 0 except to depart from it at time 0 and to terminate there at its final time.
- A *quantile walk* is a simple walk that is either non-negative or a first-passage bridge to a negative final value.

Other than the last, most of these names are standard. The quantile walks have been so named because they are the walks that may appear in a quantile pair. The Quantile bijection theorem has the following special case.

Corollary 2.5. *The quantile transform of a simple walk bridge (to 0) is a Dyck path. Moreover, for a uniform random bridge B of length $2n$ and a fixed Dyck path d of the same length,*

$$\mathbf{P}\{Q(B) = d\} = \frac{x}{\binom{2n}{n}}, \tag{2.4}$$

where x is the duration of the final excursion of d.

This corollary comes from the observation that a given Dyck path d belongs to exactly x quantile pairs (d, k): the associated helper variable k must be some time after the final visit of d to 0.

Definition 2.6. Given a walk w, for $j \in [1, n]$ we define the *quantile function of occupation measure* as

$$A_w(j) := w(\phi_w(j) - 1).$$

This function may also be expressed without reference to the quantile permutation:

$$A_w(j) = \min\{a \in \mathbb{R} \ : \ \#\{i \in [0, n-1] \ : \ w(i) \leq a\} \geq j\}.$$

For the walk on the left in Figure 2.1,

$$A_w(1) = -2, \ \ A_w(2) = A_w(3) = A_w(4) = -1,$$
$$A_w(5) = A_w(6) = A_w(7) = 0, \text{ and } \ A_w(8) = 1.$$

Figure 2.4 illustrates the relationship between the quantile transform and the quantile function of occupation measure. The graph of w is shown on the left and that of $Q(w)$ is on the right. The increments which contribute to $Q(w)(6)$ are shown in both graphs as numbered, solid arrows, and those that do not contribute are shown as dashed arrows. The time $j = 6$ is marked off with a vertical dotted line on the left. Increments with their left endpoints strictly below $A_w(j)$ do contribute to $Q(w)(6)$, increments which originate at exactly $A_w(j)$ may or may not contribute, and increments which originate strictly above $A_w(j)$ do not contribute.

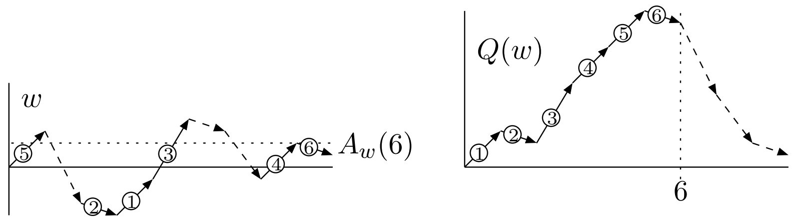


Figure 2.4: The value $Q(w)(6)$ is the sum of increments of w which originate below $A_w(6)$, as well as some which originate exactly at $A_w(6)$.

Theorem 2.7. *For any walk w of length n,*

$$\begin{array}{ll} Q(w)(j) \geq 0 & \text{for } j \in [0,\ \phi_w^{-1}(n)), \text{ and} \\ Q(w)(j) > Q(w)(n) & \text{for } j \in [\phi_w^{-1}(n),\ n). \end{array} \tag{2.5}$$

Consequently, $Q(w)$ is either a non-negative walk in the case where $w(n) \geq 0$ or a first-passage bridge to a negative value in the case where $w(n) < 0$.

Proof. First we prove that for $j < \phi_w^{-1}(n)$ we have $Q(w)(j) \geq 0$. Afterwards, we prove that for $j \in [\phi_w^{-1}(n),\ n)$ we have $Q(w)(j) > Q(w)(n)$.

Fix $j < \phi_w^{-1}(n)$ and let

$$I := \{i \in [1, n] \ : \ \text{either } w(i-1) < A_w(j) \text{ or } w(i-1) = A_w(j) \text{ and } i \leq \phi_w(j)\}.$$

Thus

$$Q(w)(j) = \sum_{i \in I} x_i. \tag{2.6}$$

We partition I into maximal intervals of consecutive integers. For example, in Figure 2.5 with $j = 6$ we have $I = \{1, 2, 4, 5, 8, 9\}$, which comprises three intervals: $\{1, 2\}$, $\{4, 5\}$, and $\{8, 9\}$. We label these intervals I_1, I_2, and so on.

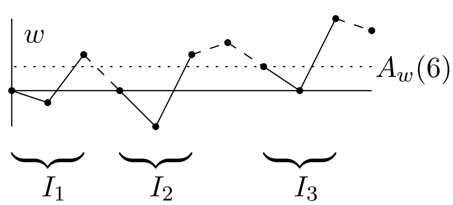


Figure 2.5: Three segments of the path of w correspond to the three intervals in I.

These intervals correspond to segments of the path of w, shown in solid lines in the figure. Each such segment begins at or below $A_w(j)$ and each ends at or above $A_w(j)$. Here we rely on our assumption that $j < \phi_w^{-1}(n)$ and thus $n \notin I$: if one of our path segments included the final increment of w then that segment might end below $A_w(j)$.

Thus, for each k we have

$$\sum_{i \in I_k} x_i \geq 0,$$

and so

$$Q(w)(j) = \sum_{i \in I} x_i = \sum_k \sum_{i \in I_k} x_i \geq 0.$$

Now fix $j \in [\phi_w^{-1}(n), n)$, and we must show that $Q(w)(j) > Q(w)(n)$. Let I^c denote $[1, n] - I$. Thus,

$$Q(w)(n) - Q(w)(j) = \sum_{i \in I^c} x_i.$$

As with I above, we partition I^c into maximal intervals of consecutive numbers, I_1^c, I_2^c, $\cdots$. These intervals correspond to segments of the path of w. Each such segment begins at or above and ends at or below $A_w(j)$. As in the previous case, here we rely on our assumption that $j \geq \phi_w^{-1}(n)$: if one of the I_k^c included the final increment then the corresponding path segment might end above $A_w(j)$.

Moreover if one of these segments begins exactly at $A_w(j)$ then it must end strictly below $A_w(j)$. In order for the segment corresponding to some block $[l, l+1, \cdots, m]$ of I^c to begin exactly at $A_w(j)$ we would need: (1) $w(l-1) = A_w(j) = w(\phi_w(j) - 1)$ and (2) $l \in I^c$. Thus, by definition of I, we would have $l \geq \phi_w^{-1}(j)$. And since $m+1 \in I$ and $m+1 > \phi_w^{-1}(j)$, we would then have $w(m) < A_w(j)$, as claimed. We conclude that for each block I_k^c,

$$\sum_{i \in I_k^c} x_i < 0.$$

Consequently,

$$Q(w)(n) - Q(w)(j) = \sum_{i \in I^c} x_i = \sum_k \sum_{i \in I_k^c} x_i < 0,$$

as desired. □

In the simple walk case, Theorem 2.7 says the following.

Corollary 2.8. *For a simple walk w of length n, the pair $(Q(w), \phi_w^{-1}(n))$ is a quantile pair.*

We note one more special case, comparable to Corollary 2.5, but which comes from considerations in our proof of Theorem 2.7.

Corollary 2.9. *The quantile transform of a Dyck path is an excursion.*

It follows from Corollary 2.8 that for a Dyck path d of length n, the quantile transform $Q(d)$ is also a Dyck path. So to prove this corollary it would suffice to show that $Q(d)$ does not visit the value 0 strictly between times 0 and n. This may be deduced from equation (2.6) and the discussion in our proof of Theorem 2.7.

Throughout the remainder of the chapter we say "walk" to refer to simple walks.

2.2 The Vervaat transform of a simple walk

The quantile transform has much in common with the (discrete) Vervaat transform V, studied in [Ver79]. Like the quantile transform, the Vervaat transform permutes the increments of a walk.

Definition 2.10. Given a walk w of length n, let

$$\tau_V(w) = \min\{j \in [0, n] : w(j) \leq w(i) \text{ for all } i \in [0, n]\}. \tag{2.7}$$

The *Vervaat permutation* ψ_w is the cyclic permutation of $[1, n]$ given by $i \mapsto i + \tau_V(w) \bmod n$, where we take $n \bmod n$ to equal n rather than 0. As with the quantile transform, we define the *Vervaat transform* V by

$$V(w)(j) = \sum_{i=1}^{j} x_{\psi_w(i)}. \tag{2.8}$$

An example of a walk and its Vervaat transform appears in Figure 2.6.

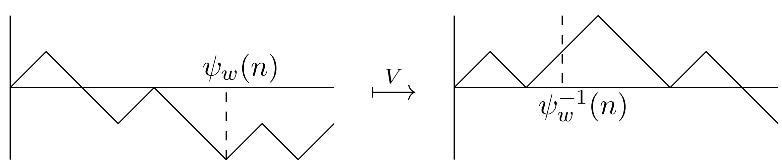


Figure 2.6: A walk transformed by V.

This transformation was studied by Vervaat because of its asymptotic properties. As scaled simple random walk bridges converge in distribution to Brownian bridge, the Vervaat transform of these bridges converges in distribution to a continuous-time version of the Vervaat transform, applied to the Brownian bridge. Vervaat used this to prove Theorem 3.4, which we discuss in the next chapter.

Surprisingly, the discrete Vervaat transform has a very similar bijection theorem to that for Q.

Definition 2.11. A *Vervaat pair* is a pair (v, k) where v is a walk of length n and k is a nonnegative integer such that $v(j) \geq 0$ for $0 \leq j \leq k$ and $v(j) > v(n)$ for $k \leq j < n$.

Theorem 2.12. *The map $w \mapsto (V(w), n - \tau_V(w))$ is a bijection between the walks of length n and Vervaat pairs.*

Proof. If we know that a pair (v, k) arises in the image of (V, K), then it is clear how to invert this map: let $y_i = v(i) - v(i-1)$ for each i; let $x_i = y_{i+k}$, where we take these indices mod n; and we define $F(v, k)$ to be the walk with increments x_i. Then $F(V(w), n - \tau_V(w)) = w$. We show that for every w the pair $(V(w), n - \tau_V(w))$ is a Vervaat pair, and that every Vervaat pair satisfies $(v, k) = (V(F(v, k)), n - \tau_V(F(v, k)))$.

Let w be a walk of length n. By definition of τ_V, for every $j \in [0, \tau_V(w))$ we have $w(j) > w(\tau_V(w))$. It follows that $V(w)(j) > v(n)$ for $j \in [n - \tau_V(w), n)$. Likewise, for $j \in [\tau_V(w), n]$ we have $w(j) \geq w(\tau_V(w))$; so it follows that $V(w)(j) \geq 0$ for $j \in [0, n - \tau_V(w)]$.

Now, consider a Vervaat pair (v, k). Then by definition of F and by the properties of the pair, for $j \in [0, n-k)$ we have $F(v, k)(j) > F(v, k)(n-k)$, and for $j \in [n-k, n]$, we have $F(v, k)(j) \geq F(v, k)(n-k)$. Thus, $\tau_V(F(v, k)) = n - k$, and the result follows. □

To our knowledge, this result has not been given explicitly in the literature. This statement strongly resembles our statement of Theorem 2.3, but we note two differences. The first is the helper variable. The helper variable in this theorem equals $\psi_w^{-1}(n)$ except in the case where w is a first-passage bridge to a negative value, in which case $\psi_w^{-1}(n) = n$ whereas $n - \tau_V = 0$; in our statement of Theorem 2.3, the helper always equals $\phi_w^{-1}(n)$ and cannot equal 0. The second difference is that the value $V(w)(k)$, where k is the helper time, must be non-negative, whereas $Q(w)(k)$ may equal -1 (see Figure 2.2). Again, this only affects the case where $w(n) < 0$.

Lemma 2.13. *For any walk v,*

$$\#\{k \ : \ (v, k) \text{ is Vervaat}\} = \#\{k \ : \ (v, k) \text{ is quantile}\}. \tag{2.9}$$

Proof. This follows from an inspection of the definitions of quantile and Vervaat pairs. In particular, let v be a walk of length n and $k \in [1, n]$. If $v(n) \geq 0$ then (v, k) is a quantile pair if and only if it is a Vervaat pair. And in the case $v(n) < 0$, the pair (v, k) is a quantile pair if and only if $(v, k-1)$ is a Vervaat pair. □

Corollary 2.14. *For $n \geq 0$ and $b \in \mathbb{Z}$ such that $n - b$ is even,*

$$\#\{(v, k) \text{ quantile} \ : \ v \text{ has length } n \text{ and } v(n) = b\} = \binom{n}{(n-b)/2}, \text{ and thus} \tag{2.10}$$

$$\#\{(v, k) \text{ quantile} \ : \ v \text{ has length } n\} = 2^n. \tag{2.11}$$

Proof. This follows from Theorem 2.12, Lemma 2.13, and the observation that both the Vervaat and quantile transforms preserve the length of a walk and its final value. □

An alternative proof of this enumeration that goes directly from the quantile transform, rather than passing by way of the Vervaat transform, appears in section 4.3.

2.3 Increment arrays

The quantile transform rearranges increments on the basis of their left endpoints; therefore, it is natural to compare the quantile transform of a walk to the walk's *increment array*, in the sense of section 1.2, which also groups increments by their start point. But in order to ease the passage to later results in this section, we reindex these arrays in terms of *level*.

Definition 2.15. Let w be a walk of length n. For $1 \leq j \leq n$ we define the *level* of (the left end of) the j^{th} increment of w to be

$$w(j-1) - \min_{0 \leq i < n} w(i).$$

The j^{th} increment of w is said to *belong to*, or to *leave*, that level. We name four important levels of a walk w, illustrated in Figure 2.7.

- The *start level* is the level of the first increment, or $-\min_{i<n} w(i)$. We typically denote this $\mathcal{S}$, or $\mathcal{S}_w$ in case of ambiguity.
- The *terminal level* is $(w(n) + \mathcal{S})$. We typically denote this $\mathcal{T}$ or $\mathcal{T}_w$.
- The *preterminal level* is the level of the final increment, or $(w(n-1)+\mathcal{S})$. We typically denote this $\mathcal{P}$ or $\mathcal{P}_w$.
- The *maximum level* is $\max_{j<n} w(j) + \mathcal{S}$. We typically denote this $\mathcal{L}$ or $\mathcal{L}_w$.

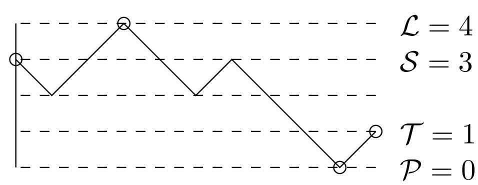


Figure 2.7: A walk with its distinguished levels labeled.

Note that if w is a first-passage bridge then no increments leave its terminal level. In this case $\mathcal{T}$ equals either -1 or $\mathcal{L} + 1$. Because $\mathcal{T}$ attains these exceptional values, the set of first-passage bridges arise as a special case throughout this chapter.

The start, preterminal, and terminal levels share the following relationship.

$$\mathcal{S} = \mathcal{T} - w(n) = \mathcal{P} - w(n-1). \tag{2.12}$$

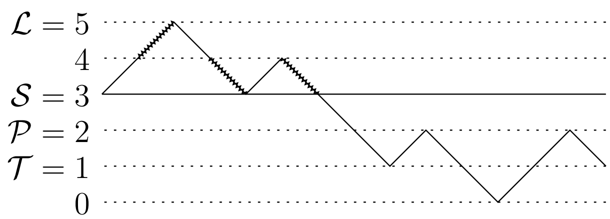
For the purposes of this chapter, we redefine increment arrays to be indexed by level rather than value. We also leave the start state out of our new definition.

Definition 2.16. An *increment array* is an indexed collection $\mathbf{x} = (\underline{x}_i)_{i=0}^{\mathcal{L}}$ of non-empty, finite sequences of ± 1s. We call the $\underline{x}_i$s the *rows* and $\mathcal{L}$ the *height* of the array. We say that an increment array $(\underline{x}_i)_{i=0}^{\mathcal{L}}$ *corresponds to a walk* w with maximum level $\mathcal{L}$ if, for every $i \in [0, \mathcal{L}]$, the sequence of increments of w at level i equals $\underline{x}_i$; i.e.

$$\underline{x}_i = (w(j+1) - w(j) \ : \ j \in [0, n), \ w(j) + \mathcal{S} = i).$$

Compare the above to Definition 1.13, specifically in the simple walk case discussed around Lemma 1.16. We may view these as increment arrays, in the sense of the earlier definition, on the graph on $\mathbb{Z}$ in which j is adjacent only to $j \pm 1$.

An example of a walk and its corresponding increment array is given in Figure 2.8. In that figure we've bolded the increments from level 4.



level i	d_i	u_i	$\underline{x}_i$
5	1	0	(-1)
4	**2**	**1**	$(\mathbf{1}, \mathbf{-1}, \mathbf{-1})$
3	1	2	(1,1,-1)
2	3	0	(-1,-1,-1)
1	1	2	(1,-1,1)
0	0	1	(1)

Figure 2.8: A walk with the corresponding increment array and up- and down-crossing counts.

Definition 2.17. Given an increment array $\mathbf{x}$, we define $u_i^{\mathbf{x}}$ and $d_i^{\mathbf{x}}$ to be the number of '1's and '-1's, respectively, that appear in $\underline{x}_i$. Correspondingly, for a walk w we define u_i^w and d_i^w to be the numbers of up- and down-steps of w from level i. We call the $u_i^{\mathbf{x}}$s and $d_i^{\mathbf{x}}$s (respectively u_i^ws and d_i^ws) the *up-* and *down-crossing counts* of $\mathbf{x}$ (resp. of w). We define the *sum of* $\mathbf{x}$, denoted $\sigma_{\mathbf{x}}$, to be the sum of all increments in the array:

$$\sigma_{\mathbf{x}} := \sum_{i=0}^{\mathcal{L}} \sum_{j \in \underline{x}_i} j = \sum_{i=0}^{\mathcal{L}} u_i - d_i. \tag{2.13}$$

Clearly, if $\mathbf{x}$ corresponds to a walk w of length n then $\sigma_{\mathbf{x}} = w(n)$, and for each i

$$u_i^{\mathbf{x}} = u_i^w \quad \text{and} \quad d_i^{\mathbf{x}} = d_i^w.$$

We now define the map α, which was alluded to in equations (2.2) and (2.3). We need this map to be injective, but we will see in Theorem 2.23 that the map from a walk to its corresponding increment array is not injective, so $\alpha(w)$ must pass some additional information.

Definition 2.18. Given an increment array $\mathbf{x} = (\underline{x}_i)_{i=0}^{\mathcal{L}}$, we may arbitrarily specify one row $\underline{x}_{\mathcal{P}}$ with $\mathcal{P} \in [0, \mathcal{L}]$ to be the *preterminal row*. We call the pair $(\mathbf{x}, \mathcal{P})$ a *marked (increment) array*, since one row has been "marked" as the preterminal row. We say that the marked array *corresponds to a walk* w if w corresponds to $\mathbf{x}$ and has preterminal level $\mathcal{P}$.

We define α to be the map which sends a walk w to its corresponding marked array.

Equation (2.12) may be restated in this setting. If an array $\mathbf{x}$ corresponds to a walk w with preterminal level $\mathcal{P}$ then the start and terminal levels of w are specified by

$$\mathcal{T} = \mathcal{P} + x^*_{\mathcal{P}}, \quad \text{and} \quad \mathcal{S} = \mathcal{T} - \sigma_{\mathbf{x}}, \tag{2.14}$$

where $x^*_{\mathcal{P}}$ denotes the final increment in the row $\underline{x}_{\mathcal{P}}$.

Definition 2.19. For a marked array $(\mathbf{x}, \mathcal{P})$ we define the indices $\mathcal{S}$ and $\mathcal{T}$ via equation (2.14). If $\mathcal{S}$ falls within $[0, \mathcal{L}]$ then we call $\underline{x}_{\mathcal{S}}$ the *start row* of $\mathbf{x}$; otherwise we say that the start row is empty. Likewise, if $\mathcal{T} \in [0, \mathcal{L}]$ then we call $\underline{x}_{\mathcal{T}}$ the *terminal row*, and if not then we say that the terminal row is empty.

In Figure 2.9 we state an algorithm to reconstitute the walk corresponding to a marked array. This is a version of the `incr_tracker` algorithm, from Figure 1.4, adapted to our redefinition of increment arrays. Figure 2.10 presents an example run of this algorithm. The example input $(\mathbf{x}, \mathcal{P})$ is shown at top of that figure, with each row below corresponding to an iteration of the loop.

```
Reconstitution(x[],P)
  ## Takes a marked array, returns a path.
  ## Treats each list x[a] as a queue.

  S = P + x[P][len(x[P])-1] - sum(x)  ## obtain S via (2.14)
  w[0] = 0
  n = 0
  While x[w[n]+S] not empty:
    w[n+1] = w[n] + x[w[n]+S].popleft()
    n = n+1

  output w
```

Figure 2.9: Algorithm builds a walk from a marked array.

We now wish to adapt Theorem 1.7 to the setting of marked arrays. We do this via Lemma 1.16.

$$\mathcal{P}=3;\quad \underline{x}_0=(1),\ \underline{x}_1=(-1,1),\ \underline{x}_2=(1),\ \underline{x}_3=(-1)$$

$$(0)\ w(0)+\mathcal{S}=1;\quad \underline{x}_0=(1),\ \underline{x}_1=(-\mathbf{1},1),\ \underline{x}_2=(1),\ \underline{x}_3=(-1)$$

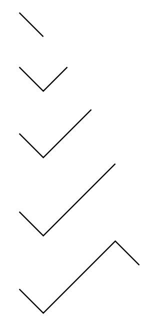
$$(1)\ w(1)+\mathcal{S}=0;\quad \underline{x}_0=(\mathbf{1}),\ \underline{x}_1=(1),\ \underline{x}_2=(1),\ \underline{x}_3=(-1)$$

$$(2)\ w(2)+\mathcal{S}=1;\quad \underline{x}_0=(),\ \underline{x}_1=(\mathbf{1}),\ \underline{x}_2=(1),\ \underline{x}_3=(-1)$$

$$(3)\ w(3)+\mathcal{S}=2;\quad \underline{x}_0=(),\ \underline{x}_1=(),\ \underline{x}_2=(\mathbf{1}),\ \underline{x}_3=(-1)$$

$$(4)\ w(4)+\mathcal{S}=3;\quad \underline{x}_0=(),\ \underline{x}_1=(),\ \underline{x}_2=(),\ \underline{x}_3=(-\mathbf{1})$$

$$(5)\ w(5)+\mathcal{S}=2;\quad \underline{x}_0=(),\ \underline{x}_1=(),\ \underline{x}_2=(),\ \underline{x}_3=()$$



[Terminates.]

Figure 2.10: Reconstitution algorithm (Fig. 2.9) run on a *valid marked array.*

Definition 2.20. A marked array has the *Bookends property* if for every $i \leq \min\{\mathcal{P},\mathcal{T}\}$ the final entry in $\underline{x}_i$ is a 1, and for each $i \geq \max\{\mathcal{P},\mathcal{T}\}$ the final entry is a -1.

A marked array has the *The Crossings property* if for each $i \in [0, \mathcal{L}+1]$

$$u_{i-1} - d_i = \mathbf{1}\{i \leq \mathcal{T}\} - \mathbf{1}\{i \leq \mathcal{S}\}, \tag{2.15}$$

where we define $u_{-1} = d_{\mathcal{L}+1} = 0$.

A marked array with the Bookends and Crossings properties is called *valid*. We call an increment array $\mathbf{x}$ valid if $(\mathbf{x},\mathcal{P})$ is valid for some $\mathcal{P}$.

The Bookends property here corresponds to the Tree property described in Definition 1.6.

Corollary 2.21. *The map α is a bijection from the set of walks to the set of valid marked arrays.*

This corollary follows immediately from Theorem 1.7 and Lemma 1.16. A different proof of Corollary 2.21 appears in [AFP13]; that other proof is based on the excursion structure of simple walks, whereas the proofs in chapter 1 of this thesis apply more generally. Note that, as a consequence of this corollary, the crossing counts for walks must satisfy equation (2.15).

We now digress from our main thread of proving the bijection between walks and quantile pairs to address the question: given a valid array $\mathbf{x}$, what can we say about the indices $\mathcal{P}$ for which $(\mathbf{x},\mathcal{P})$ is valid? We begin by asking: what does the Bookends property look like?

By the definition of $\mathcal{T}$ given in (2.14), it must differ from $\mathcal{P}$ by exactly 1. Therefore the two classifications $i \leq \min\{\mathcal{P},\mathcal{T}\}$ and $i \geq \max\{\mathcal{P},\mathcal{T}\}$ are exhaustive and non-intersecting. Given $\mathbf{x}$, there exists a $\mathcal{P}$ for which the Bookends property is satisfied if and only if, for all i below a certain threshold $\underline{x}_i$ ends in an up-step, and for all i above that threshold $\underline{x}_i$ ends in a down-step; if this is the case then $\mathcal{P}$ and $\mathcal{T}$ must stand on either side of that threshold.

Consider the following array.

$$\left.\begin{array}{r} \underline{x}_4 = (-1) \\ \underline{x}_3 = (+1, -1, -1) \end{array}\right\}$$
$$\left.\begin{array}{r} \underline{x}_2 = (-1, +1, +1, -1, +1) \\ \underline{x}_1 = (+1, +1, -1, +1) \\ \underline{x}_0 = (+1) \end{array}\right\}$$

The row-ending increments transition from 1s to -1s between rows 2 and 3. Thus, the Bookends property requires that either $\mathcal{P} = 2$ and $\mathcal{T} = 3$ or vice versa. Both of these choices are consistent with equation (2.14).

Proposition 2.22. *Given an increment array* $\mathbf{x}$*, there are at most two distinct triples* $(\mathcal{P}, \mathcal{T}, \mathcal{S})$ *that satisfy: (i) equation* (2.14)*, (ii) the Bookends property, and (iii) the property* $\mathcal{P} \in [0, \mathcal{L}]$*. Furthermore, if there are two such triples then no entry is the same in both triples.*

Proof. We begin with the special cases corresponding to first-passage bridges. First, suppose that every row of $\mathbf{x}$ ends in a '1'. Then the Bookends property and the bounds on $\mathcal{P}$ are only satisfied if $\mathcal{P} = \mathcal{L}$, and then $\mathcal{T}$ and $\mathcal{S}$ are pinned down by (2.14); in particular $\mathcal{T} = \mathcal{L} + 1$. By a similar argument, if every row ends in a '-1' then $\mathcal{P}$ must equal 0, and again $\mathcal{T}$ and $\mathcal{S}$ are specified by (2.14) with $\mathcal{T} = -1$.

Now suppose that some rows of $\mathbf{x} = (\underline{x}_i)_{i=0}^{\mathcal{L}}$ end in '1's and others in '-1's. Then there exists a $\mathcal{P}$ for which the Bookends property is satisfied if and only if there is some number $a \in [0, \mathcal{L})$ such that, for $i \leq a$ row $\underline{x}_i$ ends in a '1', and for $i > a$ row $\underline{x}_i$ ends in a '-1'. So the Bookends property and (2.14) force $(\mathcal{P}, \mathcal{T})$ to equal either $(a, a + 1)$ or $(a + 1, a)$. Thus, the two triples which satisfy all three properties are

$$(\mathcal{P}, \mathcal{T}, \mathcal{S}) = (a, a + 1, a + 1 - \sigma_{\mathbf{x}}) \text{ or } (a + 1, a, a - \sigma_{\mathbf{x}}). \tag{2.16}$$

□

We can now classify with which $\mathcal{P}$ a given $\mathbf{x}$ may form a valid marked array.

Theorem 2.23. *Let* $\mathbf{x} = (\underline{x}_i)_{i=0}^{\mathcal{L}}$ *be a valid array. If* $\sigma_{\mathbf{x}} \neq 0$ *then* $\mathbf{x}$ *corresponds to a unique walk, and if* $\sigma_{\mathbf{x}} = 0$ *then* $\mathbf{x}$ *corresponds to exactly two distinct bridges.*

Proof. By the uniqueness asserted in Corollary 2.21 it suffices to prove that if $\sigma_{\mathbf{x}} \neq 0$ (or if $\sigma_{\mathbf{x}} = 0$) then there is a unique $\mathcal{P}$ (respectively exactly two distinct values $\mathcal{P}$) for which $(\mathbf{x}, \mathcal{P})$ is valid. We proceed with three cases.

Case 1: $\sigma_{\mathbf{x}} > 0$. By Corollary 2.21, for any valid choice of $\mathcal{P}$ the resulting $\mathcal{S}$ lies within $[0, \mathcal{L}]$ – a walk must start at a level from which it has some increments. By the Crossings property,

$$u_{i-1} = d_i \text{ for } i \leq \mathcal{S} \text{ and } u_{\mathcal{S}+1} = d_{\mathcal{S}+1} + 1. \tag{2.17}$$

These two properties uniquely specify $\mathcal{S}$; and by Proposition 2.22 our choice of $\mathcal{S}$ uniquely specifies $\mathcal{P}$.

Case 2: $\sigma_{\mathbf{x}} < 0$. This dual to case 1. In this case, $\mathcal{S}$ must satisfy

$$u_i = d_{i+1} \text{ for } i \geq \mathcal{S} \text{ and } u_{\mathcal{S}-1} = d_{\mathcal{S}} - 1. \tag{2.18}$$

Again $\mathcal{S}$ is uniquely specified, and by Proposition 2.22 $\mathcal{P}$ is uniquely specified.

Case 3: $\sigma_{\mathbf{x}} = 0$. In this case, the Crossings property asserts that $u_i = d_{i+1}$ for every i; this places no constraints on $\mathcal{P}$, $\mathcal{T}$, or $\mathcal{S}$. By our assumption that $\mathbf{x}$ is valid, it therefore satisfies the crossings property regardless of $\mathcal{P}$, so the only constraints on $\mathcal{P}$ are coming from the Bookends property.

The Crossings property tells us that

$$d_0 = u_{-1} = 0 \quad \text{and} \quad u_{\mathcal{L}} = d_{\mathcal{L}+1} = 0,$$

so $\underline{x}_0$ ends in a '1' and $\underline{x}_{\mathcal{L}}$ ends in a '-1'. We observed in the proof of Proposition 2.22 that in this case there are either zero or two values $\mathcal{P}$ for which $(\mathbf{x}, \mathcal{P})$ is valid. And by our assumption that $\mathbf{x}$ is valid there are two such values. □

2.4 Partitioned walks

In this section we introduce partitioned walks and define the map β suggested in equations (2.2) and (2.3). A partitioned walk is a walk with its increments partitioned into contiguous blocks with one block distinguished. Partitioned walks correspond to marked arrays – and not just valid marked arrays – in a natural manner. Theorem 2.31, which is the main result of this section, describes the β-image of the valid marked arrays. The elements of this image set are called *quantile partitioned walks.* In section 2.5 we demonstrate a bijection between the quantile partitioned walks and the quantile pairs.

Let w be a walk of length n, and let the u_i^w and d_i^w be the up- and down- crossing counts of w from level i, as defined in the previous section.

Definition 2.24. For $j \in [0, \mathcal{L}+1]$, define t_j^w to be the number of increments of w at levels below j:

$$t_j^w := \sum_{i=0}^{j-1}(u_i + d_i).$$

So $0 = t_0^w < \cdots < t_{\mathcal{L}+1}^w = n$. We call t_j^w the j^{th} *saw tooth* of w.

Whenever it is clear from context, we suppress the superscript on the saw tooth of a walk.

Note that the helper variable employed in the quantile bijection theorem, Theorem 2.3, appears in the sequence of saw teeth:

$$\phi_w^{-1}(n) = t_{\mathcal{P}+1}. \tag{2.19}$$

This is because the n^{th} increment of w is its final increment at the preterminal level.

We are interested in the saw teeth in part because we can know the value of $Q(w)$ at t^w_j simply by knowing up- and down-crossing counts, without knowing the order of increments in each row of the array $\mathbf{x}_w$.

Lemma 2.25. *Let w be a walk with up- and down-crossing counts (u_i) and (d_i) and saw teeth (t_i). Let $\mathcal{S}$, $\mathcal{T}$, and $\mathcal{L}$ be the start, terminal, and maximum levels of w. Then*

$$Q(w)(t_j) = \sum_{i<j}(u_i - d_i) \quad \textit{for each} \ \ j \in [0, \mathcal{L}+1]. \tag{2.20}$$

This may be restated in the closed form

$$Q(w)(t_{j+1}) = u_j + (j-\mathcal{S})_+ - (j-\mathcal{T})_+ \quad \textit{for each} \ \ j \in [-1, \mathcal{L}]. \tag{2.21}$$

Proof. We note that $Q(w)(t_j)$ is a sum of all increments of w that belong to levels less than j. This proves equation (2.20). Regrouping the terms of (2.20) and applying equation (2.15) then gives equation (2.21). $\square$

Equation (2.21) is a discrete-time form of Tanaka's formula, the continuous-time version of which is discussed later around equation (3.5). Briefly, the value $Q(w)(t_{j+1})$ corresponds to the integral $\int_0^1 \mathbf{1}\{X(t) \leq a\} dX(t)$ in that it sums all increments of w which appear below the fixed level j; the term u_j corresponds to $\frac{1}{2}\ell^a$ – roughly half of the visits of a simple random walk to level j are followed by up-steps; and the latter terms $j - \mathcal{S}$ and $j - \mathcal{T}$ correspond to a and $a - X(1)$. Further discussion of the discrete Tanaka formula may be found in [Kud82, CR85a, Sza90, SS09].

Equation (2.21) takes the following form in the bridge case.

Corollary 2.26. *If w is a bridge then $Q(w)(t_{j+1}) = u_j$ for each j.*

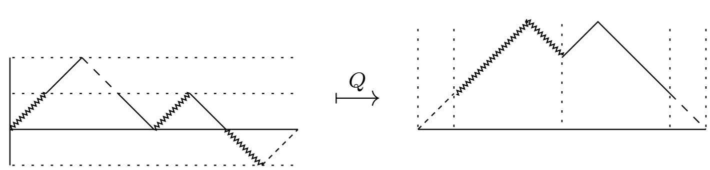
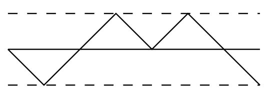


Figure 2.11: Increments emanating from a common level in w appear in a contiguous block in $Q(w)$.

The saw teeth partition the increments of $Q(w)$ into blocks in the manner illustrated in Figure 2.11: increments from the j^{th} block, between t_j and t_{j+1}, correspond to increments from the j^{th} level of w. This partition provides the link between increment arrays and the quantile transform. This is illustrated in Figure 2.12. The saw teeth are shown as vertical dotted lines partitioning the increments of $Q(w)$. Each block of this partition consists of the increments from a row of $\mathbf{x}_w$, stuck together in sequence.



$$\underline{x}_2 = (-1,-1)$$
$$\underline{x}_1 = (-1,1,1,-1)$$
$$\underline{x}_0 = (1)$$

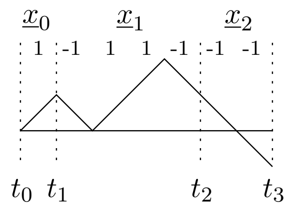


Figure 2.12: Left to right: a walk, its increment array, and its quantile transform partitioned by saw teeth.

We will now define the map β alluded to in equations (2.2) and (2.3) such that it will satisfy

$$\beta \circ \alpha(w) = (Q(w), (t_i^w)_{i=0}^{\mathcal{L}+1}, \mathcal{P}_w). \tag{2.22}$$

We define the *partitioned walks* to serve as a codomain for this map.

Definition 2.27. A *partitioned walk* is a triple $\mathbf{v} = (v, (t_i)_{i=0}^{\mathcal{L}+1}, \mathcal{P})$ where v is a walk, say of length n,

$$0 = t_0 < t_1 < \cdots < t_{\mathcal{L}+1} = n,$$

and $\mathcal{P} \in [0, \mathcal{L}]$. Here we are taking the t_j, $\mathcal{L}$, and $\mathcal{P}$ to be arbitrary numbers, rather than the saw teeth and distinguished levels of v. The name "partitioned walk" refers to the manner in which the times t_i partition the increments of v into blocks. We call the block of increments of v bounded by $t_{\mathcal{P}}$ and $t_{\mathcal{P}+1}$ the *preterminal block* of $\mathbf{v}$. We say that such a partitioned walk $\mathbf{v}$ *corresponds to a walk* w if $\mathbf{v} = (Q(w), (t_i^w)_{i=0}^{\mathcal{L}_w}, \mathcal{P}_w)$.

Definition 2.28. Define β to be the map which sends a marked array $((\underline{x}_i)_{i=0}^{\mathcal{L}}, \mathcal{P})$ to the unique partitioned walk $(v, (t_i)_{i=0}^{\mathcal{L}+1}, \mathcal{P})$ that satisfies

$$\underline{x}_i = \left(\begin{array}{r} v(t_i+1) - v(t_i),\ v(t_i+2) - v(t_i+1), \\ \cdots,\ v(t_{i+1}) - v(t_{i+1}-1) \end{array} \right) \text{ for every } i \in [0, \mathcal{L}]. \tag{2.23}$$

Define γ to be the map from partitioned walks to walk-index pairs given by

$$\gamma(v, (t_i), \mathcal{P}) := (v, t_{\mathcal{P}+1}). \tag{2.24}$$

We address the map γ in section 2.5. The map β may be thought of as stringing together increments one row at a time, as illustrated on the right in Figure 2.12, as well as in Figure 2.13. In this latter example neither the array nor the partitioned walk corresponds to any (unpartitioned) walk.

While it is clear that β is a bijection, we are particularly interested in the image of the set of *valid* marked arrays. Before we describe this image, we make a couple more definitions.

$$\begin{cases} \underline{x}_4 = (-1) \\ \underline{x}_3 = (1,1,-1) \\ \underline{x}_2 = (1,1) \\ \underline{x}_1 = (1,-1,-1) \ \bigstar \\ \underline{x}_0 = (-1,-1) \end{cases} \qquad \longleftrightarrow$$

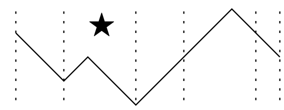


Figure 2.13: A marked array and its image under β.

Definition 2.29. Let $\mathbf{v} = (v, (t_i)_{i=0}^{\mathcal{L}+1}, \mathcal{P})$ be a partitioned walk. Motivated by the later terms in equation (2.21) we define the *trough function* for $\mathbf{v}$ to be

$$M_{\mathbf{v}}(j) := (j - \mathcal{S})_+ - (j - \mathcal{T})_+, \tag{2.25}$$

where we define the indices $\mathcal{T}$ and $\mathcal{S}$ via

$$\mathcal{T} := \mathcal{P} + v(t_{\mathcal{P}+1}) - v(t_{\mathcal{P}+1} - 1), \quad \text{and} \quad \mathcal{S} := \mathcal{T} - v(t_{\mathcal{L}+1}). \tag{2.26}$$

This is the partitioned walk analogue to equation (2.14) for marked arrays. If they exist, then we call the block of increments bounded by $t_{\mathcal{S}}$ and $t_{\mathcal{S}+1}$ the *start block*, and the block bounded by $t_{\mathcal{T}}$ and $t_{\mathcal{T}+1}$ the *terminal block*.

Definition 2.30. A partitioned walk has the *Bookends property* if for $i \leq \mathcal{T}, \mathcal{P}$, the t_{i+1}^{st} increment of v (i.e. the last increment of the i^{th} block) is an up-step; likewise, if $i \geq \mathcal{T}, \mathcal{P}$, then the t_{i+1}^{st} increment of v is a down-step.

A partitioned walk has the *Saw property* if for each $j \in [0, \mathcal{L}]$,

$$v(t_{j+1}) + v(t_j) = t_{j+1} - t_j + 2M_{\mathbf{v}}(j). \tag{2.27}$$

A partitioned walk with the Bookends and Saw properties is called a *quantile partitioned walk.*

Theorem 2.31. *The map β bijects the set of valid marked arrays with the set of quantile partitioned walks.*

The equivalence of the Bookends properties for partitioned walks versus marked arrays is clear. We complete the proof of this theorem as follows. First, we will define the *saw path* of a partitioned walk, and we use this to generate several useful restatements of the Saw property. Then we will demonstrate the equivalence of the Saw property of partitioned walks to the Crossings property of arrays; we use this to prove the theorem.

Definition 2.32. For any partitioned walk $\mathbf{v} = (v, (t_i), \mathcal{P})$, we define the *saw path* $S_{\mathbf{v}}$ to be the minimal walk that equals v at each time t_i.

The saw teeth of a walk w have been so-named because they typically coincide with the maxima of the saw path $S_{\mathbf{v}}$, where $\mathbf{v} = \beta \circ \alpha(w)$.

Lemma 2.33. *Let $\mathbf{v}$ be a partitioned walk, and let u_j and d_j denote the number of up- and down-increments of v between times t_j and t_{j+1} for each j. Then the saw property for $\mathbf{v}$ is equivalent to each of the following families of equations. For every $j \in [0, \mathcal{L}]$,*

$$M_{\mathbf{v}}(j) = -d_j + \sum_{i<j} u_i - d_i, \text{ or equivalently} \tag{2.28}$$

$$M_{\mathbf{v}}(j) = \min_{t \in [t_j, t_{j+1}]} S_{\mathbf{v}}(t). \tag{2.29}$$

Proof. By definition of the saw path

$$\min_{t \in [t_j, t_{j+1}]} S_{\mathbf{v}}(t) = -d_j + \sum_{i<j} u_i - d_i. \tag{2.30}$$

Thus, it suffices to show that the saw property is equivalent to (2.28).

First we express a few quantities in terms of the u_is and d_is:

$$t_{j+1} - t_j = u_j + d_j, \tag{2.31}$$

$$v(t_{j+1}) - v(t_j) = u_j - d_j, \text{ and} \tag{2.32}$$

$$v(t_j) = \sum_{i<j} u_i - d_i. \tag{2.33}$$

From these equations we obtain

$$v(t_{j+1}) + v(t_j) - (t_{j+1} - t_j) = -2d_j + 2v(t_j) = -2d_j + 2\sum_{i<j} u_i - d_i.$$

The saw property asserts that $2M_{\mathbf{v}}(j)$ equals the expression on the left-hand side above. The claim follows. ☐

Figure 2.14 shows two examples of

$$w \overset{\beta \circ \alpha}{\longmapsto} (Q(w), (t_i^w), \mathcal{P}).$$

The saw teeth are represented by vertical dotted lines and the preterminal block is starred. The saw path is drawn in dashed lines where it deviates below $Q(w)$. In between each pair of teeth t_j and t_{j+1} we show a horizontal dotted line at the level of $M_{\mathbf{v}}(j)$. Observe how the saw path bounces off of these horizontal lines; this illustrates equation (2.29).

In Figure 2.15 we show the saw path of a partitioned walk $\mathbf{v}$ which doesn't have Saw property. This diagram follows the same conventions as the diagrams on the right hand side in Figure 2.14.

By definition of the saw path

$$v(t) \geq S_{\mathbf{v}}(t) \text{ for every } t. \tag{2.34}$$

This gives us the following corollary to Lemma 2.33.

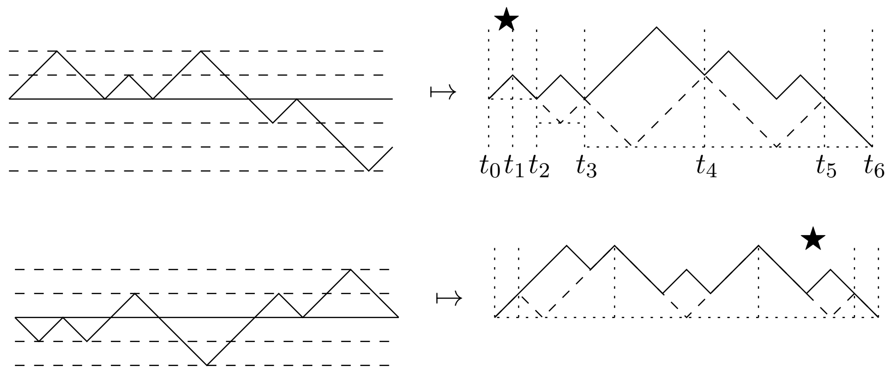


Figure 2.14: Two walks and their quantile transforms overlayed with saw teeth, saw paths, and troughs.

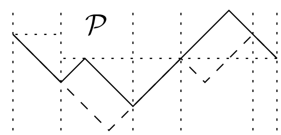


Figure 2.15: A general partitioned walk and its saw path.

Corollary 2.34. *If* $\mathbf{v} = (v, (t_i), \mathcal{P})$ *is a partitioned walk with the Saw property then for* $t \in [t_j, t_{j+1}]$,

$$v(t) \geq M_{\mathbf{v}}(j). \tag{2.35}$$

Lemma 2.35. *If* $\mathbf{v} = (v, (t_j)_{j=0}^{\mathcal{L}+1}, \mathcal{P})$ *is a partitioned walk with the Saw property then the index* $\mathcal{S}$ *of its start block falls within* $[-1, \mathcal{L}+1]$.

Proof. We consider three cases.

Case 1: $v(n) = 0$. Then $\mathcal{S} = \mathcal{T}$, and so the desired result follows from the definition of $\mathcal{T}$ in (2.26), and from the property $\mathcal{P} \in [0, \mathcal{L}]$ which is stipulated in the definition of a partitioned walk.

Case 2: $v(n) > 0$. Then $\mathcal{S} < \mathcal{T} \leq \mathcal{L} + 1$. But if $\mathcal{S} < -1$ then $M(0) > 0$. This would contradict Corollary 2.34 at $j = 0$, $t = 0$.

Case 3: $v(n) < 0$. Then $\mathcal{S} > \mathcal{T} \geq -1$. If both $\mathcal{S}, \mathcal{T} > \mathcal{L}$ then $M(\mathcal{L}) = 0$; this would contradict Corollary 2.34 at $j = \mathcal{L}$ with $t = n$. And if $\mathcal{T} \leq \mathcal{L} < \mathcal{S}$ then

$$M(\mathcal{L}) > (\mathcal{L} - \mathcal{S}) - (\mathcal{L} - \mathcal{T}) = v(n),$$

which would again contradict Corollary 2.34 at the same point. □

In fact, it follows from Theorem 2.31 that $\mathcal{S} \in [0, \mathcal{L}]$, but we require the weaker result of Lemma 2.35 to prove the theorem.

Proof of Theorem 2.31. Let $(\mathbf{x}, \mathcal{P})$ be a marked array and let $\beta(\mathbf{x}, \mathcal{P}) = \mathbf{v} = (v, (t_j)_{j=0}^{\mathcal{L}+1}, \mathcal{P})$. Clearly $(\mathbf{x}, \mathcal{P})$ has the Bookends property for arrays if and only if $\mathbf{v}$ has the Bookends property for partitioned walks. For the remainder of the proof, we assume that both have the Bookends property.

It suffices to prove that $\mathbf{v}$ has the Saw property if and only if $(\mathbf{x}, \mathcal{P})$ has the Crossings property. In fact, the Saw property is equivalent to the Crossings property even outside the context of the Bookends property, but we sidestep that proof for brevity's sake.

Let (u_j) and (d_j) denote the up- and down-crossing counts of $\mathbf{x}$; these also count the up- and down-steps of v between consecutive partitioning times t_j and t_{j+1}. Let $\mathcal{S}$ and $\mathcal{T}$ denote the start and terminal row indices for $(\mathbf{x}, \mathcal{P})$, or equivalently, the start and terminal block indices for $\mathbf{v}$.

The Saw property for $\mathbf{v}$ is equivalent to the following three conditions:

$$M_{\mathbf{v}}(-1) = 0, \quad M_{\mathbf{v}}(\mathcal{L}+1) = v(n), \text{ and} \tag{2.36}$$

$$M_{\mathbf{v}}(j) - M_{\mathbf{v}}(j-1) = u_{j-1} - d_j \text{ for each } j \in [0, \mathcal{L}+1]. \tag{2.37}$$

The Saw property implies (2.36) by way of Lemma 2.35; and given (2.36), equation (2.37) is equivalent to (2.28), which in turn is equivalent to the Saw property by Lemma 2.33.

The Crossings property for $(\mathbf{x}, \mathcal{P})$ is equivalent to those same three conditions. The validity of $(\mathbf{x}, \mathcal{P})$ implies (2.36) via Corollary 2.21: because the array corresponds to a walk, it must have $\mathcal{S} \in [0, \mathcal{L}]$. Furthermore, given (2.36) the Crossings property may be shown to be equivalent to (2.37) by substituting in the formula (2.25) for $M_{\mathbf{v}}$. $\square$

2.5 The quantile bijection theorem

In this section we give a lemma which will help us show that γ is injective on the quantile partitioned walks. We then apply this lemma to prove Theorem 2.3, the Quantile bijection theorem.

Lemma 2.36. *A partitioned walk $\mathbf{v} = (v, (t_i)_{i=0}^{\mathcal{L}+1}, \mathcal{P})$ has the Saw and Bookends properties if and only if the following two conditions hold.*

(i) For every $j \in [0, \mathcal{P}]$

$$t_j = \inf\{t \geq 0 : v(t) = t_{j+1} - t + 2M_{\mathbf{v}}(j) - v(t_{j+1})\}. \tag{2.38}$$

(ii) For every $j \in [\mathcal{P}+1, \mathcal{L}]$

$$t_{j+1} = \inf\{t \geq 0 : v(t) = t - t_j + 2M_{\mathbf{v}}(j) - v(t_j)\}. \tag{2.39}$$

Proof. The Saw property of $\mathbf{v}$ is equivalent, by algebraic manipulation, to the conditions that for $j \in [0, \mathcal{P}]$, the t_j must solve

$$v(t) + t = t_{j+1} + 2M_{\mathbf{v}}(j) - v(t_{j+1}) \tag{2.40}$$

for t, and for $j \in [\mathcal{P}+1, \mathcal{L}]$, the t_{j+1} must solve

$$v(t) - t = -t_j + 2M_{\mathbf{v}}(j) - v(t_j). \tag{2.41}$$

Now suppose that some s solves equation (2.40) for some $j \leq \mathcal{P}$. A time $r < s$ offers another solution to (2.40) if and only if

$$v(r) + r = v(s) + s.$$

This is equivalent to the condition that v takes only down-steps between the times r and s. Therefore t_j equaling the least solution to (2.40) is equivalent to the t_j^{th} increment of v being an up-step, as required by the Bookends property.

Similarly, suppose that s solves equation (2.41) for some $j \geq \mathcal{P}+1$. A time $r < s$ provides another solution if and only if

$$v(r) - r = v(s) - s,$$

which is equivalent to the condition that v takes only up-steps between r and s. Therefore t_j equaling the least solution to (2.41) is equivalent to the t_{j+1}^{st} increment of v being a down-step, as required by the Bookends property.

Equation (2.26) defines $\mathcal{T}$ from $\mathcal{P}$ in such a way that the $t_{\mathcal{P}+1}^{\text{st}}$ increment of v will always satisfy the Bookends property. Thus, if (2.38) holds for $j \in [0, \mathcal{P}]$ and (2.39) holds for every $j \in [\mathcal{P}+1, \mathcal{L}]$, then the Bookends property is met at every t_j. $\square$

Finally, we are equipped to prove the bijection theorem.

Proof of the Quantile bijection, Theorem 2.3. Definitions 2.18 and 2.28 define the maps α, β, and γ in such a way that, for a walk w of length n,

$$\gamma \circ \beta \circ \alpha(w) = (Q(w), \phi_w^{-1}(n)).$$

Corollary 2.8 asserts that this map sends walks to quantile pairs, and by Corollary 2.14 the set of walks with a given number of up- and down-steps has the same cardinality as the set of quantile pairs with those same numbers of up- and down-steps. Corollary 2.21 and Theorem 2.31 assert that that $\beta \circ \alpha$ bijects the walks with the quantile partitioned walks, so it suffices to prove that γ is injective on the quantile partitioned walks.

Now suppose that $\gamma(\mathbf{v}) = \gamma(\mathbf{v}') = (v, k)$ for some pair of quantile partitioned walks $\mathbf{v} = (v, (t_i)_{i=0}^{\mathcal{L}+1}, \mathcal{P})$, and $\mathbf{v}' = (v, (t'_i)_{i=0}^{\mathcal{L}'+1}, \mathcal{P}')$. We define

$$\widetilde{M}(i) := (i + v(n) - y_k)_+ - (i - y_k)_+, \text{ where } y_k = v(k) - v(k-1). \tag{2.42}$$

Note that, by definition 2.29,

$$\widetilde{M}(i) = M_{\mathbf{v}}(\mathcal{P} + i) = M_{\mathbf{v}'}(\mathcal{P}' + i) \text{ for every } i. \tag{2.43}$$

We prove by induction that $\mathbf{v}$ must equal $\mathbf{v}'$, and therefore that γ is injective on the quantile partitioned walks.

Base case: $t_{\mathcal{P}+1} = t'_{\mathcal{P}'+1} = k$.

Inductive step: We assume that $t_{\mathcal{P}+1-i} = t'_{\mathcal{P}'+1-i} > 0$ for some $i \geq 0$. Then by Lemma 2.36

$$t_{\mathcal{P}-i} = t'_{\mathcal{P}'-i} = \inf\{t \geq 0 : v(t) = t_{\mathcal{P}+1-i} - t + 2\widetilde{M}(-i) - v(t_{\mathcal{P}+1-i})\}.$$

Likewise, if we assume $t_{\mathcal{P}+1+i} = t_{\mathcal{P}'+1+i}$ for some $i \geq 0$ then by Lemma 2.36,

$$t_{\mathcal{P}+2+i} = t'_{\mathcal{P}'+2+i} = \inf\{t \geq 0 : v(t) = t - t_{\mathcal{P}+1+i} + 2\widetilde{M}(i+1) - v(t_{\mathcal{P}+1+i})\}.$$

By induction, $t_{\mathcal{P}+i} = t'_{\mathcal{P}'+i}$ wherever both are defined. Thus there is some greatest index $I \leq 0$ at which these simultaneously reach 0. This I must equal both $-\mathcal{P}$ and $-\mathcal{P}'$. By the same reasoning $\mathcal{L} = \mathcal{L}'$. We conclude that $\mathbf{v} = \mathbf{v}'$. □

Our proof of the bijection theorem suggests an algorithm to invert γ, passing back from a quantile pair to a quantile partitioned walk. This algorithm is stated in Figure 2.16. Along with equation (2.23) and our algorithm `Reconstitution` stated in Figure 2.9, this allows us to invert the quantile transform: given a quantile pair (v, k) of length n, these algorithms identify the unique walk w with $(Q(w), \phi_w^{-1}(n)) = (v, k)$. An example run of the `Saw_Teeth` algorithm is depicted in Figure 2.17.

Theorem 2.3 and Lemma 2.13 have the following corollary.

Corollary 2.37. *For any walk v,*

$$\#\{w \ : \ V(w) = v\} = \#\{w \ : \ Q(w) = v\}. \tag{2.44}$$

Equation (2.44) is a key result as we pass into the continuous-time setting in chapter 3.

2.6 Further problems

The quantile transform of simple random walks has been fully characterized in this chapter. But as for the quantile transforms of other discrete random walks, our only result thus far is Theorem 2.7. A natural next candidate for study would be skip-free random walks.

Let us relax our definition of quantile pairs to allow any sort of finite-length walk.

Conjecture 2.38. *The number of skip-free walks of length n with final value at most n equals the number of quantile pairs (v, k) with v being a skip-free walk of length n with final value at most n.*

This has been computationally verified up to $n = 8$.

Conjecture 2.39. *The map $w \mapsto (Q(w), \phi_w^{-1}(n))$ is a bijection between the set of skip-free left random walks of length n and the set of quantile pairs (v, k) with v being a skip-free walk of length n.*

```
def Saw_Teeth(y,k):
  ## Takes a quantile pair, returns quantile partitioned walk.

  N = len(y)-1
  kthStep = y[k]-y[k-1]
  PminusT = -1*kthStep  ## preterminal lvl - terminal lvl
  PminusS = PminusT + y[N]  ## preterminal lvl - start lvl

  t = [k]  # A list of saw teeth.

  while (t[0] != 0):  ## Finds saw teeth, going left from k
    M = max(PminusS + len(t) - 1,0) - max(PminusT + len(t) - 1,0)
    j = 0  ## must find inf(j>=0 : y[j] = t[0]-j-y[t[0]]+2M)
    while y[j] != t[0] - j - y[t[0]] + 2*M:
      j += 1
    t = [j] + t

  P = len(t)-2  ## preterminal lvl
  T = P - PminusT  ## terminal lvl
  S = P - PminusS  ## start lvl

  while (t[-1] != N):  ## Finds saw teeth, going right from k
    M = max(len(t) - 1 - S,0) - max(len(t) - 1 - T,0)
    j = t[-1]+1  ## find inf(j > t[-1] : y[j] = j-t[i]-y[t[i]]+2M)
    while y[j] != j - t[-1] - y[t[-1]] + 2*M:
      j += 1
    t += [j]

  return [y,t,P]
```

Figure 2.16: Working Python code to invert γ from quantile pairs to quantile partitioned walks.

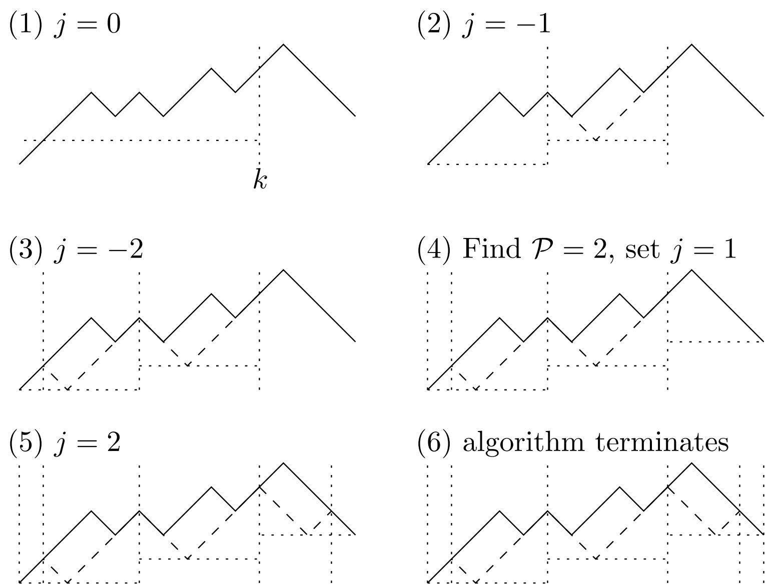


Figure 2.17: Illustrated example of the Saw_Teeth algorithm.

This has been computationally verified for skip free walks of length up to 6 with final value up to 6. Lemma 1.17 should be helpful in resolving the latter conjecture.

Beyond skip-free walks, we might study the quantile transform of general walks of finite length on the reals. Moreover, we should examine the quantile transform of upward-transient walks of infinite length; even in the simple walk case, we have no results in this vein.

Chapter 3

The quantile transform of Brownian motion

In this chapter we demonstrate that the quantile transforms of certain simple random walks converge to an expression involving Brownian local times. This allows us to pass the connection to the Vervaat transform in Corollary 2.37 through the limit to give a novel generalization of Jeulin's description of Brownian local times.

Let $(B(t),\ t \in [0,1])$ denote standard real-valued Brownian motion. Let $(B^{\mathrm{br}}(t),\ t \in [0,1])$ denote a standard Brownian bridge and $(B^{\mathrm{ex}}(t),\ t \in [0,1])$ a standard Brownian excursion – see, for example, Mörters and Peres[MP10] or Billingsley[Bil68] for the definitions of these processes. When we wish to make statements or definitions which apply to all three of B, B^{br}, and B^{ex}, we use $(X(t),\ t \in [0,1])$ to denote a general pick from among these. Finally, we use '$\stackrel{d}{=}$' to denote equality in distribution.

Definition 3.1. We use $\ell_t(a)$ to denote an a.s. jointly continuous version of the *(occupation density) local time* of X at level a, up to time t. That is

$$\ell_t(a) = \lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} \int_0^t \mathbf{1}\{|X(s) - a| < \epsilon\} ds. \tag{3.1}$$

The existence of an a.s. jointly continuous version is well known, and is originally due to Trotter[Tro58] in the case $X = B$; the generalization to B^{br} and B^{ex} is well known. We often abbreviate

$$\ell(a) := \ell_1(a).$$

Let $F(a)$ denote the *cumulative distribution function* (or *CDF*) *of occupation measure*,

$$F(a) := \int_{-\infty}^a \ell(y) dy = \mathrm{Leb}\{s \in [0,1] : X(s) \leq a\}. \tag{3.2}$$

By the continuity of X, the function F is strictly increasing in between its escape from 0 and arrival at 1. Thus we may define a continuous inverse of F, the *quantile function of* *occupation measure*,

$$A(s) := \inf\{a : F(a) > s\} \text{ for } s \in [0,1), \tag{3.3}$$

and we extend this function continuously to define $A(1) := \max_{s\in[0,1]} X(s)$. Note that, analogously, $A(0) = \min_{s\in[0,1]} X(s)$.

Recall that for a walk w, the value $Q(w)(j)$ is the sum of increments from w which appear at the j lowest values in the path of w. Heuristically, at least, the continuous-time analogue to this is the formula

$$Q(X)(t) = \int_0^1 \mathbf{1}\{X(s) \leq A(t)\} dX(s). \tag{3.4}$$

This formula would define $Q(X)(t)$ as the sum of bits of the path of X which emerge from below a certain threshold – the exact threshold below which X spends a total of time t. But it is unclear how to make sense of the integral: it cannot be an Itô integral because the integrand is not adapted. Perkins[Per82, p. 107] allows us to make sense of this and similar integrals. We quote Tanaka's formula:

$$\int_0^1 \mathbf{1}\{X(s) \leq a\} dX(s) = \frac{1}{2}\ell(a) + (a)_+ - (a - X(1))_+, \tag{3.5}$$

where $(c)_+$ denotes $\max(c, 0)$. For more on Tanaka's formula see e.g. Karatzas and Shreve[KS91, p. 205]. We will discuss Perkins' results further around Theorem 3.23, but for now we mention his technique only as motivation. His paper would have us define

$$\int_0^1 \mathbf{1}\{X(s) \leq A(t)\} dX(s) := \int_{-\infty}^{\infty} \mathbf{1}\{a \leq A(t)\} dJ(a) \tag{3.6}$$

$$= \int_{-\infty}^{\infty} \mathbf{1}\{F(a) \leq t\} dJ(a), \tag{3.7}$$

where $J(a)$ equals the right-hand side of (3.5), which is a semi-martingale with respect to a certain naturally arising filtration. This motivates us in the following definition.

Definition 3.2. The *quantile transform of Brownian motion / bridge / excursion* is

$$Q(X)(t) := \frac{1}{2}\ell(A(t)) + (A(t))_+ - (A(t) - X(1))_+. \tag{3.8}$$

Note that in the bridge and excursion cases this expression reduces to

$$Q(X)(t) := \frac{1}{2}\ell(A(t)). \tag{3.9}$$

The connection to Tanaka's formula echoes the discrete Tanaka's formula (2.21) which arose in the preceding chapter. This formula plays a central role in our proof of the main result of this chapter, which relates the quantile transform to the continuous-time Vervaat transform.

Definition 3.3. Let $\tau_m(X)$ denote the time of the (first) arrival of $(X(t),\ t \in [0,1])$ at its minimum. We suppress the argument X where it is clear from context. The *Vervaat transform* maps X to the process $V(X)$ given by

$$V(X)(t) := \begin{cases} X(\tau_m + t) - X(\tau_m) & \text{for } t \in [0, 1-\tau_m) \\ X(\tau_m + t - 1) + X(1) - X(\tau_m) & \text{for } t \in [1-\tau_m, 1]. \end{cases} \tag{3.10}$$

This is the continuous-time analogue to the discrete Vervaat transform described in section 2.2. This transform should be thought of as partitioning the increments of X into two segments, prior and subsequent to τ_m, and swapping the order of these segments.

Theorem 3.4. *$\tau_m(B^{br})$ is independent of $V(B^{br})$ and has* Uniform$[0,1]$ *distribution (Biane, 1986[Bia86]), and*

$$(V(B^{br})(t),\ t \in [0,1]) \stackrel{d}{=} (B^{ex}(t),\ t \in [0,1]). \qquad \textit{(Vervaat, 1979[Ver79])} \tag{3.11}$$

For discussions of the Vervaat and other related transformations, see [BP94] and references therein. For further extensions of Biane's result, see [Cha99], which also relates to the quantile permutation associated with a random walk and the identities discussed later in section 4.2.

The main theorem of this chapter, stated below, is a continuous-time analogue to Corollary 2.37.

Theorem 3.5. *We have* $(Q(B), B(1)) \stackrel{d}{=} (V(B), B(1))$.

We reduce this theorem to Theorem 3.8 early in section 3.1.

We may use properties of Brownian bridge to give a unique family of distributions for $Q(B)$ and $V(B)$ conditional on $B(1) = a$ that is weakly continuous in a. In the case $B(1) = 0$, Theorem 3.5 specializes to the following.

Theorem 3.6 (Jeulin, 1985)**.** *If ℓ and A denote the local time and the quantile function of occupation measure, respectively, of a Brownian bridge or excursion, then*

$$\left(\frac{1}{2}\ell(A(t)),\ t \in [0,1]\right) \stackrel{d}{=} (B^{ex}(t),\ t \in [0,1]). \tag{3.12}$$

This assertion for Brownian excursions appeared in Jeulin's monograph [Jeu85, p. 264] but without a clear, explicit proof; a proof appears in [BY87, p. 49]. Jeulin's theorem was applied by Biane and Yor[BY87] in their study of principal values around Brownian local times. Aldous[Ald98] made use of a related identity to study Brownian excursion conditioned on its local time profile; and Aldous, Miermont, and Pitman[AMP04], while working in the continuum random tree setting, discovered a version of Jeulin's result for a more general class of Lévy processes. Leuridan[Leu98] and Pitman[Pit99] have given related descriptions of Brownian local times up to a fixed time, as a function of level.

In section 3.1 we prove Theorem 3.5 by demonstrating that the quantile transforms of certain simple random walks strongly converge to that of Brownian motion. This proof appeared in [AFP13]. In section 3.2 we summarize the main results of a preprint by Lupus, Pitman, and Tang[LPT13] describing $Q(B)$, or equivalently, $V(B)$. And in section 3.3 we discuss directions for further work in this area.

3.1 Strong convergence

In proving Theorem 3.5 we will call upon classic limit results relating Brownian motion and its local times to their analogues for simple random walk. The work here falls into the broader scheme of limit results and asymptotics relating random walk local times to Brownian local times. We rely heavily on two results of Knight[Kni62, Kni63] in this area. Much else has been done around local time asymptotics; in particular, Csáki, Csörgő, Földes, and Révész have collaborated extensively, as a foursome and as individuals and pairs, in this area. We mention a small segment of their work: [Rév90, Rév81, CR84, CR85a, CR85b, CCFR09]. See also Bass and Khoshnevisan[BK93b, BK93a] and Szabados and Székeley[SS05].

Definition 3.7. For each $n \geq 1$ let $\tau_n(0) := 0$ and

$$\tau_n(j) := \inf\{t > \tau_n(j-1) : B(t) - B(\tau_n(j-1)) = \pm 2^{-n}\} \text{ for } j \in (0, 4^n]. \tag{3.13}$$

We define a walk

$$S_n(j) := 2^n B(\tau_n(j)) \text{ for } j \in [0, 4^n] \text{ and} \tag{3.14}$$

$$\bar{S}_n(t) := 2^{-n} S_n([4^n t]) \text{ for } t \in [0, 1], \tag{3.15}$$

where we take the square brackets to denote the floor function. From elementary properties of Brownian motion, $(S_n(j),\ j \geq 0)$ is a simple random walk. We call the sequence of walks S_n the *simple random walks embedded in* B. Since we will be dealing with the quantile transformed walk $Q(S_n)$, we define a rescaled version:

$$\overline{Q(S_n)}(t) := 2^{-n} Q(S_n)([4^n t]).$$

Note that $\tau^n_{4^n}$ is the sum of 4^n independent, *Exp*(4^n)-distributed variables. By a Borel-Cantelli argument, the $\tau_n(4^n)$ converge a.s. to 1. So the walks S_n depend upon the behavior of B on an interval converging a.s. to $[0, 1]$ as n increases.

Theorem 3.8. *As n increases, $\overline{Q(S_n)}$ a.s. converges uniformly to $Q(B)$.*

After a few more definitions and a summary of relevant results from the literature, we will quickly reduce Theorem 3.5 reduces to the above limit theorem. Then, the remainder of this section will be dedicated to proving this theorem.

Definition 3.9. We define the *(discrete) local time* of $S_n(j)$ at level $x \in \mathbb{R}$

$$L_n(x) := \sum_{j=0}^{4^n-1} (1-(x-[x]))\mathbf{1}\{S_n(j)=[x]\} + (x-[x])\mathbf{1}\{S_n(j)=[x]+1\}$$

This is a linearly interpolated version of the standard discrete local time. We also require a rescaled version,

$$\bar{L}_n(x) := 2^{-n}L_n(2^n x).$$

Note that for $x \in \mathbb{Z}$ we get

$$L_n(x) = \#\{j \in [0,4^n) : S_n(j) = x\} \text{ and}$$
$$\bar{L}_n(2^{-n}x) = \text{Leb}\{t \in [0,1] \ : \ \bar{S}_n(t) = 2^{-n}x\}.$$

The quantile transform $Q(S_n)$ can be thought of as an interpolated, time-shifted version of the discrete Tanaka's formula; see the remarks around equation (2.21). Previous authors – Szabados and Szekely[SS09, p. 208-9] and references therein – have stated convergence results for this discrete Tanaka's formula. However, these results are not applicable in our situation due to the random time change $A(t)$ that appears in our continuous-time formulae.

We require several limit theorems from the literature, relating simple random walk and its local times to Brownian motion, summarized below.

Theorem 3.10.

$$\bar{S}_n(\cdot) \to B(\cdot) \text{ a.s. uniformly} \qquad \text{(Knight, 1962[Kni62]).} \tag{3.16}$$

$$\begin{aligned} \min_t\{\bar{S}_n(t)\} &\to \min_{t\in[0,1]} B_t \text{ and} \\ \max_t\{\bar{S}_n(t)\} &\to \max_{t\in[0,1]} B_t \end{aligned} \qquad \text{(corollary to above).} \tag{3.17}$$

$$\bar{L}_n(\cdot) \to \ell(\cdot) \text{ a.s. uniformly} \qquad \text{(Knight, 1963[Kni63]).} \tag{3.18}$$

Equation (3.16) is an a.s. variant of Donsker's Theorem, which is discussed in standard textbooks such as Durrett[Dur10] and Kallenberg[Kal02]. Equation (3.17) is a corollary to the Knight result: both max and min are continuous with respect to the uniform convergence metric. The map from a process to its local time process, on the other hand, is not continuous with respect to uniform convergence; thus, equation (3.18) stands as its own result. An elementary proof of this latter result, albeit with convergence in probability rather than a.s., can be found in [Rév81], along with a sharp rate of convergence. Knight[Kni97] gives a sharp rate of convergence under the L^2 norm.

We can now reduce our generalization of Jeulin's theorem to our limit theorem for the quantile transform.

Proof of Theorem 3.5 from Theorem 3.8. Let $\overline{V(S_n)}(t) := 2^{-n}V(S_n)([4^n t])$. Equation (3.16) tells us that

$$\inf\{t \in [0,1] \ : \ \bar{S}_n(t) = \min_s \bar{S}_n(s)\}$$

converges a.s. to the time of the minimum of B. From this observation and (3.16) we see that $V(\bar{S}_n)$ converges a.s. uniformly to $V(B)$; this was proved for Brownian bridge by Vervaat[Ver79]. By Corollary 2.37 we have $\overline{Q(S_n)} \stackrel{d}{=} \overline{V(S_n)}$, and by Theorem 3.8 the $\overline{Q(S_n)}$ converge in distribution to $Q(B)$. Thus $Q(B) \stackrel{d}{=} V(B)$ as desired. □

Definition 3.11. The *cumulative distribution function (CDF) of occupation measure* for S_n is given by

$$F_n(y) := \int_{-\infty}^{y} L_n(x)dx \text{ and}$$
$$\bar{F}_n(y) := 4^{-n} F_n(2^n y) = \int_{-\infty}^{y} \bar{L}_n(x)dx.$$

Compare these to F, the CDF of occupation measure for B, defined in equation (3.2). Also note that at integers k,

$$\begin{aligned} F_n(k) &= \sum_{j<k} L_n(j) + \frac{1}{2} L_n(k) \\ &= \#\{i \in [0, 4^n) : S_n(i) < k\} + \frac{1}{2}\#\{i \in [0, 4^n) : S_n(i) = k\}. \end{aligned} \tag{3.19}$$

Equations (3.18) and (3.17) have the following easy consequence.

Corollary 3.12. *As n increases the $\bar{F}_n$ a.s. converge uniformly to F.*

Because Brownian motion is continuous and simple random walk cannot skip levels, the CDFs F and F_n are strictly increasing between the times where they leave 0 reach their maxima, 1 or 4^n respectively. This admits the following definitions.

Definition 3.13. We define the *quantile functions of occupation measure*

$$A_n(t) := F_n^{-1}(t) \text{ for } t \in (0, 4^n), \text{ and}$$
$$\bar{A}_n(t) := \bar{F}_n^{-1}(t) \text{ for } t \in (0, 1),$$

and we extend these continuously to define $A_n(0)$, $\bar{A}_n(0)$, $A_n(4^n)$ and $\bar{A}_n(1)$.

Compare these to A defined in equation (3.3).

Lemma 3.14. *As n increases the $\bar{A}_n$ a.s. converge uniformly to A.*

Proof. In passing a convergence result from a function to its inverse it is convenient to appeal to the Skorokhod metric. For continuous functions, uniform convergence on a compact interval $I \subset \mathbb{R}$ is equivalent to convergence under the Skorohod metric (see [Bil68]). Let i denote the identity map on I, let $|| \cdot ||$ denote the uniform convergence metric, and let Λ

denote the set of all increasing, continuous bijections on I. The Skorokhod metric may be defined as follows:

$$\sigma(f,g) := \inf_{\lambda\in\Lambda} \max\{||i-\lambda||,\ ||f-g\circ\lambda||\}. \tag{3.20}$$

Thus, it suffices to prove a.s. convergence under σ.

Fix $\epsilon > 0$. By the continuity of A, there is a.s. some $0 < \delta < \epsilon$ sufficiently small so that

$$A(\delta) - \min_{[0,1]} B(t) < \epsilon \text{ and } \max_{[0,1]} B(t) - A(1-\delta) < \epsilon.$$

And by Equation (3.17) and Corollary 3.12 there is a.s. some n so that, for all $m \geq n$,

$$\begin{aligned} \min_{t\in[0,1]} \bar{S}_m(t) &< A(\delta); \\ \max_{t\in[0,1]} \bar{S}_m(t) &> A(1-\delta); \text{ and} \\ \sup_y |\bar{F}_m(y) - F(y)| &< \epsilon. \end{aligned}$$

We show that $\sigma(\bar{A}_n, A) < 3\epsilon$.

We seek a time change $\lambda : [0,1] \to [0,1]$ which is close to the identity and for which $\bar{A}_n \circ \lambda$ is close to A. Ideally, we would like to define $\lambda = \bar{F}_n \cdot A$ so as to get $\bar{A}_n \circ \lambda = A$ exactly. But there is a problem with this choice: because $\bar{S}_n$ and B may not have the exact same max and min, $\bar{F}_n \circ A$ may not be a bijection on $[0,1]$. We turn this map into a bijection by manipulating its values near 0 and 1.

We define the random time change on $[0,1]$

$$\lambda(t) := \begin{cases} \frac{t}{\delta}\bar{F}_n(A(\delta)) & \text{for } 0 \leq t < \delta \\ \bar{F}_n(A(t)) & \text{for } \delta \leq t \leq 1-\delta \\ 1 + \frac{1-t}{\delta}(\bar{F}_n(A(1-\delta)) - 1) & \text{for } 1-\delta < t \leq 1. \end{cases} \tag{3.21}$$

By our choice of n we get

$$\bar{F}_n(A(\delta)) > 0 \text{ and } \bar{F}_n(A(1-\delta)) < 1.$$

Thus λ is a bijection.

We now show that it is uniformly close to the identity. Since $t = F(A(t))$, our conditions on n give us

$$||\lambda(t) - t||_{t\in[\delta,1-\delta]} \leq ||barF_n(A(t)) - F(A(t))|| < \epsilon.$$

For t near 0

$$||\lambda(t) - t||_{t<\delta} \leq |\lambda(\delta) - F(A(\delta))| < \epsilon,$$

and likewise for $t > 1-\delta$.

Next we consider the difference between A and $\bar{A}_n \circ \lambda$. These are equal on $[\delta, 1-\delta]$. For $t < \delta$ we get

$$A(t) \in [(\min_t B_t),\ A(\delta)] \text{ and } \bar{A}_n \circ \lambda(t) \in [(\min_t \bar{S}_n(t)),\ A(\delta)].$$

By our choices of n and δ, the lower bounds on these intervals both lie within 2ϵ of δ. A similar argument works for $t > 1-\delta$. Thus $A(t)$ lies within 2ϵ of $\bar{A}_n \circ \lambda(t)$.

We conclude that $\sigma(\bar{A}_m, a) < 3\epsilon$ for $m \geq n$. □

For our purpose, the important consequence of the preceding lemma is the following.

Corollary 3.15. *As n increases the $\bar{L}_n \circ \bar{A}_n$ a.s. converge uniformly to $\ell_1 \circ A$.*

General results for convergence of randomly time-changed random processes can be found in Billingsley[Bil68], but in the present case the proof of Corollary 3.15 from equation (3.18) and Lemma 3.14 is an elementary exercise in analysis, thanks to the a.s. uniform continuity of ℓ.

We now make use of the up- and down-crossing counts described in Definition 2.17, and of the saw teeth in Definition 2.24. For our present purpose it is convenient to re-index these sequences.

Definition 3.16. Let $m_n = \min_{j<4^n} S_n(j)$. For each $i \geq m_n$ we define u_i^n to be the number of up-steps of S_n which go from the value i to $i+1$. Likewise, let d_i^n denote the number of down-steps of S_n from value i to $i-1$. Finally, let

$$t_i^n = \sum_{j<i} (u_j^n + d_j^n). \tag{3.22}$$

We call these quantities *up-* and *down-crossing counts* and *saw teeth.*

Note the strict inequality in the bound on summation index j in the definition of m_n.

Comparing the sequences $(u_i^{S_n})$ and $(t_i^{S_n})$ of Definitions 2.17 and 2.24 with the sequences (u_i^n) and (t_i^n) defined above, we have

$$u_i^n = u_{i+m_n}^{S_n} \quad \text{and} \quad t_i^n = t_{i+m_n}^{S_n}.$$

Note that

$$L_n(k) = u_k^n + d_k^n = t_{k+1}^n - t_k^n. \tag{3.23}$$

At saw tooth times, the quantile transform $Q(S_n)$ is uniformly well approximated by a formula based on discrete local time.

Lemma 3.17. *Let A_k^n denote $A_n(t_k^n)$. As n increases the following quantities a.s. vanish uniformly in k:*

(i) $2^{-n}|L_n(k) - 2u_k^n|$,

(ii) $2^{-n}|F_n(k) - t_k^n|$,

(iii) $|F(2^{-n}k) - 4^{-n}t_k^n|$,

(iv) $2^{-n}|A_k^n - k|$, *and*

(v) $2^{-n}\left|Q(S_n)(t_k^n) - \left(\frac{1}{2}L_n(A_k^n) + (A_k^n)_+ - (A_k^n - S_n(4^n))_+\right)\right|$.

Proof. We will show that it suffices to prove the convergence of (i). Indeed, the convergence of (ii) follows from that of (i) by equation (3.19), which gives us

$$F_n(k) = t_n^k + \left(\frac{1}{2}L_n(k) - u_k^n\right) \tag{3.24}$$

for integers k; (iii) then follows by Corollary 3.12. The convergence of (iv) follows from that of (ii) by Lemma 3.14 and the uniform continuity of a. And finally, (v) then follows from the others by the discrete Tanaka formula, equation (2.21). Note that by re-indexing, we have replaced the $\mathcal{S}$ and $\mathcal{T}$ from that formula, which are the start and terminal levels, with 0 and $S_n(4^n)$ respectively, which are the start and terminal values of S_n. Thus, it suffices to prove the convergence of (i).

If we condition on $L_n(k)$ then u_k^n is distributed as $Binomial(L_n(k), \frac{1}{2})$. Our intuition going forward is this: if $L_n(k)$ is large then $(L_n(k) - 2u_k^n)/\sqrt{L_n(k)}$ approximates a standard Gaussian distribution. Throughout the remainder of the proof, for each n let $binom(n)$ denote a $Binomial(n, \frac{1}{2})$ variable that is independent of B (we expand our probability space as necessary). Fix $\epsilon > 0$ and let

$$C_1 = 1 + \max_t |B(t)| \text{ and } C_2 = 1 + \max_x \ell(x).$$

Let M be sufficiently large so that for all $n \geq M$,

$$\mathbf{P}\left\{|2^n C_2 - 2binom(2^n C_2)| > 2^n\epsilon\right\} < \sqrt{2/\pi}\exp(-2^{n-1}\epsilon^2/C_2).$$

Such an M must exist by the central limit theorem and well-known bounds on the tails of the normal distribution. Let $N \geq M$ be sufficiently large so that for all $n \geq N$,

$$\max_t |S_n(t)| < 2^n C_1 \text{ and } \max_x L_n(x) < 2^n C_2.$$

Equations (3.18) (3.17) indicate that N is a.s. finite.

We now apply the Borel-Cantelli Lemma.

$$\begin{aligned}
&\sum_{n>M}\sum_{k}\mathbf{P}\left\{|L_n(k)-2u_k^n|>2^n\epsilon;\ n>N\right\}\\
&\leq \sum_{n>M} 2^{n+1}C_1 \max_k \mathbf{P}\left\{|L_n(k)-2u_k^n|>2^n\epsilon;\ n>N\right\}\\
&< \sum_{n>M} 2C_1 e^n \max_{y\leq 2^n C_2} \mathbf{P}\{|y-2binom(y)|>2^n\epsilon\}\\
&< \sum_{n>M} C_1\sqrt{\frac{8}{\pi}}\exp(n-(2^{n-1}\epsilon^2/C_2))<\infty.
\end{aligned}$$

The claimed convergence follows by Borel-Cantelli. □

Our proof implicitly appeals to the branching process view of Dyck paths. This perspective may be originally attributable to Harris[Har52] and was implicit in the Knight papers [Kni62, Kni63] cited earlier in this section. For more on this perspective and its history, see [Pit99] and the references therein.

In order to prove Theorem 3.8, we must extend the convergence of (v) in the previous lemma to times between the saw teeth. The convergence of (iii) leads to a helpful corollary.

Corollary 3.18. *The sequence* $\min_k |t-4^{-n}t_k^n|$ *a.s. converges to 0 uniformly for* $t\in[0,1]$.

Proof. Since $\min_k t_k^n = 0$ and $\max_k t_k^n = 4^n$, it suffices to prove that $4^{-n}\sup_k(t_k^n - t_{k-1}^n)$ a.s. converges to 0. This follows from: the uniform continuity of F, the uniform convergence of the $\bar{F}_n$ to F asserted in Corollary 3.12, and the uniform vanishing of $|\bar{F}_n(k)-2^{-n}t_k^n|$ asserted in Lemma 3.17. □

We now prove a weak version of Theorem 3.8 before demonstrating the full result.

Lemma 3.19. *Let* Z_n *be the process with*

$$Z_n(t_k^n) = Q(S_n)(t_k^n)$$

for each k*, and which is linearly interpolated in between the points, and let* $\bar{Z}_n$ *be the obvious rescaling. As* n *increases,* $\bar{Z}_n$ *a.s. converges uniformly to* $Q(B)$.

Proof. Let

$$\bar{X}_n(t) := \frac{1}{2}\bar{L}_n(\bar{A}_n(t)) + (\bar{A}_n(t))_+ - (\bar{A}_n(t)-\bar{S}_n(1))_+, \tag{3.25}$$

and let $\bar{Y}_n$ denote the process which equals $\bar{X}_n$ at the (rescaled) saw teeth $4^{-n}t_k^n$ and is linearly interpolated between these times. We prove the lemma by showing that the following differences of processes go to 0 uniformly as n increases: (i) $\bar{X}_n - Q(B)$, (ii) $\bar{Y}_n-\bar{X}_n$, and (iii) $\bar{Z}_n-\bar{Y}_n$.

The uniform vanishing of (i) follows from equations (3.16) and (3.18), Lemma 3.14, and Corollary 3.15. That of (iii) is equivalent to item (v) in Lemma 3.17. Finally, each of the three terms on the right in equation (3.25) converge uniformly to uniformly continuous processes, so by Corollary 3.18, $(\bar{Y}_n - \bar{X}_n)$ a.s. vanishes uniformly as well. □

Before the technical work of extending this lemma to a full proof of Theorem 3.8 we mention a useful bound. For a simple random walk bridge $(D(j), j \in [0, 2n])$,

$$\mathbf{P}(\max_{j\in[0,2n]} |D(j)| \geq c\sqrt{2n}) \leq 2e^{-c^2}. \tag{3.26}$$

This formula may be obtained via the reflection principle and some approximation of binomial coefficients; we leave the details to the reader. The Brownian analogue to this bound appears in Billingsley[Bil68, p. 85]:

$$\mathbf{P}(\sup_{t\in[0,1]} |B^{\mathrm{br}}(t)| > c) \leq 2e^{-2c^2}. \tag{3.27}$$

We now complete our proof that $\overline{Q(S_n)}$ converges to $Q(B)$, which will in turn complete our proof of Theorem 3.5.

Proof of Theorem 3.8. Let Z_n and $\bar{Z}_n$ be as in Lemma 3.19. After that lemma it suffices to prove that $(\overline{Q(S_n)} - \bar{Z}_n)$ vanishes uniformly as n increases. By definition, this difference equals 0 at the saw teeth. Moreover, we deduce from Corollary 2.21 and Theorem 2.31 that conditional on Z_n, the walk $Q(S_n)$ is a simple random walk conditioned to equal Z_n at the saw teeth t^n_k and with some constraints, coming from the Crossings property, on its $(t^n_k)^{\text{th}}$ steps.

We must bound the fluctuations of $Q(S_n)$ in between the saw teeth. Heuristic arguments suggest that these fluctuations ought to have size on the order of $2^{n/2}$; we need only show that they grow uniformly slower than 2^n. We prove this via a Borel-Cantelli argument. There are many ways to bound the relevant probabilities of "bad behavior;" we proceed with a coupling argument.

For each (n, k) for which t^n_k is defined – i.e. with $k \in [\min S_n,\ \max S_n]$ – we define several processes and stopping times. These objects appear illustrated together in figure 3.1. First, for $j \in [0, L_n(k) - 1]$ we define

$$\begin{aligned}\hat{W}^n_k(j) &:= Q(S_n)(t^n_k + j) - Q(S_n)(t^n_k) \text{ and}\\ \check{W}^n_k(j) &:= Q(S_n)(t^n_k + j) - Q(S_n)(t^n_{k+1} - 1).\end{aligned}$$

Recall from equation (3.23) that $L_n(k)$ is the difference between consecutive saw teeth. We only define these walks up to time $L^n_k - 1$ so as to sidestep issues around constrained final increments and the Bookends property. Observe that

$$\max_{j\in[t^n_k, t^n_{k+1}]} |Q(S_n)(j) - Z_n(j)| \leq 1 + \max_{j\in[0,L_n(k)-1]} \{|\hat{W}^n_k(j)|, |\check{W}^n_k(j)|\},$$

so it suffices to bound the fluctuations of the $\hat{W}$ and $\check{W}$.

We further define

$$\Delta_k^n := Q(S_n)(t_{k+1}^n - 1) - Q(S_n)(t_k^n)$$

Observe that

$$\begin{aligned} \hat{W}_k^n(0) = 0, \text{ and } \hat{W}_k^n(L_n(k) - 1) = \Delta_k^n; \text{ whereas} \\ \check{W}_k^n(0) = -\Delta_k^n, \text{ and } \check{W}_k^n(L_n(k) - 1) = 0. \end{aligned} \tag{3.28}$$

If $L_n(k)$ is an odd number then we may define a simple random walk bridge D_k^n that has random length $L_n(k) - 1$ but is otherwise independent of S_n (we enlarge our probability space as necessary to accommodate these processes). In the next paragraph we deal with the case where $L_n(k)$ is even. Let

$$\begin{aligned} \hat{T}_k^n &:= \min\{j \ : \ D_k^n(j) + \Delta_k^n = \hat{W}_k^n(j)\} \text{ and} \\ \check{T}_k^n &:= \max\{j \ : \ D_k^n(j) - \Delta_k^n = \check{W}_k^n(j)\}. \end{aligned}$$

These stopping times must be finite, thanks to the values of $\hat{W}$ and $\check{W}$ observed in (3.28). Finally we define the coupled walks.

$$\hat{D}_k^n(j) = \begin{cases} D_k^n(j) + \Delta_k^n & \text{for } j \in [0, \hat{T}_k^n] \\ \hat{W}_k^n(j) & \text{for } j \in (\hat{T}_k^n, L_n(k) - 1]. \end{cases} \tag{3.29}$$

$$\check{D}_k^n(j) = \begin{cases} \check{W}_k^n(j) & \text{for } j \in [0, \check{T}_k^n] \\ D_k^n(j) - \Delta_k^n & \text{for } j \in (\check{T}_k^n, L_n(k) - 1]. \end{cases} \tag{3.30}$$

Conditional on $L_n(k)$, the $\hat{D}_k^n$ and $\check{D}_k^n$ remain simple random walk bridges, albeit vertically translated. These are illustrated in Figure 3.1.

In the case where $L_n(k)$ is even rather than odd, we modify the above definitions by making D_k^n a bridge to -1 if $\Delta_k^n > 0$ (or 1 respectively if $\Delta_k^n < 0$) instead of 0 and including appropriate '+1's (respectively '-1's) into the definitions of $\hat{T}_k^n$ and $\hat{D}_k^n$ so that the final value of $D_k^n + \Delta_k^n + 1$ (resp. -1) aligns with that of $\hat{W}_k^n$.

Fix $\epsilon > 0$. We may bound the extrema of $\hat{W}_k^n$ and $\check{W}_k^n$ by bounding the extrema of $\hat{D}_k^n$ and $\check{D}_k^n$. In particular, we have the following event inclusions.

$$\begin{aligned} &\{\max_j \{|\hat{W}_k^n(j)|, \ |\check{W}_k^n(j)|\} \geq 2^{n+1}\epsilon\} \\ &\subseteq \{\max_j \{|\hat{D}_k^n(j)|, \ |\check{D}_k^n(j)|\} \geq 2^{n+1}\epsilon\} \\ &\subseteq \{|\Delta_k^n| + 1 \geq 2^n\epsilon\} \cup \{\max_j |D_k^n(j)| \geq 2^n\epsilon\}. \end{aligned} \tag{3.31}$$

First we use previous results from this section to prove that a.s. only finitely many of the Δ_k^n are large. Then we make a Borel-Cantelli argument to do the same for the $\max_j |D_k^n(j)|$.

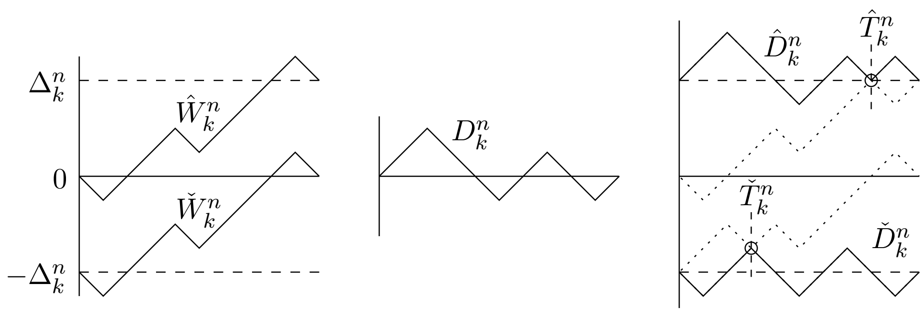


Figure 3.1: Objects from the coupling argument.

By the continuity of $Q(B)$, there is a.s. some $\delta \in (0, \epsilon^2)$ sufficiently small so that

$$\max_{|t-s|<\delta} |Q(B)(t) - Q(B)(s)| < \epsilon.$$

And there is a.s. some N sufficiently large so that for $n \geq N$:

$$\begin{aligned}\sup_j |S_n(j)| &< n2^n, \\ \max_k L_n(k) &< 3^n\delta, \text{ and} \\ \sup_t |\bar{Z}_n(t) - Q(B)(t)| &< \epsilon.\end{aligned}$$

The first two of these bounds follow from the continuity of ℓ and equations (3.16) and (3.18); the third follows from Lemma 3.19. The second and third of these imply that for $n \geq N$,

$$\begin{aligned}|\Delta_k^n| &\leq |Z_n(t_{k+1}^n) - 2^n Q(B)(4^{-n}t_{k+1}^n)| + 2^n|Q(B)(4^{-n}t_{k+1}^n) - Q(B)(4^{-n}t_k^n)| \\ &\quad + |2^n Q(B)(4^{-n}t_k^n) - Z_n(t_k^n)| \\ &\leq 3 \cdot 2^n \epsilon.\end{aligned}$$

So, folding constants into ϵ, there is a.s. some largest n for which any of the $|\Delta_k^n|$ exceed $2^n\epsilon$.

We proceed to our Borel-Cantelli argument to bound fluctuations in the D_k^n.

$$\begin{aligned}&\sum_n \sum_k \mathbf{P}\{\max_j |D_k^n(j)| > 2^n\epsilon;\ n > N\} \\ &\leq \sum_n 2^{n+1} n \max_{|k|<2^n n} \mathbf{P}\{\max_j |D_k^n(j)| > 2^n\epsilon;\ n > N\} \\ &\leq \sum_n 2^{n+1} n \max_{l \leq [3^n\delta]} \mathbf{P}\{\max_j |D_0^n(j)| > 2^n\epsilon \mid L_n(0) = l\} \\ &\leq \sum_n 2^{n+2} n e^{-(\frac{4}{3})^n} < \infty.\end{aligned}$$

The last line above follows from (3.26). We conclude from the Borel-Cantelli Lemma that a.s. only finitely many of the D_k^n exceed $2^n\epsilon$ in maximum modulus. So by the event inequality (3.31), a.s. only finitely many of the W_k^n exceed $2^{n+1}\epsilon$ in maximum modulus. □

3.2 A description of $V(B)$

Previous literature has predominately studied the Vervaat transform as it applies to Brownian bridge, although several papers have generalized this in various directions; see the introduction to [LPT13] for a discussion of references. In order to benefit from the generalization from Jeulin's Theorem 3.6 to Theorem 3.5, we must better understand the quantile or Vervaat transform of general one-dimensional Brownian motion.

Lupus, Pitman, and Tang[LPT13] have studied path properties of $V(B)$. In this section we state several of their main results.

For $a \in \mathbb{R}$ let $B^a = (B^a(t),\ t \in [0,1])$ denote a one-dimensional Brownian bridge to final value a; this is a Brownian motion conditioned to end at value a at time 1 – see Mörters and Peres[MP10] for more details on this process. For $a < 0$ let

$$Z_a := \inf\{t > 0\ :\ V(B^a)(t) = 0\}.$$

Theorem 3.20 ([LPT13])**.** *Let $a < 0$. Then Z^a has probability density*

$$f_{Z^a}(t) = \frac{|a|}{\sqrt{2\pi t(1-t)^3}} \exp\left(-\frac{a^2 t}{2(1-t)}\right), \tag{3.32}$$

and conditional on Z^a, the path $V(B^a)$ may be decomposed into two (conditionally) independent pieces:

- $(V(B^a)(t)\ :\ t \in [0, Z^a])$, *which is a Brownian excursion of length Z^a, and*
- $(V(B^a)(t)\ :\ t \in [Z^a, 1])$, *which is a Brownian first-passage bridge to value a, of length $1 - Z_a$.*

The decomposition in the above theorem should be compared to the decomposition of a Vervaat walk around the helper variable k in a Vervaat pair (v, k) – refer back to Theorem 2.12. We describe this comparison in terms of the discrete Vervaat transform of a simple walk. The transform works by splitting the increments of the walk into two blocks, and then swapping the order of these blocks. But we could further subdivide the walk, as shown in Figure 3.2.

Suppose that the walk w has minimum value m and final value $a < 0$. We decompose the walk into a first-passage bridge to value $m - a$, followed by a first-passage bridge from $m - a$ down to m, and then (a vertical translation of) a non-negative walk from m to a. Label these blocks of increments I, II, and III. The Vervaat transform shuffles these blocks by placing part III first, then I, then II, as in Figure 3.2. The Vervaat bijection theorem, Theorem 2.12, decomposes the transformed walk around the break between parts III and I. Theorem 3.20 instead decomposes the transformed walk around the break between parts II and III. This decomposition is more natural in the sense that it breaks the walk at a stopping time with respect to the filtration induced by the transformed walk; and it is not too difficult to work

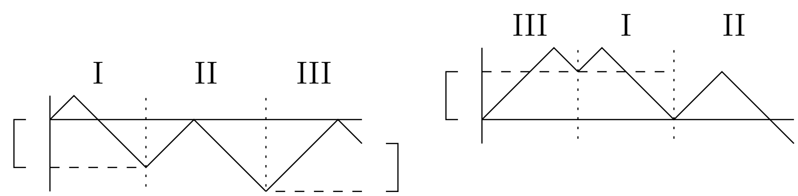


Figure 3.2: The Vervaat transform, in terms of a decomposition of a walk into three blocks of increments.

with, since the path segment subsequent to this stopping time corresponds to a contiguous segment, part II, of the original walk w.

Because, unlike simple random walk, B almost surely visits its minimum only once, we have the following duality relation for V. For $a > 0$,

$$(V(B^a)(t),\ t \in [0,1]) \stackrel{d}{=} (V(B^{-a})(1-t) + a,\ t \in [0,1]). \tag{3.33}$$

For $a > 0$ let

$$\hat{Z}^a := \sup\{t < 1\ :\ V(B^a)(t) = a\}.$$

This duality relation gives the following corollary.

Corollary 3.21 ([LPT13]). *Let $a > 0$. Then $\hat{Z}^a$ has probability density*

$$f_{\hat{Z}^a}(t) = f_{Z^{-a}}(1-t), \tag{3.34}$$

and conditional on $\hat{Z}^a$, the path $V(B^a)$ may be decomposed into two (conditionally) independent pieces:

- $(V(B^a)(t)\ :\ t \in [0, \hat{Z}^a])$, *which is a Bessel-3 bridge of length $\hat{Z}^a$ from 0 to a, and*
- $(V(B^a)(t)\ :\ t \in [\hat{Z}^a, 1])$, *which is a Brownian excursion above value a, of length $1 - \hat{Z}_a$.*

Perhaps most interestingly, the above decompositions lead to two properties of $V(B)$ as a stochastic process.

Theorem 3.22 ([LPT13]). *For $a \in \mathbb{R}$ the processes $V(B^a)$ and $V(B)$ are semi-martingales, and they are not Markov processes with respect to their own induced filtrations.*

There is a simple argument to the effect that $V(B)$ is not Markov with respect to its induced filtration. Fix some $0 < t_1 < t_2 < 1$.

$$\mathbf{P}\{V(B)(1) \geq 0 \mid V(B)(t_2) > 0\} > 0, \text{ but}$$
$$\mathbf{P}\{V(B)(1) \geq 0 \mid V(B)(t_1) = 0,\ V(B)(t_2) > 0\} = 0.$$

As for the semi-martingale property, Lupus et. al. prove this by appealing to the semi-martingale decompositions of the two components of $V(B^a)$ in Theorem 3.20. Alternatively, it is possible to prove the semi-martingale property by appealing to Theorem 3.5 and the following theorem of Perkins. Let $\mathcal{E}_x$ denote the sigma field induced by excursions of B below value x. Roughly speaking, this is

$$\mathcal{E}_x \approx \sigma(B(t) \wedge x,\ t \in [0,1]);$$

see [Per82, p. 2-3] for details.

Theorem 3.23 (Perkins, 1982[Per82])**.** *The process in x*

$$J(x) = \int_0^1 \mathbf{1}\{B(t) \leq x\} dB(t) = \frac{1}{2}\ell(x) + (x)_+ - (x - B(1))_+ \tag{3.35}$$

is a semi-martingale with respect to the filtration $(\mathcal{E}_x)_{x\in\mathbb{R}}$.

From this theorem we obtain

$$Q(B)(t) = J(A(t)) = \int_{-\infty}^{\infty} \mathbf{1}\left\{\text{Leb}\{s\ :\ B(s) < x\} < t\right\} dJ(x) \tag{3.36}$$

is a semi-martingale with respect to that same filtration. The semi-martingale property does, in fact, pass down to the filtration induced by $Q(B)$ [Tan13], although we leave the details to the reader.

Lupus et. al. also give an interesting decomposition of $V(B^a)$ in terms of excursions.

Theorem 3.24 ([LPT13])**.** *Let $a < 0$. Let $X = (X(t),\ t \in [0,1])$ be a reflecting Brownian bridge conditioned to have local time $|a|$ at value 0, and let $U \sim$ Uniform$[0,1]$ be independent of X. Let*

$$G_U := \sup\{t < U\ :\ X(t) = 0\} \quad \textit{and} \quad D_U := \inf\{t > U\ :\ X(t) = 0\}.$$

We define a new process $\hat{X}$ by swapping the excursion of X straddling time U out to the front of the process:

$$\hat{X}(t) = \begin{cases} X(G_U + t) & \textit{for } t \in [0, D_U - G_U], \\ X(t - (D_U - G_U)) & \textit{for } t \in (D_U - G_U, D_U], \textit{ and} \\ X(t) & \textit{for } t \in (D_U, 1]. \end{cases}$$

Let ℓ_t denote the occupation density local time of $\hat{X}$ at level 0, up to time t. Then

$$(\hat{X}(t) - \ell_t,\ t \in [0,1]) \stackrel{d}{=} (V(B^a)(t),\ t \in [0,1]). \tag{3.37}$$

This result should be understood in light of the identity of Bertoin, Chaumont, and Pitman[BCP03] that $(|B^{\mathrm{br}}(t)| - \ell_t,\ t \in [0,1])$ is a Brownian first-passage bridge to a negative value, up to a change of measure in the final value. The type of size-biased selection and rearrangement of excursions described by the above theorem was first studied by Perman, Pitman, and Yor[PPY92].

Finally, we have two moments of $V(B)(t)$.

Proposition 3.25 ([LPT13])**.** *For $t \in [0,1]$,*

$$\mathbf{E}[V(B)(t)] = \sqrt{\frac{8}{\pi}}(\sqrt{t} + \sqrt{1-t} - 1),\ \textit{and} \tag{3.38}$$

$$\mathit{Var}[V(B)(t)] = 3t + \frac{4-8t}{\pi}\mathit{arcsin}\sqrt{t} - \frac{4}{\pi}\sqrt{t(1-t)} \tag{3.39}$$

Lupus et. al. obtain these by computation on densities, conditioning and deconditioning. In section 4.4 we provide a combinatorial derivation of the first of these moments.

3.3 Further problems

Even without knowing the exact distribution of $Q(S)$ for S a non-simple random walk, under suitable conditions it may be possible to prove asymptotic results akin to our limit theorem, Theorem 3.8.

Conjecture 3.26. *Let $X_1, X_2, \cdots$ be a sequence of i.i.d. random variables with mean 0 and finite variance, and for each n, let $S_n = (S_n(j),\ j \in [0,n])$ be a walk with these increments. Then, under the appropriate rescaling, the walks $Q(S_n)$ converge in distribution to $Q(B)$.*

If the X_j also have finite fourth moments and we Skorokhod embed these walks in a single Brownian motion, then it may be possible to obtain strong convergence, as in Theorem 3.8. I have attempted to prove these convergence results, but I've become stuck due to issues with proving convergence of certain Riemann-like sums to $Q(B)(t)$. These sums cannot approximate an Itô integral, since the integral would have a non-adapted integrand. Two papers of Bass and Khoshnevisan[BK93b, BK93a] may be helpful in getting around this issue.

We mention two other directions in which to continue this work. Firstly, in [Ald98], Aldous studies Brownian excursion conditioned on its local time profile at time 1. The paper makes use of a result closely related to Jeulin's Theorem. It may be possible to extend his results to general Brownian motion by replacing his appeal to Jeulin with an appeal to Theorem 3.5. Secondly, in [AMP04], Aldous, Miermont, and Pitman give a version of Jeulin's theorem for a continuum random tree corresponding to a Lévy process. It may be interesting to try to relate their result to Theorem 3.5, or more generally, to relate their work to the quantile transform.

Chapter 4

Related enumerative identities

In this chapter we discuss enumerative identities for simple walks. As in chapter 2, we will say "walk" to mean "simple walk" – a process $(w(j),\ j \in [0,n])$ with $w(0) = 0$ and $|w(j+1) - w(j)| = 1$ for each j. This chapter is themed around an approach to enumerative identities via path decompositions and bijective path transformations.

In section 4.1 we begin with some well-known enumerations and related path transformations and decompositions. In section 4.2 we discuss some of the literature around the quantile permutation, which provides context for chapter 2. In section 4.3 we offer an alternative proof of Corollary 2.14, which enumerates quantile pairs, that doesn't go by way of the Vervaat transform. In section 4.4 we deduce the distribution of the j^{th} increment of the quantile transform of a simple random walk. And in section 4.5 we deduce two more identities related to instruction sets and random walk local times.

4.1 Classic identities, transformations, and decompositions

We begin by enumerating several classes of walks; all of these enumerations are well known and most are discussed in Feller[Fel68, p. 72-77]. A discussion of these and other formulae in this vein may also be found in [EK99]. Most of these classes of walks have been introduced in Definition 2.4.

Proposition 4.1.

$$\#\{\textit{walks of length } n\} = 2^n. \tag{4.1}$$

$$\#\{\textit{bridges of length } n \textit{ to value } b\} = \begin{cases} \binom{n}{\frac{n+b}{2}} & \textit{if } n+b \textit{ is even, or} \\ 0 & \textit{if } n+b \textit{ is odd.} \end{cases} \tag{4.2}$$

$$\#\{\textit{non-negative walks of length } n\} = \binom{n}{\lfloor \frac{n}{2} \rfloor}. \tag{4.3}$$

$$\#\left\{ \begin{array}{c} \textit{first-passage bridges of} \\ \textit{length } n \textit{ to negative values} \end{array} \right\} = \binom{n-1}{\lfloor \frac{n-1}{2} \rfloor}. \tag{4.4}$$

$$\#\{\textit{Dyck paths of length } 2n\} = \frac{1}{n+1}\binom{2n}{n} =: C_n. \tag{4.5}$$

$$\#\{\textit{Positive excursions of length } 2n\} = C_{n-1}. \tag{4.6}$$

We use C_n to denote the n^{th} *Catalan number*, whose formula appears above. Richard Stanley maintains a list of classes of combinatorial objects enumerated by the Catalan numbers on his website[Sta] as an online addendum to exercise 6.19 of his textbook [Sta99, p. 219-229]; as of May 2013, this list has over 200 entries.

The enumerations (4.1) and (4.2) are elementary. Below, we derive (4.3) from (4.2) and derive (4.4) from (4.3) via bijective path transformations. We will obtain (4.5) as a special case of (4.12) in the next section. Then (4.6) follows easily from (4.5).

Derivation of (4.3) *from* (4.2). Let w be a bridge of length n to value 0 or -1. For $j \in [1, n]$ we define

$$y_j = (w(j) - w(j-1)) \cdot \begin{cases} 1 & \text{if } w(j) \geq 0 \\ -1 & \text{otherwise.} \end{cases}$$

Let T_1 be the path transformation that takes w to the walk with increments y_j:

$$T_1(w)(i) = \sum_{j \leq i} y_j.$$

Then the excursions of w above 0 and below -1 correspond to the excursions of $T_1(w)$ above its future-minimum process $m(i) = \min_{j \geq i} T_1(w)(j)$.

- Increments of w from 0 to -1 correspond to last-exit increments of $T_1(w)$ departing from even values.
- Increments of w from -1 to 0 correspond to last-exit increments of $T_1(w)$ departing from odd values.
- Excursions of w above 0 appear as excursions of $T_1(w)$ above even values of its future-minimum process m.

- Excursions of w below -1 appear as excursions of $T_1(w)$ above odd values of its future-minimum process m.

This connection between w and $T_1(w)$ is illustrated in Figure 4.1. It is clear from this that T_1 is a bijection between bridges to 0 or -1 and everywhere non-negative walks. Therefore the enumeration (4.3) follows from (4.2). $\square$

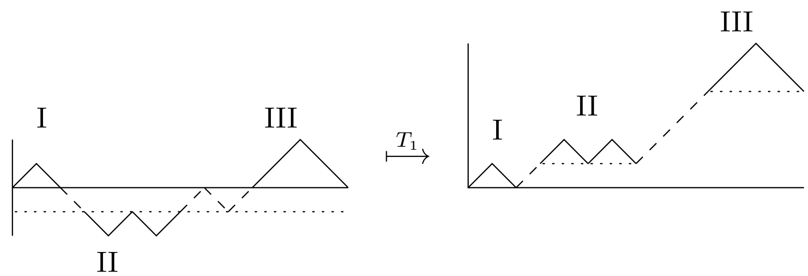


Figure 4.1: A bijective path transformation from bridges to non-negative walks.

Derivation of (4.4) *from* (4.3). Let w be a non-negative walk of length n. For each $j \in [1, n]$ let $y_j = w(n-j) - w(n-j+1)$, and let $y_{n+1} = -1$. Let T_2 be the path transformation that maps w to the walk with increments y_j. Then the property $w(i) \geq w(0)$ for each $i \in [1, n]$ translates to $T_2(w)(n-i) \geq T_2(w)(n)$, and thus $T_2(w)(j) > T_2(w)(n+1)$ for every $j \in [0, n]$.

Geometrically T_2 may be viewed as reflecting the path of w horizontally, sliding it down to start at the origin, and then adding a final down-increment. This is illustrated in Figure 4.2. This transformation T_2 is a bijection from the non-negative walks of length n to the first-passage bridges to negative values of length $n+1$. Therefore the enumeration (4.4) follows from (4.3). $\square$

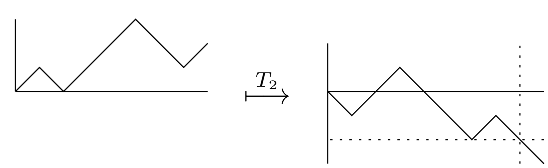


Figure 4.2: A bijective path transformation from non-negative walks to first-passage bridges to negative values.

The enumerations listed in Proposition 4.1 are of particular interest because the types of walks considered are useful as units of path decompositions. Here are a few examples.

- A Dyck path may be decomposed into a first-passage bridge to its maximum and a vertically translated non-positive walk away from the maximum down to 0. This is a discrete version of Williams' decomposition of Brownian excursions; see [RW94] for a discussion of this decomposition or [LG86] for the elementary approach via random walks.

- A walk ending at a positive value may be decomposed at its last exit from 0 into a bridge prior, an up-step, and then a non-negative walk above the value 1.
- A bridge may be decomposed into a sequence of excursions above and below 0.
- A bridge may also be decomposed into a sequence of Dyck paths with alternating signs.

The definition 2.2 of quantile pairs naturally suggests a decomposition of the quantile walks – non-negative walks and first-passage bridges to negative values – at their helper variable times.

Proposition 4.2. *Let F be the map that takes in a walk-index pair (v, k) of length n and returns a triple (v_1, y, v_2) with*

$$v_1 = (v(j),\ j \in [0, k-1]),\ \ y = v(k) - v(k-1),\ \ \textit{and}\ \ v_2 = (v(j+k) - v(k),\ j \in [0, n-k]).$$

Then F is a bijection from quantile pairs to triples with the first entry a non-negative walk, the second entry ± 1, and the last entry a first-passage bridge to a non-positive value (to a negative value except in the case where $k = n$, in which case it is the walk of length 0).

This follows immediately from the definition of the quantile pairs.

4.2 Previous results on the quantile permutation

Similar transformations to the quantile transform have arisen in the study of the fluctuations of random walks. For example, let $x_1, x_2, \cdots$ be a sequence of real numbers, and define $S(0) = 0$ and

$$S(n) := \sum_{j=1}^{n} x_j. \tag{4.7}$$

So the x_i are the increments of the process S. Fix some level $l \geq 0$. We define $S^-(n)$ (and respectively $S^+(n)$) to be the sum of the first n increments of S which originate at or below (resp. strictly above) the value l. That is, an increment x_i of S is an increment of S^- only if $S(i-1) \leq l$. This is illustrated in Figure 4.3; in that example, the increments $x_3,\ x_7,\ x_8$ and x_9 contribute to $S^+(4)$. For the sake of brevity we omit a more formal definition, which may be found in [BCY97]. We call the map $S \mapsto S^-$ the *BCY transform* (with parameter l).

The BCY transform resembles the quantile transform in that it sums increments below some level. But whereas the quantile transform may only be applied to a walk which has

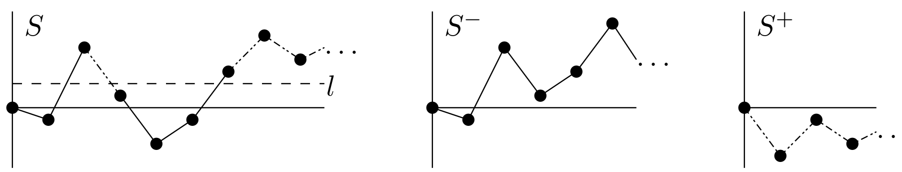


Figure 4.3: The BCY transform.

finite length or is upwardly transient, the BCY transform applies equally well to any walk of infinite length.

There are two major differences between the BCY and quantile transforms. Firstly, in the case of the BCY transform, the process S^- comprises all those increments which appear in S below some previously fixed level l; whereas in the case of the quantile transform, $Q(S)(j)$ comprises (roughly) those increments which appear in S below a variable level that increases with j. Secondly, the increments of S^- appear in the same order in which they appeared in S, whereas the increments of $Q(S)$ appear in order of the value at which they appear in S.

If we suppose that the x_i are i.i.d. random variables then by the strong Markov property, S^- has the same distribution as S [BCY97, Lemma 2]. But this is not the case for $Q(S)$; Theorem 2.3 indicates that for S a simple random walk, $Q(S)$ tend to rise at early times and fall later.

As a further example, the path transformation studied by Chaumont[Cha99] resembles the concatenation of S^- followed by S^+, but with some delicate changes. To define it, we require different notation from that introduced earlier. Recall that the final increment of the walk w has no bearing on ϕ_w. We require a version of the permutation that accounts for this increment. We draw from the notation of Port[Por63] and Chaumont[Cha99]; this notation is used only in this section and nowhere else in the paper.

Definition 4.3. Let the increment sequence $(x_i)_{i=1}^{\infty}$ and the process S be as above. Let $(S^n(j),\ j \in [0,n])$ denote the restriction of S to its n initial increments. For $k \in [0,n]$ we define M^S_{nk} and L^S_{nk} so that

$$(M^S_{n0}, L^S_{n0});\ (M^S_{n1}, L^S_{n1});\ \cdots;\ (M^S_{nn}, L^S_{nn})$$

is the increasing lexicographic reordering of the sequence

$$(S(0),0);\ (S(1),1);\ \cdots;\ (S(n),n)$$

We call the permutation

$$(0,1,\cdots,\ n) \mapsto (L^S_{n0},\ \cdots,\ L^S_{nn})$$

the *quantile permutation of vertices* of S^n (whereas ϕ_{S^n} might be thought of as a quantile permutation *of increments*). We define

$$R^S_{nk} := \#\{i \le L^S_{nk}\ :\ S(i) \le M^S_{nk}\}.$$

We suppress the superscript when it is clear from context which process is being discussed. Both the BCY and Chaumont transforms are motivated by the following theorem.

Theorem 4.4 (Wendel, 1960[Wen60]; Port, 1963[Por63]; Chaumont, 1999[Cha99])**.** *Suppose that $x_1, \cdots, x_n$ are exchangeable real-valued random variables, and let S denote the process with these increments. Fix $k \in [0, n]$ and let S' denote the process*

$$S'(j) = S(k+j) - S(k) \text{ for } j \in [0, n-k].$$

Then

$$\begin{pmatrix} S(n) \\ M^S_{nk} \\ L^S_{nk} \\ R^S_{nk} \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} S(k) + S'(n-k) \\ M^S_{kk} + M^{S'}_{n-k,0} \\ L^S_{kk} + L^{S'}_{n-k,0} \\ L^S_{kk} \end{pmatrix}. \tag{4.8}$$

The identity in the first two coordinates in equation (4.8) is due to Wendel; Port made the satisfying extension of the result to the third coordinate. For more discussion of related results such as Sparre Andersen's Theorem[And53b, And53a] and Spitzer's Combinatorial Lemma[Spi56], see Port[Por63]. Andersen's Theorem and a related result of Port are used later, in section 4.4.

Chaumont made the suggestive extension of (4.8) to the fourth coordinate and presented the first path-transformation-based proof Port's result. Let the x_i and S be as in Theorem 4.4 and fix some $k \in [0, n]$. Chaumont's transformation works by partitioning the increments of S into four blocks.

$$\begin{aligned} I_1 &:= \{i \in [1, L_{nk}] \ : \ S(i-1) \le M_{nk}\}, \\ I_2 &:= \{i \in (L_{nk}, n] \ : \ S(i) < M_{nk}\}, \\ I_3 &:= \{i \in [1, L_{nk}] \ : \ S(i-1) > M_{nk}\}, \text{ and} \\ I_4 &:= \{i \in (L_{nk}, n] \ : \ S(i) \ge M_{nk}\} \end{aligned}$$

The *Chaumont transform* sends S to the process $\tilde{S}$ whose increments are the x_i with $i \in I_1$, followed by those with $i \in I_2$, then I_3, and finally I_4, with the increments within each block arranged in order of increasing index. Details may be found in [Cha99, p. 3-4]. This transformation is illustrated in Figure 4.4, in which increments belonging to I_1 and I_2 are shown as solid lines, whereas those belonging to I_3 and I_4 are shown as dotted.

If S has exchangeable random increments then S and $\tilde{S}$ have the same distribution; as with the BCY transform, this presents a marked difference from the quantile transform. Chaumont demonstrates that if we substitute $\tilde{S}$ for S on the right-hand side of equation (4.8) then we get identical equality, rather than identity in law.

Theorem 4.4 admits various continuous-time versions. Before stating some of these, we state a loose continuous-time analogue to the quantile permutation, due to Chaumont[Cha00].

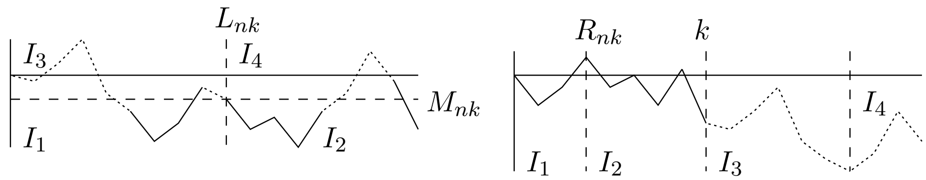


Figure 4.4: On the left a process S and on the right its Chaumont transform.

Definition 4.5. For $(X(t),\ t \in [0,1])$ a continuous, real-valued stochastic process with continuous local time, as in equation (3.1), and let A be its quantile function of occupation measure, as in equation (3.3). We define

$$m_X(s) := \inf\{t \in [0,1]\ :\ \ell_t(A(s)) > U\ell_1(A(s))\} \text{ for } s \in (0,1), \tag{4.9}$$

where U is an independent *Uniform*$[0,1]$ random variable. We define $m_X(0)$ to equal the minimum value visited by X, and $m_X(1)$ to equal the maximum.

Note that $X(m_s^X) = A(s)$; this corresponds to our definition in the previous chapter $A_w(j) = w(\phi_w(j))$. Despite this, the analogy between m and the quantile permutation is flawed because m requires additional randomization in its definition. But there can be no bijection from $[0,1]$ to itself with all of the properties we would want in a quantile permutation; so we must settle for m.

Theorem 4.6. *Let $(X(t),\ t \in [0,1])$ be a Lévy process, and let A be the quantile function of its occupation measure, as in equation* (3.3)*. Fix $T \in [0,1]$ and define*

$$X' := (X(t),\ t \in [0,T]) \text{ and } X'' := (X(t+T) - X(T),\ t \in [0,1-T]).$$

Then

$$(X(1),\ A(T)) \stackrel{d}{=} (X'(T) + X''(1-T),\ \sup_{t\in[0,T]} X'(t) + \inf_{t\in[0,1-T]} X''(t)) \tag{4.10}$$

(Dassios, 1996[Das95, Das96]). If X is Brownian bridge plus drift, then

$$\begin{pmatrix} X(1) \\ A(T) \\ m(T) \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} X'(T) + X''(1-T) \\ \sup_{t\in[0,T]} X'(t) + \inf_{t\in[0,1-T]} X''(t) \\ m_{X'}(T) + m_{X''}(0) \end{pmatrix}. \tag{4.11}$$

(Chaumont, 2000[Cha00]).

Various path transformation-based proofs of (4.10) were obtained by Embrechts, Rogers, and Yor[ERY95] in the Brownian case and by Bertoin et. al.[BCY97] in the Lévy case. Chaumont proved (4.11) with a continuous-time analogue to the Chaumont transform described above. These results have applications to finance in the pricing of Asian options. For a discussion of these applications see Dassios[Das95, Das96, Das05] and references therein.

4.3 Enumeration of quantile pairs

In this section we give an alternate proof – one not dependent on the Vervaat bijection theorem – that there are as many quantile pairs (v, k) in which v has u up-steps and d-down steps as there are walks with u up-steps and d down-steps. We begin with notation.

Let $q(u, d)$ denote the number of quantile pairs (v, k) in which v has exactly u up-steps and d down-steps. For $u \geq d$ let $walk_+(u, d)$ denote the number of everywhere non-negative walks with u up-steps and d down-steps. For $u \neq d$ let $fpb(u, d)$ denote the number of first-passage bridges with u up-steps and d down-steps.

The following two formulae are well known and can be found in Feller[Fel68, p. 72-77] and in Eğecioğlu and King[EK99].

Proposition 4.7.

$$walk_+(u, d) = \binom{u+d}{u} - \binom{u+d}{u+1} \text{ for } u \geq d, \text{ and} \tag{4.12}$$

$$fpb(u, d) = \binom{u+d-1}{u \wedge d} - \binom{u+d-1}{(u \wedge d) - 1} \text{ for } u \neq d, \tag{4.13}$$

where we define $\binom{n}{-1} = \binom{n}{n+1} = 0$.

Our derivation of (4.4) from (4.3) in the previous section also suffices to derive (4.13) from (4.12), so we need only prove (4.12). We follow the approach of Eğecioğlu and King[EK99].

Proof. Suppose $u \geq d$. The right-hand side of (4.12) counts all walks with the appropriate number of up- and down-steps and then subtracts off some of these walks. It suffices to prove that $\binom{u+d}{u+1}$ is the number of walks with u up-steps and d down-steps that at some time pass below 0. Let A denote the set of such walks and let B denote the set of walks with $u+1$ up-steps and $d-1$ down-steps. It now suffices to establish a bijection from A to B.

Take $w \in A$ and let k be the time of its first visit to -1. We define $y_j = w(j-1) - w(j)$ for $j \in [1, k]$ and $y_j = w(j) - w(j-1)$ for $j \in (k, u+d]$. Let T_3 be the path transformation the maps w to the walk with increments y_j. Then $T_3(w) \in B$ and has its first-passage to 1 at time k. Since $u+1 > d-1$, it follows that every walk in B must eventually visit 1. This transformation T_3 is a bijection from A to B. □

Note that in the case $u = d$, equation (4.12) gives a count from Dyck paths. This expression easily yields the formula for the Catalan numbers C_n given by (4.5).

Moving on, we will require a version of the Cycle lemma.

Lemma 4.8 (Cycle lemma, Dvoretzky and Motzkin, 1947[DM47])**.** *A uniformly random first-passage bridge to some level $-b$, with $b > 0$, may be decomposed into b consecutive, exchangeable random first-passage bridges to level -1. If we condition on the lengths of these first-passage bridges then they are independent and uniformly distributed in the sets of first-passage bridges to -1 of the appropriate lengths.*

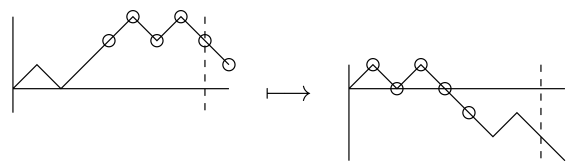


Figure 4.5: A duality relationship between positive and negative quantile walks.

Versions of this lemma have been rediscovered many times. For more on this topic see [DZ90] and [Pit98, p. 172-3] and references therein.

Finally, we require the following duality formula.

Lemma 4.9. *For any non-negative integers u and d,*

$$q(u, d+1) - \binom{u+d}{d+1} = q(d, u+1) - \binom{u+d}{u+1}. \tag{4.14}$$

Proof. The formula is trivial in the case $u = d$. Moreover, it suffices to prove the formula in the case $u > d$, since the case $u < d$ follows by swapping variables.

We define a bijective path transformation T_4 that transforms a non-negative walk ending in a down-step to a first-passage bridge down. This transformation offers a duality between two classes of quantile pairs.

Let w be a non-negative walk that ends in a down-step with a total of u up-steps and $d+1$ down-steps. For $j \in [1, u+d]$ let $y_j = w(u+d-j) - w(u+d+1-j)$ and define $y_{u+d+1} = -1$. Define T to be the path transformation that sends v to the walk with increments y_j. Then $T_4(v)$ has d up-steps and $u+1$ down-steps. This transformation is very similar to that used in the previous section to deduce (4.4) from (4.3). This transformation is illustrated in Figure 4.5.

Fix $u > d$. The transformation T_4 bijectively maps: (1) non-negative walks that end in down-steps and take u up-steps and $d+1$ down-steps to (2) first-passage bridges that take d up-steps and $u+1$ down-steps. This map has the additional property that the final excursion of v above its final value corresponds to the initial Dyck path of $T_4(v)$ above 0, prior to its first visit to -1; see Figure 4.5. Therefore v belongs to exactly one more quantile pair than $T_4(v)$ does:

$$\#\{k : (v, k) \text{ is quantile}\} = \#\{k : (T_4(v), k) \text{ is quantile}\} + 1. \tag{4.15}$$

This gives the following identity for $u > d$:

$$q(u, d+1) - \text{walk}_+(u-1, d+1) = q(d, u+1) + \text{fpb}(d, u+1). \tag{4.16}$$

The second term on the right corresponds to the "+1" from equation (4.15). The second term on the left accounts for quantile pairs involving non-negative walks that end in up-steps, which have not been included in the domain of our path transformation T_4. Subbing in the known counts (4.12) and (4.13) gives the desired result. $\square$

We now have all of the elements needed to re-prove our enumeration of quantile pairs.

Proposition 4.10. *For any non-negative integers u and d,*

$$q(u,d) = \binom{u+d}{u}. \tag{4.17}$$

Proof. We prove the result in the case $u < d$ and then use equation (4.14) to pass our result to the case where $u \geq d$.

Suppose $u < d$. Let $(W_j,\ j \in [0,n])$ denote a uniform random first passage bridge conditioned to have u up-steps and d down-steps, where u and d are fixed. Let τ denote the first-arrival time of W at -1; this is the random number of quantile pairs to which W belongs. By the Cycle Lemma, W may be decomposed into $d-u$ exchangeable first passage bridges to -1. Thus,

$$\mathbf{E}(\tau) = \frac{u+d}{d-u}.$$

So

$$\begin{aligned}
q(u,d) &= \mathbf{E}(\tau)\mathrm{fpb}(u,d) \\
&= \frac{u+d}{d-u}\left(\binom{u+d-1}{d-1} - \binom{u+d-1}{u-1}\right) \\
&= \frac{u+d}{d-u}\left(\binom{u+d}{d}\frac{d}{u+d} - \binom{u+d}{u}\frac{u}{u+d}\right) = \binom{u+d}{d},
\end{aligned}$$

as desired.

Now suppose $u \geq d$. By equation (4.14) and the previous case

$$\begin{aligned}
q(u,d) &= \binom{u+d}{u+1} - \binom{u+d-1}{u+1} + \binom{u+d-1}{u-1} \\
&= \binom{u+d-1}{u} + \binom{u+d-1}{u-1} = \binom{u+d}{u}.
\end{aligned}$$

$\square$

4.4 An increment of a Q-transformed simple random walk

Let $\Omega = \{-1,1\}^n$ with the discrete sigma-algebra $\mathcal{F}$ and uniform probability measure $\mathbf{P}$. Let $X_1, \cdots, X_n$ be the one-dimensional projection maps, and for $0 \leq a < b \leq n$ let

$$S_a^b := \sum_{i=a+1}^{b} X_i.$$

For $a = b$ we define $S_a^b = 0$; thus

$$S := (S_0^j,\ j \in [0, n])$$

is the simple random walk with increments $(X_i)_{i=1}^n$. Let ϕ denote the quantile permutation associated with S, as in Definition 2.1. Recall that this permutation is determined by $(X_i)_{i=1}^{n-1}$, and it is independent of X_n. For each $i \in [1, n]$ let $\hat{X}_i = X_{\phi(i)}$. In this section we give an exact, combinatorial formula for the distribution of $\hat{X}_i$.

Throughout this section we use the notation

$$h(j) := \binom{j}{[j/2]}.$$

We find it convenient to further define $h(-1) = 1$.

Proposition 4.11. *For $1 < k < n$,*

$$\mathbf{P}\{\hat{X}_k = -1\} = \frac{1}{2} - 2^{-k}h(k-1) + 2^{-(n-k)}h(n-k-1) - 2^{-n}h(k-1)h(n-k-1). \quad (4.18)$$

In the $k = 1$ and $k = n$ cases we get

$$\mathbf{P}\{\hat{X}_1 = -1\} = 2^{-n}h(n-2) \text{ and} \quad (4.19)$$

$$\mathbf{P}\{\hat{X}_n = -1\} = 1 - 2^{-n}h(n-1). \quad (4.20)$$

Before proceeding to the proof, we define some notation and mention two useful results from the literature around the quantile permutation.

For $a \le b$, let

$$N_{a,b}^{+} = \#\{j \in [a,b] \ :\ S_a^j \ge 0\}; \qquad N_{a,b}^{++} = \#\{j \in [a,b] \ :\ S_a^j > 0\};$$
$$N_{a,b}^{-} = \#\{j \in [a,b] \ :\ S_a^j \le 0\}; \text{ and } \ N_{a,b}^{--} = \#\{j \in [a,b] \ :\ S_a^j < 0\}.$$

We will also abuse this notation in the following manner. For $a \le b$ let

$$N_{b,a}^{+} = \#\{j \in [a,b] \ :\ S_j^b \ge 0\},$$

and likewise for N^{++}, N^{-}, and N^{--}.

We adopt the standard convention of denoting an event by writing a condition in braces; for example, instead of writing $\{\omega \in \Omega \ :\ N_{0,k}^{+} = k\}$ we will say $\{N_{0,k}^{+} = k\}$. Similarly, we write $\#\{N_{0,k}^{+} = k\}$ to denote $\#\{\omega \in \Omega \ :\ N_{0,k}^{+} = k\}$, and likewise for denoting probability.

Theorem 4.12 (Andersen's theorem, 1953[And53b, And53a])**.**

$$N_{0,n-1}^{+} \overset{d}{=} \phi(n) \text{ and} \quad (4.21)$$

$$N_{0,n-1}^{--} \overset{d}{=} \phi(1) - 1.$$

Interestingly, this theorem implies that

$$\phi^{-1}(1) \overset{d}{=} \phi(1) \quad \text{and} \quad \phi^{-1}(n) \overset{d}{=} \phi(n); \tag{4.22}$$

the first identity follows from $\phi^{-1}(1) = N^{--}_{0,n-1} + 1$, and the second from the distributional identity

$$\phi^{-1}(n) = N^{+}_{n-1,0} \overset{d}{=} N^{+}_{0,n-1}.$$

The computation in this section is driven by a decomposition due to Port. The following appears as equation 5.3 in [Por63].

Lemma 4.13 (Port's decomposition, 1963)**.**

$$\{\phi(k) = l\} = \bigcup_{j=1\vee(l+k-n)}^{l\wedge k} \{N^{+}_{l-1,0} = j\} \cap \{N^{--}_{l-1,n-1} = k - j\}. \tag{4.23}$$

Port's decomposition follows from the derivation

$$\begin{aligned}\phi^{-1}(l) &= \#\{j \in [0, l-1] \ : \ S_0^j \leq S_0^{l-1}\} + \#\{j \in [l-1, n-1] \ : \ S_0^j < S_0^{l-1}\} \\ &= N^{+}_{l-1,0} + N^{--}_{l-1,n-1}.\end{aligned}$$

This manipulation is illustrated in Figure 4.6.

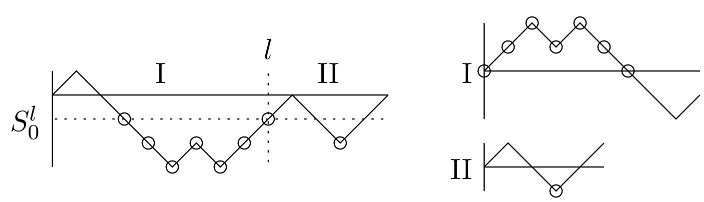


Figure 4.6: An illustration of Port's decomposition. Circled times contribute to $\phi^{-1}(l)$.

We note a particularly useful case of (4.23).

$$\{\phi(n) = l\} = \{N^{+}_{l-1,0} = l\} \cap \{N^{--}_{l-1,n-1} = n - l\}.$$

Thus, for $l \in [1, n]$,

$$\begin{aligned}\mathbf{P}\{\phi(n) = l\} &= \mathbf{P}\{N^{+}_{l-1,0} = l\}\mathbf{P}\{N^{--}_{l-1,n-1} = n - l\} \\ &= 2^{1-n} h(l-1)h(n-l-1)\end{aligned} \tag{4.24}$$

by equations (4.3) and (4.4). (In the $l = n$ case we are appealing to our definition $h(-1) = 1$.) So by Andersen's theorem 4.12,

$$\mathbf{P}\{N^{+}_{0,k-1} = l\} = 2^{1-k} h(l-1)h(k-l-1) \ \text{ for } \ l \in [1, k]. \tag{4.25}$$

Proof of Proposition 4.11. We begin with the relatively easy cases $k = 1$ and $k = n$. In the latter case, $\hat{X}_n = 1$ if and only if S is a first-passage bridge to a positive value. In the former case, $\hat{X}_1 = -1$ if and only if $(S_0^j, j \leq n-1)$ is a first-passage bridge to a negative value and moreover, $X_n = -1$. The given formulae then follow from equation (4.4).

Henceforth, assume $1 < k < n$. Now, let

$$p_k^i := \mathbf{P}\{\phi(k) = i;\ X_i = -1\}, p_k := \mathbf{P}\{\hat{X}_k = -1\}, \quad \text{and} \quad p_k^* := \mathbf{P}\{\hat{X}_k = -1;\ \phi(k) < n\}.$$

So $p_k = \sum_{i=1}^n p_k^i$ and $p_k^* = p_k - p_k^n$.

In the context of $\hat{X}_k$ equaling -1, we can tinker with Port's formula. In particular, given $X_i = -1$ with $i < n$,

$$N_{i-1,n-1}^{--} = N_{i,n-1}^{-} \geq 1.$$

Thus, for $i < n$, (4.23) combined with (4.25) gives us

$$\begin{aligned} p_k^i &= \mathbf{P}\left(\bigcup_{j=1\vee(i+k-n)}^{i\wedge(k-1)} \{N_{i-1,0}^{+} = j\} \cap \{X_i = -1\} \cap \{N_{i,n-1}^{-} = k-j\}\right) \\ &= \sum_{j=1\vee(i+k-n)}^{i\wedge(k-1)} 2^{1-i} h(j-1) h(i-j-1) \cdot \frac{1}{2} \cdot 2^{1+i-n} h(k-j-1) h(n-1-i-k+j) \\ &= 2^{1-n} \sum_{j=1\vee(i+k-n)}^{i\wedge(k-1)} h(j-1)h(i-j-1)h(k-j-1)h(n-1-i-k+j). \end{aligned}$$

Note that we have reduced the upper bound of our union from that in (4.23); in Port's decomposition the index goes as high as $i \wedge k$, but here it is bounded by $i \wedge (k-1)$. This is a consequence of our supposition that $X_i = -1$, in particular because that forces $N_{i-1,n-1}^{--} \geq 1$.

Summing the above terms gives p_k^*.

$$\begin{aligned} p_k^* &= 2^{1-n} \sum_{i=1}^{n-1} \sum_{j=1\vee(i+k-n)}^{i\wedge(k-1)} h(j-1)h(i-j-1)h(k-j-1)h(n-1-i-k+j) \\ &= 2^{1-n} \sum_{j=1}^{k-1} \sum_{i=j}^{j+n-k} h(j-1)h(i-j-1)h(k-j-1)h(n-1-i-k+j). \end{aligned}$$

This is cleaned up via a change of index from i to $x = i - j$.

$$\begin{aligned} p_k^* &= 2^{1-n} \sum_{j=1}^{k-1} \sum_{x=0}^{n-k} h(j-1)h(x-1)h(k-j-1)h(n-k-x-1) \\ &= \frac{1}{2}\left(\sum_{j=1}^{k-1} 2^{1-k} h(j-1)h(k-j-1)\right)\left(\sum_{x=0}^{n-k} 2^{1-n+k} h(x-1)h(n-k-x-1)\right). \end{aligned}$$

We now appeal again to (4.25).

$$p_k^* = \frac{1}{2}\left(\sum_{j=1}^{k-1}\mathbf{P}\{N_{0,k-1}^+ = j\}\right)\left(2^{1-n+k}h(n-k-1) + \sum_{x=1}^{n-k}\mathbf{P}\{N_{0,n-k-1}^+ = x\}\right).$$

In the second of these bracketed sums we've pulled out one term, and the remaining terms add up the probabilities of all possible outcomes for $N_{0,k-1}^+$; the first sum is similar, but instead of having an extra term, this sum is missing a term.

$$p_k^* = \frac{1}{2}(1 - 2^{1-k}h(k-1))(1 + 2^{1-n+k}h(n-k-1)).$$

Finally, we apply (4.23) to find p_k^n.

$$p_k^n = \mathbf{P}\{N_{n-1,0}^+ = k\}\mathbf{P}\{X_n = -1\} = 2^{-n}h(k-1)h(n-k-1).$$

Thus,

$$p_k = \frac{1}{2}(1 - 2^{1-k}h(k-1))(1 + 2^{1-n+k}h(n-k-1)) + 2^{-n}h(k-1)h(n-k-1).$$

Expanding out the product gives equation (4.18). □

We can use this result to give a novel proof of the expected value of $V(B)(t)$, which we have quoted in the previous chapter.

Derivation of (3.38) *via Proposition 4.11.* The result is trivial in the cases $t = 0$ and $t = 1$, so we assume $0 < t < 1$. Let S_n, $\bar{S}_n$, and $\overline{Q(S_n)}$ be as in chapter 3. From Theorems 3.5 and 3.8, it suffices to compute the limit of $2^{-n}\mathbf{E}(Q(S_n)[4^n t])$. From Proposition 4.11,

$$\begin{aligned}\mathbf{E}(Q(S_n)(j)) = 1 &- 2^{-4^n}h(4^n-2)\\ &+ 2\sum_{k=1}^{j} 2^{-k}h(k-1) - 2^{k-4^n}h(4^n-k-1) + 2^{-4^n}h(k-1)h(4^n-k-1).\end{aligned}$$

Stirling's formula gives the approximation

$$h(n) = 2^n\sqrt{\frac{2}{n\pi}}(1 - O(n^{-1})). \tag{4.26}$$

We leave it to the reader to verify that the error term in this expression does not contribute to our limiting formula. We will write $\sim$ to indicate that limit in n of the difference between two expressions is 0. We plug this formula into our sum and drop the first two inconsequential terms to obtain

$$\begin{aligned}
2^{-n}\mathbf{E}(Q(S_n)[4^n t]) &\sim 2^{-n}\sum_{k=1}^{[4^n t]} \sqrt{\frac{2}{(k-1)\pi}} - \sqrt{\frac{2}{(4^n-k-1)\pi}} + \sqrt{\frac{1}{(k-1)(4^n-k-1)\pi^2}} \\
&\sim 2^{-n}\sqrt{2/\pi}\int_0^{[4^n t]} \left(s^{-\frac{1}{2}} - (4^n-s)^{-\frac{1}{2}} + (s(4^n-s)\pi)^{-\frac{1}{2}}\right) ds \\
&\sim \sqrt{2/\pi}\int_0^{t} \left(s^{-\frac{1}{2}} - (1-s)^{-\frac{1}{2}} + 2^{-n}(s(1-s)\pi)^{-\frac{1}{2}}\right) ds \\
&\sim \sqrt{2/\pi}(2\sqrt{s} + 2\sqrt{1-s})|_{s=0}^{t} \\
&= \sqrt{8/\pi}(\sqrt{t} + \sqrt{1-t} - 1),
\end{aligned}$$

as desired. $\square$

4.5 Enumerations based on increment arrays

In this section we offer two more enumerative identities related to increment arrays and simple walk local times.

Definition 4.14. Let w be a walk of length n. Let a denote the minimum of w up to time $n-1$ and b its maximum in that time frame. For each i from a to b, let u_i^w denote the number of up-increments that w takes from value i to $i+1$ (as in Definition 2.17, but indexed by value rather than "level"). The sequence $\underline{u}^w = (u_i^w \ : \ i \in [a,b])$ is called the *up-crossing profile* of w. Likewise, we denote the *down-crossing profile* $\underline{d}^w$, with each d_i^w denoting the number of increments from i to $i-1$.

Theorem 1.7 and Lemma 1.16 have the following combinatorial corollary.

Corollary 4.15. *Let $\underline{u} = (u_i \ : \ i \in [a,b])$ be a finite sequence of non-negative integers with $a \le 0 \le b$, and let $t \in [a-1, b+1]$ with the properties that $u_b = \mathbf{1}\{t = b+1\}$ and $u_i > 0$ for every i with the exceptions of: (i) $t < i < 0$, and possibly (ii) $i = b$. The number of walks with up-crossing profile $\underline{u}$ and terminal value t is*

$$\prod_{j=a+1}^{b-1} \binom{u_j + u_{j-1} - \mathbf{1}\{j>0\} - \mathbf{1}\{j<t\}}{u_j - \mathbf{1}\{j<t\}}. \tag{4.27}$$

Proof. In light of Theorem 1.7, there is a bijective correspondence between simple walks and a certain class of increment arrays $(\mathbf{x}, 0)$ described by Lemma 1.16. In particular, given the up-crossing profile $\underline{u} = (u_i \ : \ i \in [a,b])$ for the walk: (i) the Crossings property indicates that its number of down-steps d_i from each level i must satisfy (1.8), which asserts

$$d_i = u_{i-1} + \mathbf{1}\{i \le 0\} - \mathbf{1}\{i \le t\};$$

and (ii) the Tree property indicates that each stack $\underline{x}_i$ of increments must end with a '1' if $i < t$ or a '-1' if $i > t$. In light of the Tree property, the number of ways to order the increments in stack $\underline{x}_i$ is

$$\binom{u_i + d_i - 1}{u_i - 1} \text{ if } i < t \quad \text{or} \quad \binom{u_i + d_i - 1}{u_i} \text{ if } i > t.$$

Subbing in with the Crossings property and merging these two formulae, the number of ways to order the increments in $\underline{x}_i$ with $i \neq t$ is

$$\binom{u_i + u_{i-1} - 1 + \mathbf{1}\{i \leq 0\} - \mathbf{1}\{i \leq t\}}{u_i - \mathbf{1}\{i < t\}} = \binom{u_i + u_{i-1} - \mathbf{1}\{i > 0\} - \mathbf{1}\{i < t\}}{u_i - \mathbf{1}\{i < t\}}. \tag{4.28}$$

Note that there is a unique way to order the increments in $\underline{x}_a$ and $\underline{x}_b$. By the Crossings property, $d_a = \mathbf{1}\{t = a - 1\}$, and in the case $d_a = 1$, the tree property requires that this solitary down-step be the final increment of $\underline{x}_a$; a corresponding argument indicates the unique ordering of $\underline{x}_b$.

The case $a < i = t < b$ differs from that described by (4.28) in that $\underline{x}_t$ may end with either a $+1$ or a -1. Thus, there are

$$\binom{u_t + d_t}{u_t} = \binom{u_t + u_{t-1} + \mathbf{1}\{t \leq 0\} - 1}{u_t}$$

possible orderings of the increments of $\underline{x}_t$. This and (4.28) give the desired result. □

The Dyck path case $a = t = 0$ of this corollary is well-known from the branching process perspective. It gives a formula for a conditional distribution of a certain Galton-Watson branching process via the connection between the branching process, a random plane tree, and the Dyck path exploration process for said tree. See [Pit98, Pit99] and references therein for more details on these connections. Briefly, let $(Z_i)_{i=0}^{\infty}$ be a Galton-Watson branching process with offspring distribution Geometric($\frac{1}{2}$) and Z_0 also Geometric($\frac{1}{2}$), and let $Z = \sum_i Z_i$. Now let $(u_i)_{i=0}^{b-1}$ be a sequence of positive integers, let $u_b = 0$, and let $n = \sum_i u_i$. Then

$$\mathbf{P}\{Z_i = u_i \text{ for } i \in [0, b];\ Z_b = 0 \mid Z = n\} = \frac{1}{C_n} \prod_{j=1}^{b-1} \binom{u_j + u_{j-1} - 1}{u_j}, \tag{4.29}$$

where C_n is the n^{th} Catalan number, as in equation (4.5).

Another interesting special case of Corollary 4.15 is the bridge case $t = 0$, in which the number of simple walk bridges with up-crossing profile $\underline{u}$ is

$$\prod_{j=a+1}^{b-1} \binom{u_j + u_{j-1} - \mathbf{1}\{j \neq 0\}}{u_j - \mathbf{1}\{j < 0\}}. \tag{4.30}$$

Now we proceed to a second result. For the purposes of the following, we adopt a variation on our discrete local time notation from previous chapters. This result may be found in Révesz [Rév90, Theorem 9.3], although our approach via path transformations appears to be novel.

Theorem 4.16. *For a walk w of length n, let $L_y(w)$ denote the total number of visits of w to value y and let $L^*_y(w)$ denote the number of visits strictly prior to time n. If $n, k, y \in \mathbb{Z}$ satisfy $k \geq 1$ and $n \geq |y| + (2k-1)$ then*

$$\#\{\textit{walks } w \textit{ of length } n \ : \ L_y(w) = L^*_y(w) = k\} = 2^k \binom{n-k}{\lfloor \frac{1}{2}(n+|y|) \rfloor}. \tag{4.31}$$

Moreover,

$$\#\{\textit{walks } w \textit{ of length } n \ : \ L^*_y(w) = k\} = \begin{cases} 2^k \dbinom{n-k}{\frac{1}{2}(n+|y|-1)} & \textit{if } n+y \textit{ is odd} \\ 2^{k+1} \dbinom{n-k-1}{\frac{1}{2}(n+|y|)-1} & \textit{if } n+y \textit{ is even.} \end{cases} \tag{4.32}$$

And if we relax our condition on n to $n \geq |y| + (2k-2)$ then still

$$\#\{\textit{walks } w \textit{ of length } n \ : \ L_y(w) = k\} = \begin{cases} 2^{k-1} \dbinom{n-k+1}{\frac{1}{2}(n+|y|)} & \textit{if } n+y \textit{ is even} \\ 2^{k} \dbinom{n-k}{\frac{1}{2}(n+|y|-1)} & \textit{if } n+y \textit{ is odd.} \end{cases} \tag{4.33}$$

Proof. Equations (4.32) and (4.33) follow from (4.31) and the following:

$$\begin{aligned} &\#\{\text{walks } w \text{ of length } n \ : \ L^*_y(w) = L_y(w) - 1 = k\} \\ &= 2^k \left[\binom{n-k-1}{\frac{1}{2}(n+|y|)-1} - \binom{n-k-1}{\frac{1}{2}(n+|y|)} \right] \mathbf{1}\{n+y \text{ is even}\}. \end{aligned} \tag{4.34}$$

This formula above counts the walks of length n that visit value y exactly k times strictly prior to terminating at value y. The two formulae (4.32) and (4.33) may be obtained by adding (4.31) to (4.34); to obtain (4.33) we must appeal to (4.34) substituting in '$k+1$' for 'k'.

Now it suffices to prove (4.34) and (4.31). We do this with a pair of path transformations.

Let w be a walk of length n with $L^*_y = L_y - 1 = k$. We decompose w at its visits to y into a first passage bridge to value y (only if $y \neq 0$), then k excursions away from y. We transform w into a first-passage bridge $T_5(w)$ of length $n-k$ up to value $|y|+k$, decomposed around its first passages to $|y|, \ |y|+1, \ \cdots, \ |y|+k-1$. Informally, we build up $T_5(w)$ as follows.

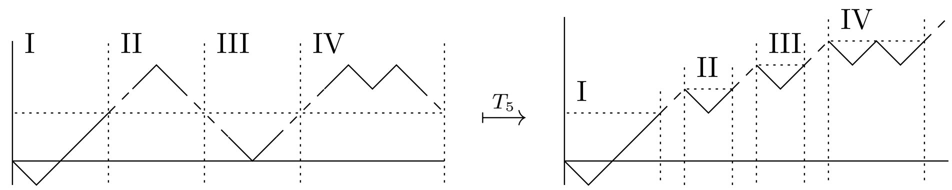


Figure 4.7: A walk with $L_2^* = L_2 - 1 = 3$ transformed into a first passage bridge to value 5.

(i) If $y \neq 0$ then we transform the first-passage of w to y into a first passage to $|y|$ by flipping its sign if necessary.

(ii) For each j from 1 to k we take the j^{th} excursion of w away from y and append it as a Dyck path below level $|y| + j - 1$, prior to a first-passage increment of $T_5(w)$ up to $|y| + j$.

Thus, the excursions of w away from y correspond to the Dyck paths of $T_5(w)$ below its past-maximum process. This transformation is illustrated in Figure 4.7.

More formally, let $t_1 < t_2 < \cdots < t_{k+1} = n$ be the times of visits of w to value y. Let $x_1, \cdots, x_n$ be the increments of w. For $1 \leq i \leq n - k$ we define

$$z_i = \begin{cases} \operatorname{sign}(y)x_i & \text{for } 1 \leq i \leq t_1 \\ -\operatorname{sign}(x_{t_j+1})x_{i+j} & \text{for } t_j - (j-1) < i \leq t_{j+1} - j. \end{cases} \tag{4.35}$$

We define $T_5(w)$ to be the walk with increments $(z_i)_{i=1}^{n-k}$.

This transformation is a surjection from walks of length n with $L_y^* = L_y = k$ to first-passage bridges of length $n - k$ to value $|y| + k$. This map is 2^k-to-1, since it ignores the signs of the excursions of w away from y. This proves (4.34).

Now, to prove (4.31) let w be a walk of length n with $L_y^*(w) = L_y(w) = k$, so $w(n) \neq y$. We decompose w into two segments, prior and subsequent to its final visit to y at time t_k. Let

$$w_1(j) = w(j) \text{ for } j \in [0, t], \text{ and } w_2(j) = |w(j + t_k + 1) - w(t_k + 1)| \text{ for } j \in [0, n - t_k - 1].$$

Since t_k is the time of the final visit of w to y, the absolute value signs inthe definition of w_2 ensure that a walk that would be either everywhere non-negative or everywhere non-positive is guaranteed to be the former. Moreover, w_1 is a walk with $L_y^*(w_1) + 1 = L_y(w_1) = k$. We define $T_6(w)$ to be $T_1^{-1}(w_2)$ appended to the end of $T_5(w_1)$, where T_1 is the transformation described around Figure 4.1. Formally,

$$T_6(w)(j) = \begin{cases} T_5(w_1)(j) & \text{for } j \in [0, t_k - k] \\ (|y| + k - 1) + T_1^{-1}(w_2)(j - t_k + k - 1) & \text{for } j \in [t_k - (k-1), n - k] \end{cases}$$

The transformation T_5 maps w_1 to a first-passage bridge of length $t_k - (k-1)$ to value $|y| + k - 1$, and T_1^{-1} maps w_2 to a bridge of length $n - t_k - 1$ to either 0 or 1, depending on parity. Note that, since all excursions have even length, the parity of t_k equals that of y. So attached, these walks $T_5(w_1)$ and $T_1^{-1}(w_2)$ form a walk of length $n-k$ to final value $|y|+k-\mathbf{1}\{n+y \text{ is odd}\}$, decomposed at it's first passage time to $|y|+k-1$. This map is surjective and, like T_6, it is 2^k-to-1. This proves (4.31). □

4.6 Further problems

The instruction sets of chapter 1 are combinatorially suggestive. It may be interesting to try to mine this perspective for further enumerative results, particularly in relation to local times.

In section 4.4 we use Port's decomposition, Lemma 4.13, to find a formula for $\mathbf{E}[B(t)]$. It would be interesting to try to obtain further moments of $B(t)$ with refined versions of this decomposition.

Also in section 4.4 we explore a connection between Andersen's theorem 4.12 and the quantile permutation and its inverse. Formula (4.22) mentions the most immediate implication of this connection; this ought to be explored further.

Bibliography

[AFP13] S. Assaf, N. Forman, and J. Pitman. The quantile transform of simple walks and brownian motion. Submitted to Probab. Theory Relat. Fields, 2013.

[Ald90] David J. Aldous. The random walk construction of uniform spanning trees and uniform labelled trees. *SIAM J. Discrete Math.*, 3(4):450–465, 1990.

[Ald98] D.J. Aldous. Brownian excursion conditioned on its local time. *Elect. Comm. in Probab*, 3:79–90, 1998.

[AMP04] David Aldous, Grégory Miermont, and Jim Pitman. The exploration process of inhomogeneous continuum random trees, and an extension of jeulins local time identity. *Probab. Theory Relat. Fields*, 129:182–218, 2004.

[And53a] Erik Sparre Andersen. On the fluctuations of sums of random variables. *Math. Scand.*, 1:263–285, 1953.

[And53b] E.S. Andersen. On sums of symmetrically dependent random variables. *Skand. Akturaetid.*, 36:123–138, 1953.

[BCP03] Jean Bertoin, Loïc Chaumont, and Jim Pitman. Path transformations of first passage bridges. *Electron. Comm. Probab.*, 8:155–166 (electronic), 2003.

[BCY97] J. Bertoin, L. Chaumont, and M. Yor. Two chain-transformations and their applications to quantiles. *J. Appl. Probab.*, 34(4):882–897, 1997.

[Bia86] Ph. Biane. Relations entre pont et excursion du mouvement brownien réel. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(1):1–7, 1986.

[Bil68] Patrick Billingsley. *Convergence of probability measures.* John Wiley & Sons Inc., New York, 1968.

[BK93a] Richard F. Bass and Davar Khoshnevisan. Rates of convergence to Brownian local time. *Stochastic Process. Appl.*, 47(2):197–213, 1993.

[BK93b] Richard F. Bass and Davar Khoshnevisan. Strong approximations to Brownian local time. In *Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992)*, volume 33 of *Progr. Probab.*, pages 43–65. Birkhäuser Boston, Boston, MA, 1993.

[Bol98] Béla Bollobás. *Modern graph theory*, volume 184 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.

[BP94] Jean Bertoin and Jim Pitman. Path transformations connecting Brownian bridge, excursion and meander. *Bull. Sci. Math.*, 118(2):147–166, 1994.

[Bro89] Andrei Broder. Generating random spanning trees. In *Thirtieth Annual Symposium on Foundations of Computer Science*, pages 442–447. IEEE, 1989.

[BS13] Matthew Baker and Farbod Shokrieh. Chip-firing games, potential theory on graphs, and spanning trees. *J. Combin. Theory Ser. A*, 120(1):164–182, 2013.

[BY87] P. Biane and M. Yor. Valeurs principales associées aux temps locaux Browniens. *Bulletin des sciences mathématiques*, 111(1):23–101, 1987.

[CCFR09] Endre Csáki, Miklós Csörgő, Antónia Földes, and Pál Révész. Random walk local time approximated by a Brownian sheet combined with an independent Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(2):515–544, 2009.

[Cha99] L. Chaumont. A path transformation and its applications to fluctuation theory. *J. London Math. Soc. (2)*, 59(2):729–741, 1999.

[Cha00] L. Chaumont. An extension of Vervaat's transformation and its consequences. *J. Theoret. Probab.*, 13(1):259–277, 2000.

[CR84] M. Csörgő and P. Révész. Three strong approximations of the local time of a Wiener process and their applications to invariance. In *Limit theorems in probability and statistics, Vol. I, II (Veszprém, 1982)*, volume 36 of *Colloq. Math. Soc. János Bolyai*, pages 223–254. North-Holland, Amsterdam, 1984.

[CR85a] M. Csörgő and P. Révész. On strong invariance for local time of partial sums. *Stochastic Process. Appl.*, 20(1):59–84, 1985.

[CR85b] M. Csörgő and P. Révész. On the stability of the local time of a symmetric random walk. *Acta Sci. Math. (Szeged)*, 48(1-4):85–96, 1985.

[Das95] Angelos Dassios. The distribution of the quantile of a Brownian motion with drift and the pricing of related path-dependent options. *Ann. Appl. Probab.*, 5(2):389–398, 1995.

[Das96] Angelos Dassios. Sample quantiles of stochastic processes with stationary and independent increments. *Ann. Appl. Probab.*, 6(3):1041–1043, 1996.

[Das05] Angelos Dassios. On the quantiles of Brownian motion and their hitting times. *Bernoulli*, 11(1):29–36, 2005.

[DF80] P. Diaconis and D. Freedman. de Finetti's theorem for Markov chains. *Ann. Probab.*, 8(1):115–130, 1980.

[DF84] P. Diaconis and D. Freedman. Partial exchangeability and sufficiency. In *Statistics: applications and new directions (Calcutta, 1981)*, pages 205–236. Indian Statist. Inst., Calcutta, 1984.

[DF91] P. Diaconis and W. Fulton. A growth model, a game, an algebra, Lagrange inversion, and characteristic classes. *Rend. Sem. Mat. Univ. Politec. Torino*, 49(1):95–119 (1993), 1991. Commutative algebra and algebraic geometry, II (Italian) (Turin, 1990).

[DM47] A. Dvoretzky and Th. Motzkin. A problem of arrangements. *Duke Math. J.*, 14:305–313, 1947.

[Dur10] Richard Durrett. *Probability: Theory and Examples.* Cambridge Series in Statistical and Probabilistic Mechanics. Cambridge University Press, New York, 4 edition, 2010.

[DZ90] Nachum Dershowitz and Shmuel Zaks. The cycle lemma and some applications. *European J. Combin.*, 11(1):35–40, 1990.

[EK99] Ömer Eğecioğlu and Alastair King. Random walks and Catalan factorization. In *Proceedings of the Thirtieth Southeastern International Conference on Combinatorics, Graph Theory, and Computing (Boca Raton, FL, 1999)*, volume 138, pages 129–140, 1999.

[Eri96] Kimmo Eriksson. Chip-firing games on mutating graphs. *SIAM J. Discrete Math.*, 9(1):118–128, 1996.

[ERY95] P. Embrechts, L. C. G. Rogers, and M. Yor. A proof of Dassios' representation of the α-quantile of Brownian motion with drift. *Ann. Appl. Probab.*, 5(3):757–767, 1995.

[Fel68] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., New York, 1968.

[Har52] T. E. Harris. First passage and recurrence distributions. *Trans. Amer. Math. Soc.*, 73:471–486, 1952.

[HLM$^+$08] Alexander E. Holroyd, Lionel Levine, Karola Mészáros, Yuval Peres, James Propp, and David B. Wilson. Chip-firing and rotor-routing on directed graphs. In *In and out of equilibrium. 2*, volume 60 of *Progr. Probab.*, pages 331–364. Birkhäuser, Basel, 2008.

[HP10] Alexander E. Holroyd and James Propp. Rotor walks and Markov chains. In *Algorithmic probability and combinatorics*, volume 520 of *Contemp. Math.*, pages 105–126. Amer. Math. Soc., Providence, RI, 2010.

[Jeu85] T. Jeulin. Application de la theorie du grossissement a l'etude des temps locaux browniens. In Th. Jeulin and M. Yor, editors, *Grossissements de filtrations: exemples et applications*, volume 1118 of *Lecture Notes in Mathematics*, pages 197–304. Springer Berlin / Heidelberg, 1985. 10.1007/BFb0075775.

[Kal02] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[Kni62] Frank B. Knight. On the random walk and Brownian motion. *Trans. Amer. Math. Soc.*, 103:218–228, 1962.

[Kni63] F. B. Knight. Random walks and a sojourn density process of Brownian motion. *Trans. Amer. Math. Soc.*, 109:56–86, 1963.

[Kni97] Frank B. Knight. Approximation of stopped Brownian local time by diadic crossing chains. *Stochastic Process. Appl.*, 66(2):253–270, 1997.

[KS91] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[Kud82] R. Kudzhma. Itô's formula for a random walk. *Litovsk. Mat. Sb.*, 22(3):122–127, 1982.

[Leu98] Christophe Leuridan. Le théorème de Ray-Knight à temps fixe. In *Séminaire de Probabilités, XXXII*, volume 1686 of *Lecture Notes in Math.*, pages 376–396. Springer, Berlin, 1998.

[LG86] J.-F. Le Gall. Une approche élémentaire des théorèmes de décomposition de Williams. In *Séminaire de Probabilités, XX, 1984/85*, volume 1204 of *Lecture Notes in Math.*, pages 447–464. Springer, Berlin, 1986.

[LP10] Lionel Levine and James Propp. What is . . . a sandpile? *Notices Amer. Math. Soc.*, 57(8):976–979, 2010.

[LP13] R. Lyons and Y. Peres. *Probability on Trees and Networks*. Cambridge Univ. Press, 2013. In preparation. Current version available at `http://mypage.iu.edu/~rdlyons/`.

[LPT13] Titus Lupu, Jim Pitman, and Wenpin Tang. The Vervaat transform of Brownian bridges and Brownian motion. *Pre-print*, 2013. arXiv:1310.3889 [math.PR].

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

[Per82] Edwin Perkins. Local time is a semimartingale. *Z. Wahrsch. Verw. Gebiete*, 60(1):79–117, 1982.

[Pit98] Jim Pitman. Enumerations of trees and forests related to branching processes and random walks. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, volume 41 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 163–180. Amer. Math. Soc., Providence, RI, 1998.

[Pit99] Jim Pitman. The SDE solved by local times of a Brownian excursion or bridge derived from the height profile of a random tree or forest. *Ann. Probab.*, 27(1):261–283, 1999.

[Por63] Sidney C. Port. An elementary probability approach to fluctuation theory. *J. Math. Anal. Appl.*, 6:109–151, 1963.

[PPY92] Mihael Perman, Jim Pitman, and Marc Yor. Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Related Fields*, 92(1):21–39, 1992.

[PW98] James Gary Propp and David Bruce Wilson. How to get a perfectly random sample from a generic Markov chain and generate a random spanning tree of a directed graph. *J. Algorithms*, 27(2):170–217, 1998. 7th Annual ACM-SIAM Symposium on Discrete Algorithms (Atlanta, GA, 1996).

[Rév81] P. Révész. Local time and invariance. In *Analytical methods in probability theory (Oberwolfach, 1980)*, volume 861 of *Lecture Notes in Math.*, pages 128–145. Springer, Berlin, 1981.

[Rév90] Pál Révész. *Random walk in random and nonrandom environments*. World Scientific Publishing Co. Inc., Teaneck, NJ, 1990.

[RW94] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 1*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Ltd., Chichester, second edition, 1994. Foundations.

[Spi56] Frank Spitzer. A combinatorial lemma and its application to probability theory. *Trans. Amer. Math. Soc.*, 82:323–339, 1956.

[SS05] Tamás Szabados and Balázs Székely. An elementary approach to Brownian local time based on simple, symmetric random walks. *Period. Math. Hungar.*, 51(1):79–98, 2005.

[SS09] Tamás Szabados and Balázs Székely. Stochastic integration based on simple, symmetric random walks. *J. Theoret. Probab.*, 22(1):203–219, 2009.

[Sta] Richard P. Stanley. http://www-math.mit.edu/ rstan/ec/. Accessed 22 September 2013.

[Sta99] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[Sza90] Tamás Szabados. A discrete Itô's formula. In *Limit theorems in probability and statistics (Pécs, 1989)*, volume 57 of *Colloq. Math. Soc. János Bolyai*, pages 491–502. North-Holland, Amsterdam, 1990.

[Tan13] Wenpin Tang. Personal communication, 2013.

[Tro58] H. F. Trotter. A property of Brownian motion paths. *Illinois J. Math.*, 2:425–433, 1958.

[vAEdB51] T. van Aardenne-Ehrenfest and N. G. de Bruijn. Circuits and trees in oriented linear graphs. *Simon Stevin*, 28:203–217, 1951.

[Ver79] Wim Vervaat. A relation between Brownian bridge and Brownian excursion. *Ann. Probab.*, 7(1):143–149, 1979.

[Wen60] J. G. Wendel. Order statistics of partial sums. *Ann. Math. Statist.*, 31:1034–1044, 1960.