Random walks

Thomas Budzinski

December 1, 2025

Introduction

For $d \geq 1$, let $(e_i)_{1 \leq i \leq d}$ denote the canonical basis of $\mathbb{R}^d$. The *simple random walk* on $\mathbb{Z}^d$ is the random process $(S_n)_{n \geq 0}$ defined by

$$S_n = X_1 + \cdots + X_n,$$

where the variables X_i are independent and identically distributed with uniform law on

$$\{e_1, -e_1, \ldots, e_d, -e_d\}$$

. This process can also be described as a Markov chain on $\mathbb{Z}^d$ with a very simple transition matrix. In particular, it is irreducible and is recurrent for $d \in \{1, 2\}$ and transient for $d \geq 3$.

The are (at least!) two natural ways to widely generalize this process:

- we can replace the uniform measure on $\{e_1, -e_1, \ldots, e_d, -e_d\}$ by more general distributions on $\mathbb{Z}^d$ or even $\mathbb{R}^d$,
- or, if we view $\mathbb{Z}^d$ as a graph and (S_n) as a process choosing S_{n+1} uniformly at random among the neighbours of S_n.

In probability, *"random walks"* is a generic expression that encompasses both of these generalizations. In both cases, we can ask many natural questions such as recurrence or transience, typical displacement ("how far is S_n from S_0?"), or fine study of the set of points visited by the walk.

The first part of the course will be dedicated to the first generalization, i.e. random walks on $\mathbb{R}^d$. In the second part, we will study random walks on general graphs, which is the second generalization.

Contents

1 Random walks on $\mathbb{R}^d$ 3

1.1 Recurrence and transience . . . 3
1.1.1 Definitions and first properties . . . 3
1.1.2 Recurrence and transience . . . 5
1.1.3 The one-dimensional case . . . 8
1.2 Characteristic function and applications . . . 11
1.2.1 Motivation . . . 11
1.2.2 The Chung–Fuchs criterion . . . 12
1.2.3 Applications . . . 14
1.2.4 The local central limit theorem . . . 16
1.3 Ladder times and cyclic lemma . . . 19
1.3.1 Ladder times and heights . . . 19
1.3.2 Duality . . . 20
1.3.3 The cyclic lemma . . . 21
1.3.4 Kemperman's formula and applications . . . 23
1.4 Applications to random trees . . . 26
1.4.1 Definitions . . . 26
1.4.2 The contour function . . . 27
1.4.3 The Lukasiewicz path . . . 32
1.4.4 Galton–Watson trees . . . 34
1.5 Range and intersections of random walks . . . 36
1.5.1 Linear growth of the range . . . 36
1.5.2 The range in dimension 2 . . . 38
1.5.3 Intersections of random walks . . . 41

2 Random walks on general graphs 45

2.1 Electric networks . . . 45
2.1.1 Networks, voltages and flows . . . 45
2.1.2 Effective resistances . . . 49
2.1.3 Energy and variational principles . . . 51
2.1.4 Resistances in infinite networks . . . 54
2.1.5 Recurrence and transience criteria and applications . . . 57
2.1.6 Invariance under quasi-isometries . . . 61
2.2 Spectral radius and isoperimetric inequalities . . . 64

Chapter 1

Random walks on $\mathbb{R}^d$

1.1 Recurrence and transience

1.1.1 Definitions and first properties

Let us start by giving the main definition for the first part of the course.

Definition 1. Let $d \geq 1$, and let μ be a probability measure on $\mathbb{R}^d$. A *random walk* on $\mathbb{R}^d$ with step distribution μ is a process $(S_n)_{n\geq 0}$ of the form

$$S_n = X_1 + \cdots + X_n,$$

where the X_i are independent and identically distributed with law μ.

Our goal will be tu understand the behaviour of these processes, and in particular how it depends on the dimension d and the step distribution μ.

Remark 2. It is possible (for example if μ has a density with respect to the Lebesgue measure) that the state space of S is uncountable. Therefore, the general theory of Markov chain does not apply to our setting. Therefore, the main goal of this Section 1.1 will be to reprove some of its results (strong Markov property, recurrence/transience dichotomy) with adapted definitions.

We denote by $\mathcal{F}_n$ the σ-algebra generated by $(X_1, \ldots, X_n)$, so that $(\mathcal{F}_n)_{n\geq 0}$ is the natural filtration associated to the process (S_n).

Definition 3.

- A *stopping time* is a random variable $\tau : \Omega \to \mathbb{N} \cup \{+\infty\}$ such that for all $n \geq 0$, we have $\{\tau \leq n\} \in \mathcal{F}_n$.
- For any stopping time τ, we denote by $\mathcal{F}_\tau$ the σ-algebra consisting of all the events A such that for all $n \geq 0$, we have $A \cap \{\tau \leq n\} \in \mathcal{F}_n$.

We can now formulate the strong Markov property.

Proposition 4 (Strong Markov property)**.** Let τ be a stopping time such that $\tau < +\infty$ almost surely, and let $S_n^{(\tau)} = S_{\tau+n} - S_\tau$. Then $S^{(\tau)}$ has the same law as S and is independent from $\mathcal{F}_\tau$.

This Markov property is not an immediate conseqquence of the general one of Markov chains because the state space may be uncountable. However, the proof is exactly the same.

Proof. Let $A \in \mathcal{F}_\tau$ and $k \geq 1$. Let also $f : \mathbb{R}^k \to \mathbb{R}$ be measurable and bounded. We can write

$$\begin{aligned}\mathbb{E}\left[\mathbb{1}_A f\left(S_1^{(\tau)}, \dots, S_k^{(\tau)}\right)\right] &= \sum_{n\geq 0} \mathbb{E}\left[\mathbb{1}_{A\cap\{\tau=n\}} f\left(X_{n+1}, \dots, X_{n+1} + \cdots + X_{n+k}\right)\right] \\ &= \sum_{n\geq 0} \mathbb{P}\left(A \cap \{\tau = n\}\right) \mathbb{E}\left[f\left(X_{n+1}, \dots, X_{n+1} + \cdots + X_{n+k}\right)\right] \\ &= \mathbb{P}(A)\mathbb{E}\left[f\left(S_1, \dots, S_k\right)\right].\end{aligned}$$

It follows that $\left(S_1^{(\tau)}, \dots, S_k^{(\tau)}\right)$ is independent from A for all k, so $S^{(\tau)}$ is independent from $\mathcal{F}_\tau$. Finally, by taking $A = \Omega$, we find that $S^{(\tau)}$ has the same law as S. □

The Hewitt–Savage 0-1 law. The usual Kolmogorv 0-1 law says that any event invariant under changing the values of finitely many increments has probability 0 or 1. Unfortunately, the assumption is too strong to apply to many natural events, such as "hitting an interval infinitely many times". Therefore, the following stronger statement will be very useful to us.

Theorem 1. Let $(X_i)_{i\geq 1}$ be a sequence of i.i.d. random variables with values in a set E. Let $A \in \sigma\left(X_i | i \geq 1\right)$ be an event which is invariant under any permutation of the X_i with finite support (that is, any bijection $\sigma : \mathbb{N}^* \to \mathbb{N}^*$ such that $\sigma(i) = i$ except for finitely many i). Then $\mathbb{P}(A) \in \{0, 1\}$.

Proof. The result is easy to prove if we further assume that A only depends on finitely many coordinates. Therefore, we will approximate A by events satisfying this property.

More precisely, let $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. We first claim that there are events $A_n \in \mathcal{F}_n$ such that $\mathbb{P}(A \Delta A_n) \to 0$ as $n \to +\infty$, where the symetric difference $A\Delta A_n$ stands for $(A \backslash A_n) \cup (A_n \backslash A)$. Indeed, the result is obvious if A itself belongs to $\bigcup_n \mathcal{F}_n$. Moreover, it is straightforward to check that the set of events satisfying this approximation property form a σ-algebra[1], so it must be $\mathcal{F}_\infty$.

Now, let us fix $n \geq 1$. We know that any event $B \in \mathcal{F}_\infty$ is of the form

$$\left\{(X_i)_{i\geq 1} \in \widetilde{B}\right\}$$

for some $\widetilde{B} \subset E^{\mathbb{N}^*}$ measurable. We then define the event

$$\phi_n(B) = \left\{(X_{n+1}, \dots, X_{2n}, X_1, \dots, X_n, X_{2n+1}, X_{2n+2}, \dots) \in \widetilde{B}\right\}.$$

Now, if A_n is the approximation of A described above, we can write

$$\mathbb{P}\left(A \Delta \Phi_n(A_n)\right) = \mathbb{P}\left(\phi_n(A) \Delta \phi_n(A_n)\right) = \mathbb{P}\left(\phi(A \Delta A_n)\right) = \mathbb{P}\left(A \Delta A_n\right) \xrightarrow[n\to+\infty]{} 0,$$

where the first equality uses the assumption on A and the third uses that the permuted process

$$(X_{n+1}, \dots, X_{2n}, X_1, \dots, X_n, X_{2n+1}, X_{2n+2}, \dots)$$

has the same law as $(X_i)_{i\geq 1}$. The last equation means that both A_n and $\phi_n(A_n)$ are good approximations of A, so their intersection $A_n \cap \phi_n(A_n)$ also is. More precisely, we have

$$\mathbb{P}\left(A \Delta (A_n \cap \phi_n(A_n))\right) \leq \mathbb{P}\left(A \Delta A_n\right) + \mathbb{P}\left(A \Delta \phi_n(A_n)\right) \xrightarrow[n\to+\infty]{} 0,$$

[1] Contrary to what I said in class, the monotone class theorem is not needed here.

so

$$\mathbb{P}(A) = \lim_{n\to+\infty} \mathbb{P}\left(A_n \cap \phi_n(A_n)\right) = \lim_{n\to+\infty} \mathbb{P}\left(A_n\right)\mathbb{P}\left(\phi_n(A_n)\right) = \mathbb{P}(A)^2,$$

so $\mathbb{P}(A) \in \{0, 1\}$. □

Example 5. If $E \subset \mathbb{R}^d$, the event

$$A = \{\text{there are infinitely many times } n \text{ such that } S_n \in E\}$$

satisfies the assumption of the Hewitt–Savage 0-1 law, but not the assumption of the Kolmogorov version. Most of the upcoming applications will be events of this form.

1.1.2 Recurrence and transience

If for example our step distribution μ is included in the lattice $\mathbb{Z}^d$, then we can see (S_n) as a Markov chain on $\mathbb{Z}^d$, which is either recurrent or transient. On the other hand, if for example μ has a density with respect to the Lebesgue measure, then almost surely, we have $S_n \neq 0$ for all n, which would mean that S would be transient regardless of its "large scale" properties if we blindly copy the "discrete" definition. Therefore, it seems to make more sense to replace "visit a point infinitely often" by "approach a point infinitely often" in the definition of recurrence.

More precisely, for $x = (x_i)_{1\leq i\leq d} \in \mathbb{R}^d$, we write $|x| = \max_{1\leq i\leq d} |x_i|$.

Definition 6. Let $x \in \mathbb{R}^d$.

- We say that x is a *possible value* of S if

$$\forall \varepsilon > 0,\ \exists n \geq 0,\ \mathbb{P}\left(|S_n - x| < \varepsilon\right) > 0.$$

- We say that x is a *recurrent value* of S if

$$\forall \varepsilon > 0,\ \mathbb{P}\left(|S_n - x| < \varepsilon \text{ for infinitely many } n\right) > 0.$$

We write $\mathcal{P}_\mu$ for the set of possible values, and $\mathcal{R}_\mu$ for the set of recurrent values.

Remark 7. By the Hewitt–Savage law, the probability appearing in the definition of a recurrent value is either 0 or 1, so x is recurrent if and only if this probability is 1 for any $\varepsilon > 0$.

Example 8. For the simple random walk on $\mathbb{Z}^d$, we would have $\mathcal{P}_\mu = \mathbb{Z}^d$ and $\mathcal{R}_\mu$ is $\mathbb{Z}^d$ for $d \in \{1, 2\}$ and is empty for $d \geq 3$.

The goal of the next result is to guarantee the equivalence between natural definitions of a recurrent/transient random walk on $\mathbb{R}^d$.

Theorem 2. Let μ be a probability measure on $\mathbb{R}^d$. Either $\mathcal{R}_\mu$ is empty, or $\mathcal{R}_\mu = \mathcal{P}_\mu$ is a closed subgroup of $\mathbb{R}^d$. We say that S is *transient* in the first case, and *recurrent* in the second.

Proof. • We first notice that if $x \notin \mathcal{R}_\mu$, there is $\varepsilon > 0$ such that almost surely, the ball of radius ε around x is visited finitely many times. It follows that if $|x - y| < \frac{\varepsilon}{2}$, then the ball of radius $\frac{\varepsilon}{2}$ around y is visited finitely many times, so the complement of $\mathcal{R}_\mu$ is open, so $\mathcal{R}_\mu$ is closed.

- We now assume $\mathcal{R}_\mu \neq \emptyset$. We will show

$$\forall x \in \mathcal{P}_\mu,\ \forall y \in \mathcal{R}_\mu,\ y - x \in \mathcal{R}_\mu. \tag{1.1}$$

  The rough idea behind this equation is that S must approach y even if we condition it to hit a small ball around x, and to notice that by the Markov property, this means that a shifted version of S approaches $y - x$. More precisely, for any $z \in \mathbb{R}^d$ and any $m \geq 1$ and $\delta > 0$, we write

$$p_{\delta,m}(z) = \mathbb{P}\left(\forall n \geq m,\ |S_m - z| \geq \delta\right).$$

  We now fix $m \geq 1$ and $\varepsilon > 0$. Since $x \in \mathcal{P}_\mu$, let $k \geq 0$ be such that $\mathbb{P}\left(|S_k - x| < \varepsilon\right) > 0$. Then we have

$$\begin{aligned} p_{\varepsilon,m+k}(y) &\geq \mathbb{P}\left(|S_k - x| < \varepsilon \text{ and } \forall n \geq k+m,\ |S_n - S_k - (y-x)| \geq 2\varepsilon\right) \\ &= \mathbb{P}\left(|S_k - x| < \varepsilon\right) \times p_{2\varepsilon,m}(y - x) \end{aligned}$$

  by the fact that $(S_{k+i} - S_k)$ is a random walk and is independent of $\mathcal{F}_k$. Since $y \in \mathcal{R}_\mu$, the left-hand side is 0, so the right-hand side as well. But we know that the first factor is positive, so $p_{2\varepsilon,m}(y - x) = 0$ for any $m \geq 1$ and $\varepsilon > 0$, so $y - x \in \mathcal{R}_\mu$, which proves (1.1).

- We now show that $\mathcal{R}_\mu$ is a group: let $x, x' \in \mathcal{R}_\mu$. We first apply (1.1) with $y = x$ to get $0 \in \mathcal{R}_\mu$. We then apply (1.1) a second time with $y = 0$ to obtain $-x, -x' \in \mathcal{R}_\mu$. Finally, a last application of (1.1) with $y = -x'$ gives $-x - x' \in \mathcal{R}_\mu$, so $x + x' \in \mathcal{R}_\mu$.

- Finally, let $x \in \mathcal{P}_\mu$. By (1.1) with $y = 0$, we get $-x \in \mathcal{R}_\mu$, so $x \in \mathcal{R}_\mu$, so $\mathcal{R}_\mu = \mathcal{P}_\mu$ as soon as $\mathcal{R}_\mu \neq \emptyset$.

□

Remark 9. In dimensions 1, a closed subgroup of $\mathbb{R}$ is either $\mathbb{R}$ or $c\mathbb{Z}$ for some $c \geq 0$. In higher dimensions, a closed subgroup is of the form $\mathbb{Z}^i \times R^j \times \{0\}^k$ with $i + j + k = d$, up to the action of $GL_d(\mathbb{R})$.

A first characterization of recurrence/transience. For discrete Markov chain, we know that a point x is recurrent if and only if

$$\sum_{n \geq 0} \mathbb{P}_x\left(S_n = x\right) = +\infty.$$

Indeed, the left-hand side represents the expectation of the number of visits to 0, which is a geometric random variable by the strong Markov propety. Therefore, it is a.s. finite if and only if it has finite expectation. We now prove an analog of this result for random walks on $\mathbb{R}^d$.

Theorem 3. Fix $\varepsilon > 0$. A random walk S on $\mathbb{R}^d$ is transient if and only if

$$\sum_{n \geq 0} \mathbb{P}\left(|S_n| < \varepsilon\right) < +\infty.$$

Proof. We will first prove that the finiteness of the sum in the statement does not depend on the choice of $\varepsilon > 0$ (this is an immediate consequence of our Theorem, but we will need to prove this first). For this, we will prove the following inequality for any $\varepsilon > 0$ and integer $m \geq 1$:

$$\sum_{n \geq 0} \mathbb{P}\left(|S_n| < m\varepsilon\right) \leq (2m)^d \sum_{n \geq 0} \mathbb{P}\left(|S_n| < \varepsilon\right). \tag{1.2}$$

To show this inequality, we will partition the "big cube" $[-m\varepsilon, m\varepsilon)^d$ into $(2m)^d$ little cubes with side length ε (this is where the choice of the norm $|\cdot|$ is important). The little cubes will be indexed by $\mathbf{k} \in \{-m, \dots, m-1\}^d$. We get

$$\sum_{n\geq 0} \mathbb{P}\left(|S_n| < m\varepsilon\right) \leq \sum_{n\geq 0} \sum_{\mathbf{k}\in\{-m,\dots,m-1\}^d} \mathbb{P}\left(S_n \in \varepsilon\mathbf{k} + [0,\varepsilon)^d\right),$$

where the inequality comes from replacing the open ball $(-m\varepsilon, m\varepsilon)^d$ by $[-m\varepsilon, m\varepsilon)$. In order to study the contribution of a fixed $\mathbf{k} \in \{-m, \dots, m-1\}^d$, we write

$$T_{\mathbf{k}} = \inf\{\ell \geq 0 | S_\ell \in \varepsilon\mathbf{k} + [0,\varepsilon)^d\}.$$

We can now write

$$\begin{aligned}
\sum_{n\geq 0} \mathbb{P}\left(S_n \in \varepsilon\mathbf{k} + [0,\varepsilon)^d\right) &= \sum_{n\geq 0}\sum_{\ell=0}^{n} \mathbb{P}\left(S_n \in \varepsilon\mathbf{k} + [0,\varepsilon)^d \text{ and } T_{\mathbf{k}} = \ell\right) \\
&\leq \sum_{\ell\geq 0}\sum_{n\geq \ell} \mathbb{P}\left(T_{\mathbf{k}} = \ell \text{ and } |S_n - S_\ell| < \varepsilon\right) \\
&= \sum_{\ell\geq 0}\sum_{j\geq 0} \mathbb{P}\left(T_{\mathbf{k}} = \ell\right)\mathbb{P}\left(|S_j| < \varepsilon\right) \qquad\qquad \leq \sum_{j\geq 0} \mathbb{P}\left(|S_j| < \varepsilon\right),
\end{aligned}$$

where the second line uses that two points in the same little cube must be close, and the third line uses independence. By summing over $\mathbf{k}$, we obtain (1.2). In particular, if the sum is finite for some $\varepsilon > 0$, then it is finite for all $\varepsilon > 0$.

We now finish the proof of the theorem. First, if $\sum_{n\geq 0} \mathbb{P}\left(|S_n| < \varepsilon\right) < +\infty$, then the expected number of vists of S in a certain ball is finite, so the number of visits is a.s. finite, which means that S is transient. On the other hand, if S is transient, we can find $\varepsilon > 0$ small such that

$$p_\varepsilon := \mathbb{P}\left(\forall n \geq 1,\, |S_n| > \varepsilon\right) > 0.$$

As in the usual discrete setting, we will try to compare the number of visits of a certain small ball with a geometric variable with parameter $1 - p_\varepsilon$. For this, we define the stopping times τ_i by $\tau_0 = 0$ and, for $i \geq 0$,

$$\tau_{i+1} = \inf\{n > \tau_i \,\big|\, |S_n| < \frac{\varepsilon}{2}\}.$$

For any i, we have

$$\mathbb{P}\left(\tau_{i+1} = +\infty | \tau_i < +\infty\right) \geq \mathbb{P}\left(\forall n \geq 1,\, \left|S_n^{(\tau_i)}\right| > \varepsilon\right) = p_\varepsilon,$$

where the inequality comes from the triangle inequality and the last equality from the strong Markov property. By induction on i, this becomes

$$\mathbb{P}\left(\tau_i < +\infty\right) \leq (1 - p_\varepsilon)^i$$

for all $i \geq 0$, so

$$\sum_{n\geq 0} \mathbb{P}\left(|S_n| < \varepsilon\right) = \mathbb{E}\left[\sum_{i\geq 0} \mathbb{1}_{\tau_i < +\infty}\right] < +\infty.$$

This proves the finiteness of the sum for some $\varepsilon > 0$, and the first part of the proof extends this to all $\varepsilon > 0$. □

1.1.3 The one-dimensional case

We conclude this section with a more precised study of the one-dimensional case. Note that the projection of a random walk on $\mathbb{R}^d$ on any of its coordinates is a one-dimensional random walk, so the results below also say something about higher-dimensional walks. We now assume $d = 1$ and $\mu \neq \delta_0$.

Proposition 10. One of the following holds almost surely:

1. $S_n \xrightarrow[n\to+\infty]{} +\infty$,
2. $S_n \xrightarrow[n\to+\infty]{} -\infty$,
3. $\liminf_{n\to+\infty} S_n = -\infty$ and $\limsup_{n\to+\infty} S_n = +\infty$.

We say that S drifts to $+\infty$ in case 1, that S drifts to $+\infty$ in case 2 and that S oscillates in case 3.

Proof. By the Kolmogorov 0-1 law, each of the three cases above has probability 0 or 1, so it is enough to show that at least one occurs. If this is not the case, note that it means that either $\liminf S$ or $\limsup S$ is finite. Note that $\liminf S$ and $\limsup S$ are both deterministic by the Hewitt–Savage 0-1 law. Without loss of generality, assume that $\liminf S = c \in \mathbb{R}$ almost surely. Then $\mathcal{R}_\mu$ contains c, so it is nonempty, so $\mathcal{R}_\mu = \mathcal{P}_\mu$ is a subgroup of $\mathbb{R}$. But on the other hand $\mathcal{R}_\mu$ is bounded from below by c, so $\mathcal{R}_\mu = \{0\}$, so $\mathcal{P}_\mu = 0$. This implies $\mu = \delta_0$, which was excluded. □

Remark 11. A recurrent walk must oscillate, but a transient walk can either drift to $+\infty$ or $-\infty$, or oscillate (see examples below).

One-dimensional random walks with finite expectation. In the case of walks witgh finite expectation, the law of large numbers will allow us to get a very simple criterion for recurrence.

Theorem 4. Assume that $\mathbb{E}[|X_1|] < +\infty$.

- If $\mathbb{E}[X_1] > 0$, then S is transient and drifts to $+\infty$.
- If $\mathbb{E}[X_1] < 0$, then S is transient and drifts to $-\infty$.
- If $\mathbb{E}[X_1] = 0$, then S is recurrent and oscillates.

Proof. The transient case follows directly from the law of large numbers: we have $\frac{S_n}{n} \to \mathbb{E}[X_1] > 0$ almost surely, so $S_n > 1$ for n large enough, and similarly for the negative case.

For the recurrent case, a bit more work is needed. Roughly speaking, the argument will be that a walk with mean 0 at time n must be in $[-\varepsilon n, \varepsilon n]$ by the law of large numbers. On the other hand, if it is transient, it cannot spend too much time in a given interval of length 1. We will get a contradiction by cutting $[-\varepsilon n, \varepsilon n]$ into small intervals of length 1.

More precisely, let $\varepsilon > 0$. By the law of large number, we know that for n large enough (say, $n \geq N_0$ where N_0 is random), we have $|S_n| \leq \varepsilon n$. It follows that, for n large enough to have

$$n \geq \max\left(N_0, |S_1|/\varepsilon, \ldots, |S_{N_0}|/\varepsilon\right),$$

we have

$$\forall 0 \leq i \leq n,\ |S_i| \leq \varepsilon n.$$

Therefore, with probability going to 1 as $n \to +\infty$, we have

$$\sum_{i=0}^{+\infty} \mathbb{1}_{|S_i| \leq \varepsilon n} \geq n,$$

so for n large enough, we can write

$$\mathbb{E}\left[\sum_{i=0}^{+\infty} \mathbb{1}_{|S_i| \leq \varepsilon n}\right] \geq \frac{n}{2}. \tag{1.3}$$

On the other hand, assume that S is transient. Then we can write $\sum_{i \geq 0} \mathbb{P}(|S_i| < 1) \leq C$ for some constant C (which, crucially, does not depend on ε). Therefore, for all $m \in \mathbb{Z}$, let $\tau_m = \inf\{i \geq 0 | S_i \in [m, m+1]\}$. By applying the strong Markov property to τ_m, we can write

$$\sum_{i \geq 0} \mathbb{P}(S_i \in [m, m+1]) = \mathbb{E}\left[\sum_{i \geq 0} \mathbb{1}_{S_i \in [m, m+1]}\right] \leq \mathbb{P}(\tau_m < +\infty)\, \mathbb{E}\left[\sum_{i \geq 0} \mathbb{1}_{|S_i| < 1}\right] \leq C.$$

Finally, by summing over $-\varepsilon n - 1 \leq m \leq \varepsilon n$, we find

$$\sum_{i \geq 0} \mathbb{P}(|S_i| \leq \varepsilon n) \leq (2\varepsilon n + 2)C.$$

This contradicts (1.3) if $\varepsilon > 0$ was chosen small enough. □

Heavy-tailed walks. We now want to say a few things about the case where $\mathbb{E}[|X_1|] = +\infty$, which is actually quite rich. We will first focus on the case where X is symetric and has a polynomial tail.

Proposition 12. Let μ be a probability measure on $\mathbb{Z}$ such that for any $k \geq 0$, we have $\mu(k) = \mu(-k) \sim k^{-\alpha}$ as $k \to +\infty$, where $\alpha \in (1, 2)$. Then the random walk S on $\mathbb{R}$ with step distribution μ is transient and oscillates.

Remark 13. We must have $\alpha > 1$ for the measure to be finite, and $\alpha < 2$ implies $\mathbb{E}[|X_1|] = +\infty$ (the case $\alpha = 2$ is a "limit" case).

Proof. First, note that S has the same law as $-S$ so S cannot drift to $+\infty$ or $-\infty$, so it oscillates. To prove transience, we will bound the probability $\mathbb{P}(S_n = 0)$ by using the fact that S makes at least one "large" jump before time n. More precisely, we fix $\varepsilon > 0$ and define

$$\tau = \inf\{i \geq 1 \big| |X_i| > n^{1+\varepsilon}\}.$$

By conditioning on τ, we can write

$$\mathbb{P}(S_n = 0) \leq \mathbb{P}(\tau > n) + \mathbb{E}\left[\mathbb{P}(S_n = 0 | \tau)\, \mathbb{1}_{\tau \leq n}\right]. \tag{1.4}$$

The first term is easy to bound: there is a constant $c > 0$ such that we have

$$\mathbb{P}(\tau > n) = \left(1 - \mathbb{P}\left(|X_1| \leq n^{1+\varepsilon}\right)\right)^n \leq \exp\left(-n \times (c + o(1))(n^{1+\varepsilon})^{1-\alpha}\right) \leq \exp(-n^{\varepsilon})$$

if ε was chosen small enough. This is summable. On the other hand, let us study the conditional distribution of $(X_i)_{i\leq 1}$ conditionally on τ. We find that the steps (X_i) are still independent and X_i has conditional law μ for $i > \tau$ and μ conditionned on being smaller than $n^{1+\varepsilon}$ for $i < \tau$. Finally, the conditional law of X_τ given τ is given by

$$\widetilde{\mu}(i) = \mathbb{1}_{|i|>n^{1+\varepsilon}} \frac{\mu(i)}{\sum_{|j|>n^{1+\varepsilon}} \mu(j)} \leq \frac{C(n^{1+\varepsilon})^{-\alpha}}{c\,(n^{1+\varepsilon})^{1-\alpha}} = O\left(n^{-1-\varepsilon}\right)$$

for some constants $c, C > 0$. Therefore, if $\tau \leq n$, we can write

$$\mathbb{P}\left(S_n = 0|\tau\right) = \mathbb{P}\left(X_\tau = -(X_1 + \cdots + X_{\tau-1} + X_{\tau+1} + \cdots + X_n)|\tau \leq n\right) \leq \sup_i \widetilde{\mu}(i) = O\left(n^{-1-\varepsilon}\right).$$

By taking the expectation, we find that the second term of (1.4) is summable, so S is transient. □

Remark 14. It is quite natural to ask what is the "slowest possible" rate of decrease for which the walk is still recurrent. It turns out that this question is quite difficult, and that some very irregular choices of μ can give rise to surprising counter-examples. For example, it was proved by Shepp [5] that for any function $\varepsilon(x) \to 0$ as $x \to +\infty$, we can find a symetric measure μ on $\mathbb{Z}$ such that $\mu\left(\mathbb{R}\backslash[-x, x]\right) \geq \varepsilon(x)$ for x large enough, but the associated random walk is still recurrent.

The Cauchy random walk. Finally, for symetric, heavy-tailed random walk with a "regular behaviour" like $\mu_k \sim k^{-\alpha}$, we have seen that the walk is recurrent for $\alpha > 2$ and transient for $\alpha < 2$. We conclude with a specific example which belongs to the "limiting" case $\alpha = 2$: the *Cauchy* random walk.

We denote by (B_t^1, B_t^2) a two-dimensional Brownian motion started from 0. We define $\tau_1 = \inf\{t \geq 0|B_t^1 = 1\}$ and $X = B_{\tau_1}^2$. In other words, the variable X describes the point at which B first hits the vertical line $\{x = 1\}$. Then we can compute that X has density $\frac{1}{\pi(1+x^2)}$ on $\mathbb{R}$ with respect to the Lebesgue measure (this is the *Cauchy* distribution), which in particular implies $\mathbb{E}[|X|] = +\infty$.

Moreover, let $(X_i)_{i\geq 1}$ be i.i.d. copies of X. By applying the strong Markov property of Brownian motion to the hitting times τ_n of the vertical lines $\{x = n\}$, we find that $S_n = X_1 + \cdots + X_n$ follows the same law as $B_{\tau_n}^2$. By invariance of Brownian motion under scaling, this has the same law as nX_1, which shows

$$\mathbb{P}\left(|S_n| < 1\right) = \sum_{n\geq 1} \mathbb{P}\left(|X_1| < \frac{1}{n}\right) \underset{n\to+\infty}{\sim} \frac{1}{\pi n}.$$

This is summable, so the random walk S is recurrent. Moreover, for any fixed constant c, the same argument applies if we replace X by $X - c$. Therefore, we have built a one-dimensional random walk S such that for any $c \in \mathbb{R}$, the walk $(S_n - cn)_{n\geq 0}$ is recurrent. This means that for any c, there are infinitely many n such that S_n is "close" to cn, which is quite counterintuitive when compared to the finite expectation case!

Stable random variables. Cauchy variables are a particular case of the following class of random variables.

Definition 15. Let $\alpha \in (0,2]$. An α-stable law μ on $\mathbb{R}$ is a law such that if $X_1, \dots, X_n$ are i.i.d. with law μ, then $X_1 + \cdots + X_n$ has the same law as $n^{1/\alpha} X_1$.

Example 16. The most famous example is the Gaussian distribution ($\alpha = 2$). The Cauchy distribution is also stable with $\alpha = 1$, and so is the law of the hitting time of 1 for a Brownian motion (with $\alpha = 1/2$). These are the only examples of stable distributions with explicit densities with respect to Lebesgue. Stable distributions replace the Gaussian one in the central limit theorem when the variables have a heavy polynomial tail.

1.2 Characteristic function and applications

1.2.1 Motivation

Let μ be a probability measure on $\mathbb{R}^d$. Let $(X_i)_{i\geq 1}$ be i.i.d. variables with distribution μ and let $(S_n)_{n\geq 0} = (X_1 + \cdots + X_n)_{n\geq 0}$ be the corresponding random walk. For $\xi \in \mathbb{R}^d$, we define

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^d} e^{i\xi\cdot x} \mu(\mathrm{d}x) = \mathbb{E}\left[e^{i\xi\cdot X}\right], \tag{1.5}$$

where X has law μ. The function $\hat{\mu}$ is called the *characteristic function* of μ. In particular, we have $\hat{\mu}(0) = 1$ and $|\hat{\mu}(\xi)| \leq 1$ for all $\xi \in \mathbb{R}^d$. Moreover, if $\mathbb{E}[|X|] < +\infty$, then differentiating (1.5) gives

$$\hat{\mu}(\xi) = 1 + i\xi \cdot \mathbb{E}[X] + o(|\xi|) \quad \text{as } \xi \to 0.$$

Similarly, if $\mathbb{E}[|X|^2] < +\infty$, by differentiating (1.5) twice, we find

$$\hat{\mu}(\xi) = 1 + i\xi \cdot \mathbb{E}[X] - \frac{\xi^t Q \xi}{2} + o\left(|\xi|^2\right) \quad \text{as } \xi \to 0,$$

where Q is the covariance matrix of X, i.e. $Q_{i,j} = \mathbb{E}\left[X_1^{(i)} X_1^{(j)}\right]$.

The reason why this characteristic function will play an important role for us is that we can easily access the characteristic function of S_n: for all $n \geq 0$ and $\xi \in \mathbb{R}^d$, we have

$$\mathbb{E}\left[e^{i\xi\cdot S_n}\right] = \mathbb{E}\left[e^{i\xi\cdot X_1} \dots e^{i\xi\cdot x_n}\right] = \hat{\mu}(\xi)^n.$$

Since the characteristic function characterizes the distribution, this is in theory sufficient to recover any information about the law of S_n. As an example, let us give an exact formula for random walks on $\mathbb{Z}^d$.

Lemma 17. If $\mu(\mathbb{Z}^d) = 1$, then for all $n \geq 0$ and $x \in \mathbb{Z}^d$, we have

$$\mathbb{P}(S_n = x) = \frac{1}{(2\pi)^d} \int_{(-\pi,\pi)^d} e^{-i\xi\cdot x} \hat{\mu}(\xi)^n \mathrm{d}\xi.$$

Proof. The result is an example of application of the Fourier inversion formula. More precisely, let us compute the right-hand side using the Fubini theorem:

$$\begin{aligned}\frac{1}{(2\pi)^d} \int_{(-\pi,\pi)^d} e^{-i\xi\cdot x} \mathbb{E}\left[e^{i\xi\cdot S_n}\right] \mathrm{d}\xi &= \frac{1}{(2\pi)^d} \int_{(-\pi,\pi)^d} \sum_{y\in\mathbb{Z}^d} e^{-i\xi\cdot x} e^{i\xi\cdot y} \mathbb{P}(S_n = y) \mathrm{d}\xi \\ &= \frac{1}{(2\pi)^d} \sum_{y\in\mathbb{Z}^d} \mathbb{P}(S_n = y) \int_{(-\pi,\pi)^d} e^{i\xi\cdot(y-x)} \mathrm{d}\xi.\end{aligned}$$

The integral is 0 if $x \neq y$ and $(2\pi)^d$ if $x = y$, so we finally find $\mathbb{P}(S_n = x)$. □

Our goals in the enxt pages will be to turn this formula into a criterion for recurrence or transience, and more generally into quantitative estimates on the quantity $\mathbb{P}(S_n = x)$.

1.2.2 The Chung–Fuchs criterion

We will prove the following general result for random walks on $\mathbb{R}^d$, which gives a characterization of recurrent random walks in terms of the characteristic function of μ.

Theorem 5. Let S be a random walk on $\mathbb{R}^d$ with step distribution μ, and fix $\alpha > 0$. Then S is recurrent if and only if

$$\lim_{r\to 1} \int_{(-\alpha,\alpha)^d} \mathrm{Re}\left(\frac{1}{1 - r\hat{\mu}(\xi)}\right) = +\infty. \tag{1.6}$$

Proof for $\mu(\mathbb{Z}^d)$ and $\alpha = \pi$. As suggested by Lemma 17, the proof is simpler for random waks with integer values. Indeed, using Lemma 17, for $|r| < 1$, we have

$$\sum_{n\geq 0} r^n \mathbb{P}\left(S_n = 0\right) = \frac{1}{(2\pi)^d} \sum_{n\geq 0} \int_{(-\pi,\pi)^d} r^n \hat{\mu}(\xi)^n \mathrm{d}\xi = \frac{1}{(2\pi)^d} \int_{(-\pi,\pi)^d} \frac{\mathrm{d}\xi}{1 - r\hat{\mu}(\xi)},$$

where we could use the Fubini formula because $|r| < 1$. The left-hand side is real, so we can take the real part on the right-hand side. When $r \to 1$, the left-hand side goes to $\sum_{n\geq 0} \mathbb{P}(S_n = 0)$. We know that S is recurrent if and only if this limit is infinite and the result follows. □

Remark 18. Our proof does actually better: in the transient case (1.6) gives the expected number of vists of S at 0. Since this is a geometric random variable by the strong Markov property, by taking the inverse, we recover the parameter of this geometric variable, which also the probability for the walk to never come back to 0. For example, for the simple random walk on $\mathbb{Z}^3$, we find

$$\begin{aligned}
\mathbb{P}\left(\forall n \geq 1,\, S_n \neq 0\right) &= \frac{(2\pi)^3}{3} \left(\int_{(-\pi,\pi)^3} \frac{\mathrm{d}\xi_1\, \mathrm{d}\xi_2\, \mathrm{d}\xi_3}{3 - \cos\xi_1 - \cos\xi_2 - \cos\xi_3} \right)^{-1} \\
&= \frac{256\pi^6}{\sqrt{6}\,\Gamma(1/24)\,\Gamma(5/24)\,\Gamma(7/24)\,\Gamma(11/24)} \\
&\approx 0.3405...
\end{aligned}$$

The first inequality follows directly from the proof, and I have no idea how to obtain the second.

We now move on to the proof for the general case. We will only treat the case $d = 1$ in details, but the general case is very similar. We want to estimate the sum $\sum_{n\geq 0} \mathbb{P}\left(|S_n| < x\right)$ for some fixed $x > 0$. The reason why the argument for $\mathbb{Z}^d$ cannot be directly adapted is that the Fourier transform of the indicator $\mathbb{1}_{|x|<1}$ is not integrable. Therefore, it is not directly possible to compute $\mathbb{P}\left(|S_n| < x\right)$ using Fourier inversion as in Lemma 17. In order to solve this problem, we will approximate this indicator by a smoother function. Therefore, we fiw a constant $\beta > 0$ and for $x \in \mathbb{R}$ we write

$$f(x) = \left(1 - \frac{|x|}{\beta}\right)^+.$$

Lemma 19.

- For any $t \in \mathbb{R}$, we have
$$\hat{f}(t) = \int_{\mathbb{R}} f(x) e^{itx} \, \mathrm{d}x = \beta \left(\frac{\sin \beta t/2)}{\beta t/2} \right)^2.$$
- For any $x \in \mathbb{R}$, we have
$$\hat{\hat{f}}(x) = \int_{\mathbb{R}} \hat{f}(t) e^{itx} \, \mathrm{d}t = 2\pi f(x).$$

Proof. The first item is a direct computation using an integration by parts, and the second item follows from the Fourier inversion formula, which applies because $\hat{f}$ is integrable. □

Since we are only interested in the finiteness of an integral, we can afford to lose constant factors, so we will crudely bound $\mathbb{1}_{|x|<\beta}$ by either $\hat{f}$ (for the upper bound) or f (for the lower bound).

More precisely, let us fix $\beta > 0$. There is a constant $c > 0$ such that fro all x, we have $\mathbb{1}_{|x|<1/\beta} \leq c\hat{f}(x)$, so we can write

$$\mathbb{P}\left(|S_n| < \frac{1}{\beta} \right) \leq c\mathbb{E}\left[\hat{f}(S_n) \right] = c \int_{\mathbb{R}} \mathbb{E}\left[e^{itS_n} \right] f(t) \, \mathrm{d}t = c \int_{\mathbb{R}} f(t) \hat{\mu}(t)^n \, \mathrm{d}t,$$

where we could use the Fubini formula because f is integrable. We can now take the real part of the right-hand side, multiply by r^n for some $|r| < 1$ and sum over n to get

$$\sum_{n \geq 0} r^n \mathbb{P}\left(|S_n| < \frac{1}{\beta} \right) \leq c \int_{\mathbb{R}} \mathrm{Re}\left(\frac{f(t)}{1 - r\hat{\mu}(t)} \right) \mathrm{d}t \leq c \int_{-\beta}^{\beta} \mathrm{Re}\left(\frac{1}{1 - r\hat{\mu}(t)} \right) \mathrm{d}t,$$

where in the end we used that f has compact support in $[-\beta, \beta]$ (this is where it is important that we bounded the indicator using $\hat{f}$ and not f). In particular, if the walk S is recurrent, the left-hand side goes to $+\infty$ as $r \to 1$, so the right-hand side also does, which proves one direction of the criterion by taking $\beta = \alpha$.

For the other direction, we need to bound the indicator from *below*, so we write $\mathbb{1}_{|x|<\beta} \geq f(x)$ (the lower bound must have compact support, so it is important to use f and not $\hat{f}$).We can then write

$$\mathbb{P}\left(|S_n| < \beta \right) \geq \mathbb{E}\left[f(S_n) \right] = \frac{1}{2\pi} \mathbb{E}\left[\hat{\hat{f}}(S_n) \right] = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(t) \hat{\mu}(t) \, \mathrm{d}t.$$

Again, we can take the real part, sum over n and apply Fubini to get, for $|r| < 1$:

$$\sum_{n \geq 0} r^n \mathbb{P}\left(|S_n| < \beta \right) \geq \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{Re}\left(\frac{\hat{f}(t)}{1 - r\hat{\mu}(t)} \right) \mathrm{d}t.$$

The function that we integrate is nonnegative and is bounded from below by a constant $c > 0$ on the interval $\left[-\frac{1}{\beta}, \frac{1}{\beta} \right]$, so we can write

$$\sum_{n \geq 0} r^n \mathbb{P}\left(|S_n| < \beta \right) \geq \frac{c}{2\pi} \int_{-1/\beta}^{1/\beta} \mathrm{Re}\left(\frac{1}{1 - r\hat{\mu}(t)} \right) \mathrm{d}t.$$

If the walk is transient, the left-hand side is bounded as $r \to 1$, so the right-hand side as well, which proves the second direction by taking $\beta = \frac{1}{\alpha}$.

Finally, the proof in any dimension $d \geq 1$ is exactly the same. We just need to replace the function $t \to f(t)$ on $\mathbb{R}$ by the function $(t_1, \ldots, t_d) \to \prod_{i=1}^d f(t_i)$ on $\mathbb{R}^d$, and all the proof adapts well.

Strong Chung–Fuchs criterion. We could wonder if the additional parameter r is really necessary in the Chung–Fuchs criterion. It turns out that we can get rid of it, but then the proof becomes much more complicated, and out of the scope of this course.

Theorem 6. Let $\alpha > 0$. The random walk S is recurrent if and only if

$$\int_{(-\alpha,\alpha)^d} \mathrm{Re}\left(\frac{1}{1-\hat{\mu}(\xi)}\right) \mathrm{d}\xi = +\infty.$$

Note that if S is transient, since $\mathrm{Re}\left(\frac{1}{1-r\hat{\mu}(\xi)}\right) \geq 0$, the Fatou lemma shows

$$\int_{(-\alpha,\alpha)^d} \mathrm{Re}\left(\frac{1}{1-\hat{\mu}(\xi)}\right) \mathrm{d}\xi \leq \liminf_{r\to 1} \int_{(-\alpha,\alpha)^d} \mathrm{Re}\left(\frac{1}{1-r\hat{\mu}(\xi)}\right) \mathrm{d}\xi < +\infty,$$

so the reverse implication is easy. The difficult implication is the direct one, and the proof can be found in [6, Chapter 8]. It is not purely Fourier-analytic, and relies on additional random walk estimates.

1.2.3 Applications

We can now apply the Chung–Fuchs criterion to obtain general results about the behaviour of random walks depending on the dimension.

Theorem 7.

- For $d = 1$, if $\mathbb{E}[|X|] < +\infty$ and $\mathbb{E}[X] = 0$, then S is recurrent.
- For $d = 2$, if $\mathbb{E}\left[|X|^2\right] < +\infty$ and $\mathbb{E}[X] = 0$, then S is recurrent.
- For $d \geq 3$, if the step distribution μ is not supported by a linear hyperplane of $\mathbb{R}^d$, then S is transient.

Note that for the one-dimensionl part, we recover a result proved in Section 1.1.3.

Proof.

- Let us start with the one-dimensional statement. We write $\hat{\mu}(t) = a(t) + ib(t)$ with $a(t), b(t) \in [-1, 1]$. Then we have

$$\mathrm{Re}\left(\frac{1}{1-r\hat{\mu}(t)}\right) = \frac{1-r\cdot a(t)}{(1-r\cdot a(t))^2 + r^2 b(t)^2} \geq \frac{1-r}{(1-r\cdot a(t))^2 + r^2 b(t)^2}.$$

But since $\mathbb{E}[X] = 0$, we know that $\hat{\mu}(t) = 1 + o(t)$ as $t \to 0$, so let us fix $\varepsilon > 0$. Then we can find $\delta > 0$ such that for $0 \leq t \leq \delta$, we have $|1 - a(t)| \leq \varepsilon t$ and $|b(t)| \leq \varepsilon t$. This implies

$$(1 - r\cdot a(t))^2 = \left((1-r) + r(1-a(t))\right)^2 \geq 2(1-r)^2 + 2\varepsilon^2 r^2 t^2,$$

so

$$\mathrm{Re}\left(\frac{1}{1-r\hat{\mu}(t)}\right) \geq \frac{1-r}{2(1-r)^2 + 3\varepsilon^2 r^2 t^2}.$$

Since $\mathrm{Re}\left(\frac{1}{1-r\hat{\mu}(t)}\right) \geq 0$ for all $t \in \mathbb{R}$, this becomes

$$\begin{aligned}
\int_1^{-1} \mathrm{Re}\left(\frac{1}{1-r\hat{\mu}(t)}\right) \mathrm{d}t &\geq \int_0^{\delta} \frac{1-r}{2(1-r)^2 + 3\varepsilon^2 r^2 t^2} \,\mathrm{d}t \\
&= \frac{1}{1-r} \int_0^{\delta} \frac{\mathrm{d}t}{2 + 3\left(\frac{t\varepsilon r}{1-r}\right)^2} \\
&= \frac{1}{1-r} \int_0^{\frac{\delta\varepsilon r}{1-r}} \frac{1}{2+3u^2} \frac{1-r}{\varepsilon r} \,\mathrm{d}u.
\end{aligned}$$

As $r \to 1$, this converges to $\frac{1}{\varepsilon}\int_0^{+\infty} \frac{\mathrm{d}u}{2+3u^2}$. Since this holds for any $\varepsilon > 0$, we have proved that the left-hand side goes to $+\infty$ as $r \to 1$.

- We move on to the two-dimensional part. We assume $\mathbb{E}\left[|X|^2\right] < +\infty$, and denote by Q the covariance matrix of X. By the Fatou lemma, we have

$$\liminf_{r\to 1} \int_{(-\pi,\pi)^2} \mathrm{Re}\left(\frac{1}{1-r\hat{\mu}(t)}\right) \mathrm{d}\xi \geq \int_{(-\pi,\pi)^2} \mathrm{Re}\left(\frac{1}{1-\hat{\mu}(\xi)}\right) \mathrm{d}\xi.$$

When $\xi \to 0$, we know that $\frac{1}{1-\hat{\mu}(\xi)} \sim \frac{2}{\xi^t Q \xi}$, so $\mathrm{Re}\left(\frac{1}{1-\hat{\mu}(\xi)}\right) \sim \frac{2}{\xi^t Q \xi}$. In particular, there is $c > 0$ such that $\mathrm{Re}\left(\frac{1}{1-\hat{\mu}(\xi)}\right) \geq \frac{c}{|\xi|^2}$ for all $\xi \in (-\pi, \pi)$. Therefore, we can write

$$\int_{(-\pi,\pi)^2} \mathrm{Re}\left(\frac{1}{1-\hat{\mu}(\xi)}\right) \mathrm{d}\xi \geq \int_{B(0,1)} \frac{\mathrm{d}\xi}{|\xi|^2} = \int_0^1 \frac{2\pi s \,\mathrm{d}s}{s^2} = +\infty,$$

where the second integral is over the ball of radius 1 around 0 and the third is obtained by taking polar coordinates. This concludes the proof.

- We now prove transience for $d \geq 3$. For this, we will need a general estimate showing that the characteristic function does not stay too close to 1.

Lemma 20. Assume that the support of μ is not included in an affine hyperplane. Then there is a constant $c > 0$ such that, for ξ in a neighbourhood of 0, we have

$$|\hat{\mu}(\xi)| \leq 1 - c|\xi|^2.$$

Proof. We first prove the lemma in the case where μ has bounded support. We can then denote by m its mean and by Q its covariance matrix. As $\xi \to 0$, we have

$$\hat{\mu}(\xi) = 1 + im \cdot \xi - \frac{\xi^t Q \xi}{2} + o\left(|\xi|^2\right).$$

It follows that

$$\begin{aligned}
|\hat{\mu}(\xi)|^2 &= 1 - \xi^t Q \xi + |m \cdot \xi|^2 + o\left(|\xi|^2\right) \\
&= 1 - \sum_{i,j=1}^{d} \xi_i \xi_j \mathbb{E}\left[X^{(i)} X^{(j)}\right] + \mathbb{E}[X \cdot \xi]^2 + o\left(|\xi|^2\right) \\
&= 1 - \mathrm{Var}(X \cdot \xi) + o\left(|\xi|^2\right) \\
&= 1 - |\xi|^2 \mathrm{Var}\left(X \cdot \frac{\xi}{|\xi|}\right) + o\left(|\xi|^2\right).
\end{aligned}$$

However, the support of μ is not included in any hyperplane. this means that for any u with $|u| = 1$, the variable $X \cdot u$ is not deterministic so $\mathrm{Var}(X \cdot u) > 0$. Therefore, we have

$$|\hat{\mu}(\xi)| \leq 1 - \frac{|\xi|^2}{2} \min_{|u|=1} \mathrm{var}(X \cdot u) + o\left(|\xi|^2\right),$$

where $\min_{|u|=1} \mathrm{var}(X \cdot u) > 0$ by compactness, which proves the lemma if μ is compactly supported.

For a general μ, let $A \subset \mathbb{R}^d$ be a bounded subset such that the restriction of μ to A is not supported by a hyperplane. This implies $\mu(A) > 0$, so we can define the probability measure $\mu_A = \frac{\mu(A\cap\cdot)}{\mu(A)}$ on A. Then can write

$$\begin{aligned}|\hat{\mu}(\xi)| &= |\mu(A)\hat{\mu}_A(\xi) + \mu(A^c)\hat{\mu}_{A^c}(\xi)| \\ &\leq \mu(A)\,|\hat{\mu}_A(\xi)| + (1-\mu(A))\,|\hat{\mu}_{A^c}(\xi)| \\ &\leq \mu(A)\left(1 - c|\xi|^2\right) + 1 - \mu(A) \\ &= 1 - c\mu(A)|\xi|^2,\end{aligned}$$

where we applied our lemma to the compactly supported measure μ_A. □

We can now conclude the proof of transience for $d \geq 3$. We know that there is $c > 0$ such that for any ξ in some neighbourhood of 0 and for any $0 < r < 1$, we have

$$\mathrm{Re}\left(\frac{1}{1 - r\hat{\mu}(\xi)}\right) = \frac{\mathrm{Re}\,(1 - r\hat{\mu}(\xi))}{\mathrm{Re}\,(1 - r\hat{\mu}(\xi))^2 + \mathrm{Im}\,(1 - r\hat{\mu}(\xi))^2} \leq \frac{1}{\mathrm{Re}\,(1 - r\hat{\mu}(\xi))}.$$

We know that there is a neighbourhood of 0 where $\mathrm{Re}\,(\hat{\mu}(\xi)) > 0$. In this neghbourhood, we can therefore write

$$\mathrm{Re}\left(\frac{1}{1 - r\hat{\mu}(\xi)}\right) \leq \frac{1}{\mathrm{Re}\,(1 - \hat{\mu}(\xi))} = \frac{1}{1 - Re\,(\hat{\mu}(\xi))} \leq \frac{1}{1 - |\hat{\mu}(\xi)|} \leq \frac{1}{|\xi|^2}.$$

This is integrable in a neighbourhood of 0 for $d \geq 3$, which proves transience in the case where μ is not supported in any affine hyperplane.

Finally, if μ is supported in an affine hyperplane which is not linear, there is a vector $v \in \mathbb{R}^d$ and a constant $a \neq 0$ such that $v \cdot X_1 = a$ almost surely. It follows that $v \cdot S_n = an$ almost surely for all n, so $|S_n| \to +\infty$ and the walk is transient.

□

1.2.4 The local central limit theorem

Let us first explain what we want to do in the one-dimensional case. Consider a random walk supported in $\mathbb{Z}$. For random walks with mean 0 and finite variance 1, the *central limit theorem* shows that, for $a < b$, we have

$$\mathbb{P}\left(S_n \in \left[a\sqrt{n}, b\sqrt{n}\right]\right) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\,\mathrm{d}t.$$

From here, it is tempting to "shrink the interval to one point" and say that for any $x \in \mathbb{Z}$ with $x \approx t\sqrt{n}$, we have

$$\mathbb{P}\left(S_n = x\right) \approx \frac{1}{\sqrt{2\pi}\sqrt{n}} e^{-t^2/2}. \tag{1.7}$$

This would be a strong reinforcement of the CLT, which we call a *local CLT*. Roughly speaking, such a result means that if we condition S_n on being in a small box of side length $\varepsilon\sqrt{n}$ with ε small, then S_n is approximately uniformly distributed over all integer points of the box. The goal of this section is to turn this into a precise statement under the right assumptions.

For all this section, we will assume $\mu(\mathbb{Z}^d) = 1$, i.e. we work with random walks with values in $\mathbb{Z}^d$. In particular, the state space is countable, so all the technology of general Markov chains

applies to our setting. We first notice that (1.7) is completly wrong for the simple random walk on $\mathbb{Z}$! Indeed, if x and n have different parity, then $\mathbb{P}(S_n = x) = 0$. To get rid of such parity issues, we need to consider *aperiodic* walks.

Definition 21. We say that μ is *aperiodic* if the associated random walk is an irreducible, aperiodic Markov chain on $\mathbb{Z}^d$. Equivalently, this means that for any $x \in \mathbb{Z}^d$, we have $\mathbb{P}(S_n = x) > 0$ for n large enough.

Since our main tool will again be Fourier transforms, we first turn this assumption into a statement on the characteristic function.

Lemma 22. Assume that μ is aperiodic. Then for $\xi \in \mathbb{R}^d$, we have

$$|\hat{\mu}(\xi)| = 1 \quad \text{if and only if} \quad \xi \in (2\pi\mathbb{Z})^d.$$

Note that for the simple random walk on $\mathbb{Z}^d$, this condition is not satisfied since

$$\hat{\mu}(\pi, \dots, \pi) = (-1)^d.$$

Proof. If $\xi \in (2\pi\mathbb{Z})^d$, an immediate computation gives $\hat{\mu}(\xi) = 1$.

On the other hand, if $|\hat{\mu}(\xi)| = 1$, then for all $n \geq 1$, we can write

$$1 = |\hat{\mu}(\xi)^n| = \left|\mathbb{E}\left[e^{i\xi \cdot S_n}\right]\right| = \left|\sum_{x \in \mathbb{Z}^d} \mathbb{P}(S_n = x)\, e^{i\xi \cdot x}\right| \leq \sum_{x \in \mathbb{Z}^d} \mathbb{P}(S_n = x) \left|e^{i\xi \cdot x}\right| = 1$$

by the triangle inequality. In particular, we must have equality in the triangle inequality, so the complex numbers $e^{i\xi \cdot x}$ such that $\mathbb{P}(S_n = x) > 0$ are positively aligned. If n is chosen large enough, the set of such x contains 0 and all the vectors e_i for $1 \leq i \leq d$ of the canonical basis of $\mathbb{R}^d$. It follows that $1, e^{i\xi_1}, \dots, e^{i\xi_d}$ are positively aligned, so they are all equal to 1, so $\xi_1, \dots, \xi_d$ are all multiples of 2π. □

This lemma has the following immediate consequence, which will be very useful later.

Lemma 23. If μ is aperiodic, then there is a constant $c > 0$ such that

$$\forall \xi \in [-\pi, \pi]^d, \quad |\hat{\mu}(\xi)| \leq 1 - c|\xi|^2.$$

Proof. If the Markov chain associated to μ is irreducible, then in particular it is not included in a hyperplane. Therefore, by Lemma 20, there are an open neighbourhood $\mathcal{N} \subset [-\pi, \pi]^d$ of 0 and a constant c_1 such that $|\hat{\mu}(\xi)| \leq 1 - c|\xi|^2$ for all $\xi \in \mathcal{N}$. On the other hand, by lemma 22 and by compactness, there is a constant $c_2 > 0$ such that for all $\xi \in [-\pi, \pi]^d \backslash \mathcal{N}$, we have

$$|\hat{\mu}(\xi)| \leq 1 - c_2 \leq 1 - \frac{c_2}{d(0, \mathcal{N}^c)^2}|\xi|^2.$$

The result follows by taking $c = \min\left(c_1, \frac{c_2}{d(0,\mathcal{N}^c)^2}\right)$. □

We can now state our local central limit theorem.

Theorem 8. Let μ be an aperiodic distribution on $\mathbb{Z}^d$ with mean $m \in \mathbb{R}^d$ and finite covariance matrix Q. Let $\gamma_Q(x) = \frac{1}{\sqrt{2\pi \det(Q)}} \exp\left(-\frac{x^t Q x}{2}\right)$ be the density on $\mathbb{R}^d$ of a centered Gaussian vector with covariance matrix Q. Then we have

$$\lim_{n\to\infty} \sup_{x\in\mathbb{Z}^d} n^{d/2} \left| \mathbb{P}\left(S_n = x\right) - n^{-d/2} \gamma_Q\left(\frac{x-nm}{\sqrt{n}}\right)\right| = 0.$$

Note that the fact that μ is aperiodic guarantees that Q is positive definite, so the determinant appearing in the denominator is nonzero. Note also that this estimate is uniform over all $x \in \mathbb{Z}^d$, even if $x - nm$ is not of order $\sqrt{n}$. This is not that surprising, since in this case we expect both $\mathbb{P}(S_n = x)$ and $\gamma_Q\left(\frac{x-nm}{\sqrt{n}}\right)$ to be very small.

Proof. We only write down the proof completely for $d = 1$. The starting point is the Fourier inversion formula of Lemma 17:

$$\mathbb{P}\left(S_n = x\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i\xi x} \hat{\mu}(\xi)^n \, \mathrm{d}\xi.$$

Intuitively, Lemma 23 guarantees that if ξ is not very close to 0, then $|\hat{\mu}(\xi)^n|$ will be very small, so we need to focus on values of ξ in a small neighbourhood of 0.

Let us first recenter the process. For this, we define ν as the law of $X_1 - \mathbb{E}[X_1]$, so that $\hat{\nu}(\xi) = e^{-i_x im}\hat{\mu}(\xi)$ for all $\xi \in \mathbb{R}$. Therefore, we can write

$$\mathbb{P}\left(S_n = x\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i\xi x} e^{imn\xi} \hat{\nu}(\xi)^n \, \mathrm{d}\xi.$$

On the other hand ν has mean 0 and variance $\sigma^2 > 0$ by irreducibility, so

$$\hat{\nu}(\xi) = 1 - \frac{\sigma^2}{2}\xi^2 + o(\xi^2) \quad \text{as } \xi \to 0.$$

Informally, we want to write $\hat{\nu}(\xi)^n \approx \exp\left(-\frac{\sigma^2}{2} n\xi^2\right)$, which means that $\hat{\nu}(\xi)^n$ should remain "large" as long as $\xi \lesssim \frac{1}{\sqrt{n}}$. Therefore, we make the change of variable $\xi = \frac{u}{\sqrt{n}}$. We obtain

$$\begin{aligned}
\mathbb{P}\left(S_n = x\right) &= \frac{1}{2\pi} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} e^{-ix\frac{u}{\sqrt{n}}} e^{im\sqrt{n}u} \hat{\nu}\left(\frac{u}{\sqrt{n}}\right)^n \frac{\mathrm{d}u}{\sqrt{n}} \\
&\approx \frac{1}{2\pi\sqrt{n}} \int_{-\infty}^{+\infty} \exp\left(i\left(\sqrt{n}m - \frac{x}{\sqrt{n}}\right)u\right) \exp\left(-\frac{\sigma^2 u^2}{2}\right) \mathrm{d}u \qquad (1.8) \\
&= \frac{1}{2\pi\sqrt{n}} \mathbb{E}\left[\exp\left(i\left(\sqrt{n}m - \frac{x}{\sqrt{n}}\right)\frac{N}{\sigma}\right)\right] \qquad \text{where } N \sim \mathcal{N}(0,1) \\
&= \frac{1}{2\pi\sqrt{n}} \exp\left(-\frac{1}{2}\left(\frac{\sqrt{n}m}{\sigma} - \frac{x}{\sigma\sqrt{n}}\right)^2\right) \\
&= \frac{1}{2\pi\sqrt{n}} \exp\left(-\frac{1}{2\sigma^2}\left(\frac{x-mn}{\sqrt{n}}\right)^2\right).
\end{aligned}$$

To conclude the proof, we need to estimate the error that we made in the approximation (1.8). By the triangle inequality, the error that we made can be bounded by

$$\frac{1}{2\pi\sqrt{n}} \int_{-\infty}^{+\infty} \left| \mathbb{1}_{|u|\le\pi\sqrt{n}} \, \hat{\nu}\left(\frac{u}{\sqrt{n}}\right)^n - e^{-\frac{\sigma^2 u^2}{2}} \right| \mathrm{d}u. \qquad (1.9)$$

In particular, this does not depend on the choice of x, so all we need to prove is that this is $o\left(\frac{1}{\sqrt{n}}\right)$ as $n \to \infty$. The term that is integrated converges simply to 0 as $n \to \infty$. Moreover, using Lemma 23, we can dominate it as follows:

$$\left| \mathbb{1}_{|u| \leq \pi\sqrt{n}} \hat{\nu}\left(\frac{u}{\sqrt{n}}\right)^n - e^{-\frac{\sigma^2 u^2}{2}} \right| \leq \mathbb{1}_{|u| \leq \pi\sqrt{n}} \left(1 - c\frac{u^2}{n}\right)^n + e^{-\frac{\sigma^2 u^2}{2}} \leq 2e^{-c'u^2}$$

for some constant $c' > 0$. This is integrable, so by dominated convergence the integral in (1.9) goes to 0, so (1.9) is $o\left(\frac{1}{\sqrt{n}}\right)$, which concludes the proof. □

1.3 Ladder times and cyclic lemma

In this section, we will only focus on random walks on $\mathbb{R}$ or even $\mathbb{Z}$. Our main goal will be tu understand the probability that a random walk stays positive until time n in terms of n. For this, the tools that we will use will be the invariance of the law of a random walk under various elementary transformations.

1.3.1 Ladder times and heights

In all this section, we will assume $d = 1$, and denote by (S_n) a random walk with step distribution μ on $\mathbb{R}$. We first define the *ladder times and heights* of S by induction (see Figure 1.1 for an illustration).

Definition 24 (Ascending ladder times and heights)**.** We write $T_0^> = 0$ and $H_0^> = 0$ and define recursively, for all $i \geq 1$:

$$T_i^> = \inf\{n > T_{i-1}^> \,|\, S_n > H_{i-1}^>\} \quad \text{and} \quad H_i^> = S_{T_i^>}.$$

We define similarly $(T_i^\geq, H_i^\geq)$ (weak ascending ladder times), $(T_i^<, H_i^<)$ (descending ladder times) and $(T_i^\leq, H_i^\leq)$ (weak descending ladder times).

Note that the times $T_i^>$ are stopping times. Moreover, the strong Markov property implies that the probability that $T_{i+1}^> < +\infty$ conditionally on $T_i^> < +\infty$ is equal to the probability that $T_1^> < +\infty$. In other words, the random variable

$$I^> = \inf\{i \geq 0 | T_i^> = +\infty\}$$

is geometric in the sense that, for all $i \geq 1$, we have

$$\mathbb{P}\left(I^> = i\right) = \mathbb{P}\left(T_1^> < +\infty\right)^{i-1} \times \mathbb{P}\left(T_1^> = +\infty\right).$$

Moreover, also by the strong Markov property applied to the stopping times $T_i^>$, the variables

$$\left(T_i^> - T_{i-1}^>, H_i^< - H_{i-1}^>\right)_{1 \leq i \leq I^>}$$

are i.i.d. and all have the same law as $(T_1^>, H_1^>)$ conditionned on $T_1^> < +\infty$. The same is true if $>$ is replaced by $<$, $\geq$ or $\leq$.

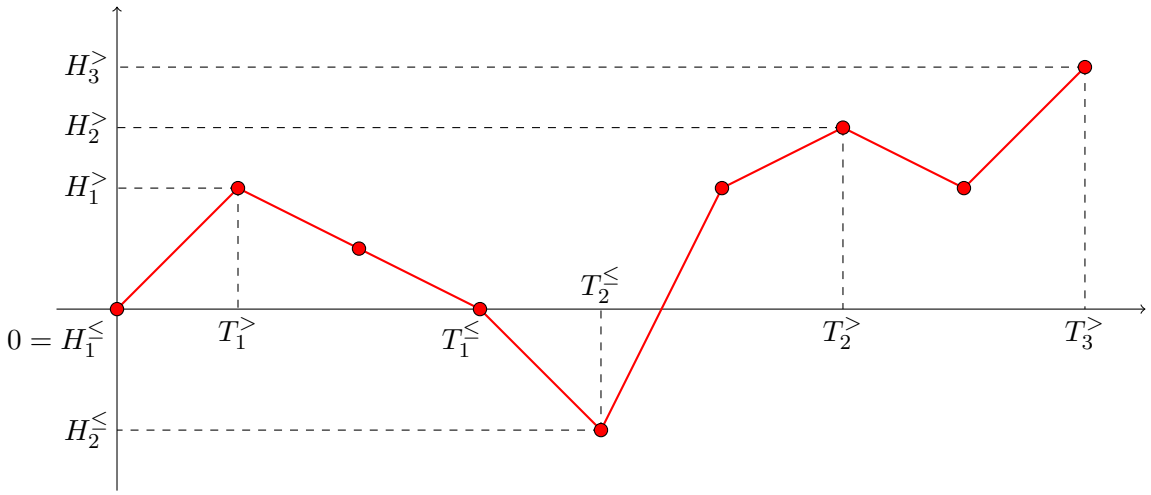

Figure 1.1: An example of walk with the ascending ladder times $(T_i^>, H_i^>)$ and heights, and the weak descending ladder times and heights $(T_i^\leq, H_i^\leq)$.

1.3.2 Duality

The first transformation of our random walk that we will be interested in simply consists of reading its steps in the reverse order. This is equivalent to rotating the graph of the walk (i.e. the red trajectory of Figure 1.1) by 180 degrees.

Lemma 25. For all $n \geq 1$, we have

$$(S_0, S_1, \ldots, S_{n-1}, S_n) \overset{(d)}{=} (S_n - S_n, S_n - S_{n-1}, \ldots, S_n - S_1, S_n)\,.$$

Proof. The increments of the left-hand side are $X_1, \ldots, X_n$, which are i.i.d. with aw μ. the increments of the right-hand side are $X_n, \ldots, X_1$, which are also i.i.d. with law μ. Note that this proof still holds in any dimension. □

Although the last result may seem very easy, it already has several interesting consequences on the expectation of ladder times. Note that the following result does not make any assumption on the step distribution!

Proposition 26.

- If S drifts to $-\infty$, then

$$\mathbb{P}\left(T_1^> = +\infty\right) > 0,\ \mathbb{P}\left(T_1^\geq = +\infty\right) > 0 \quad \text{and} \quad \mathbb{E}\left[T_1^<\right] < +\infty,\ \mathbb{E}\left[T_1^\leq\right] < +\infty.$$

  Moreover, we have $\mathbb{E}\left[T_1^<\right] = \frac{1}{\mathbb{P}\left(T_1^\geq = +\infty\right)}$.

- If S drifts to $+\infty$, then the symetric statement is true.
- If S oscillates, all the ladder times are a.s. finite with infinite expectation.

Proof. For any $n \geq 0$, we can write using duality

$$\begin{aligned}\mathbb{P}\left(T_1^< > n\right) &= \mathbb{P}\left(S_0, S_1, \dots, S_n \geq 0\right) \\ &= \mathbb{P}\left(S_n - S_n, S_n - S_{n-1}, \dots, S_n - S_1, S_n \geq 0\right) \\ &= \mathbb{P}\left(S_n \geq S_0, \dots, S_{n-1}\right) \\ &= \mathbb{P}\left(\exists i,\, T_i^{\geq} = n\right).\end{aligned}$$

Summing over n, we can compute the expectation:

$$\mathbb{E}\left[T_1^<\right] = \sum_{n\geq 0} \mathbb{P}\left(T_1^< > n\right) = \sum_{n\geq 0} \mathbb{E}\left[\sum_{i\geq 0} \mathbb{1}_{T_i^{\geq}=n}\right] = \mathbb{E}\left[\sum_{i\geq 0} \mathbb{1}_{T_i^{\geq}<+\infty}\right] = \mathbb{E}\left[I^{\geq}\right] = \frac{1}{\mathbb{P}\left(T_1^{\geq} = +\infty\right)}$$

where we used Fubini and, in the last equality, the fact that $I^{\geq}$ is geometric with parameter $\mathbb{P}\left(T_1^{\geq} = +\infty\right)$. □

1.3.3 The cyclic lemma

In this subsection, we assume that our random walk S has values in $\mathbb{Z}$ and is *negative skip-free*, i.e. for all n, we almost surely have

$$X_n \in \{-1, 0, 1, 2, \dots\}.$$

In other words, the random walk cannot make "large negative jumps". The interest of this property is that if $0 > a > b$, a negative skip-free walk must hit the value a before hitting the value b. In particular, this implies $H_i^< = -i$ for all $i \geq 0$.

The idea behind the cyclic lemma is to apply a cyclic permutation to the steps of a random walk. We will first state a deterministic version of the cyclic lemma, and then apply it to random walks.

If $x_1, \dots, x_n \in \{-1, 0, 1, 2, \dots\}$, we define a trajectory s by

$$s_i = x_1 + \cdots + x_i \quad \text{for } 0 \leq i \leq n.$$

For all $0 \leq \ell \leq n$, we define the *cyclic shift* $s^{(\ell)}$ of s by

$$s_j^{(\ell)} = \begin{cases} x_{\ell+1} + \cdots + x_{\ell+j} & \text{if } 0 \leq j \leq n-\ell, \\ x_{\ell+1} + \cdots + x_n + x_1 + \cdots + x_{j-(n-\ell)} & \text{if } n-\ell \leq j \leq n. \end{cases}$$

In other words, we make the same steps (x_i) as in the trajectory s, but in the following order:

$$x_{\ell+1}, x_{\ell+2}, \dots, x_n, x_1, x_2, \dots, x_\ell.$$

See also Figure 1.2 for an example. Note that in particular the overall sum of the steps stays the same, so $s_n^{(\ell)} = s_n$.

Lemma 27. We assume $s_n = -k$ with $k \geq 1$. Then there are exactly k values of $0 \leq \ell \leq n-1$ such that n is the hitting tme of $-k$ for the shifted walk $s^{(\ell)}$.

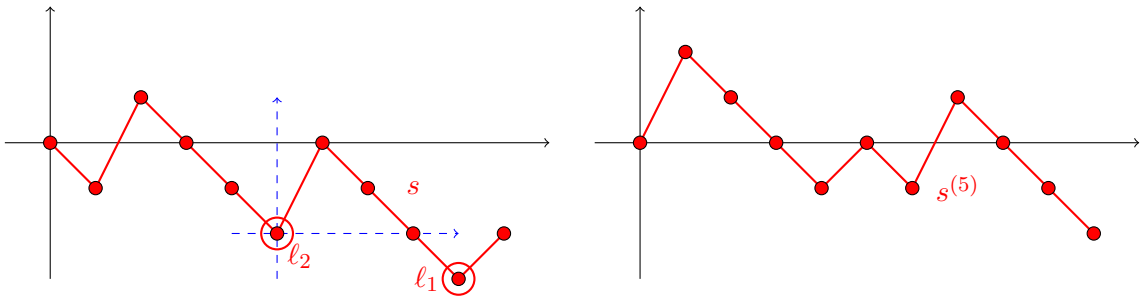

Figure 1.2: On the left, a negative skip-free trajectory s (with $n = 10$ and $k = 2$). On the right, the shifted trajectory $s^{(5)}$. The shifted trajectory is equivalent to "starting the walk from the blue axes", performing first the step on the right of the starting point (from left to right), and then the steps on its left (again from left to right). The two indices ℓ_1, ℓ_2 for which the shifted walk hits -2 first at time n correspond to the circled points.

Proof. • Looking at Figure 1.1 suggests that the "right" values of ℓ should be hitting times of negative values for s, and that these values should be as small as possible. Therefore, we write $m = \min\{s_0, s_1, \dots, s_n\} \leq -k$. For $0 \leq i \leq k-1$, we define

$$\ell_i = \min\{j \geq 0 | s_j = m + i\}.$$

Note that these indices are well-defined because s is negative skip-free and $m+i \leq m+k \leq 0$. We will check that the values of ℓ for which $s^{(\ell)}$ hits $-k$ first at time n are exactly $\ell_0, \dots, \ell_{k-1}$, which will prove the lemma.

- Let us start by checking that the ℓ_i indeed work. For this, we fix an index $0 \leq j \leq n-1$ and check that $s_j^{(\ell_i)} > -k$.
  - If $0 \leq j \leq n - \ell_i$, then we have
  $$s_j^{(\ell_i)} = s_{\ell_i+j} - s_{\ell_i} = s_{\ell_i+j} - (m+i) \geq m - (m+i) = -i > -k,$$
  where the first inequality holds by definition of m.
  - If $n - \ell_i \leq j \leq n-1$, then $j - (n - \ell_i) < \ell_i$, so by definition of ℓ_i we have $s_{j-(n-\ell_i)} > s_{\ell_i} = m + i$. Therefore, we can write
  $$s_j^{(\ell_i)} = s_n - s_{\ell_i} + s_{j-(n-\ell_i)} = -k - (m+i) + s_{j-(n-\ell_i)} > -k.$$
  - Since $s_n^{(\ell_i)} = s_n = -k$, it follows that n is the hitting time of $-k$ for $s^{(\ell_i)}$, so ℓ_i indeed works.
- On the other hand, let $0 \leq \ell \leq n-1$ be such that our hitting time condition holds. Then for any $0 \leq j < \ell$, we have
$$-k < s_{n-\ell+j}^{(\ell)} = s_n - s_\ell + s_j = -k - s_\ell + s_j,$$
so $s_j > s_\ell$. This means that s is a descending ladder time, i.e. it is the hitting time of a negative value $m+i$ with $0 \leq i \leq -m$. On the other hand, let j be the hitting time of m for s. Then we have $j \geq \ell$ (by the skip-free property) and
$$-k < s_{j-\ell}^{(\ell)} = s_j - s_\ell = m - (m+i) = -i.$$

It follows that $i < k$, so ℓ is one of the values ℓ_i for $0 \leq i \leq k-1$, which concludes the proof.

□

1.3.4 Kemperman's formula and applications

We can now move on to probabilistic consequences of the cyclic lemma. As suggested by the statement of the lemma, we first use it to describe the law of descending ladder times for a negative skip-free random walk.

Proposition 28. Let S be a negative skip-free random walk on $\mathbb{Z}$. Then $T_k^<$ is the hitting time of $-k$ and for all $k, n \geq 1$, we have

$$\mathbb{P}\left(T_k^< = n\right) = \frac{k}{n}\mathbb{P}\left(S_n = -k\right).$$

Note that the right-hand side is usually way easier to compute (using Fourier techniques or, for some simple distributions, just binomial coefficients).

Proof. First, it follows immediately from the skip-free property that $T_k^<$ is the hitting time of $-k$.

To prove the formula, we will perform a random shift, i.e. we shift S by a random index U chosen uniformly at random in $\{0, 1, \ldots, n-1\}$, independently of S. Note that for any choice of ℓ, the shifted walk $S^{(\ell)}$ has the same law as S (its increments are i.i.d. with law μ). Therefore $S^{(U)}$ has the same law as S, so

$$\mathbb{P}\left(T_k^< = n\right) = \mathbb{P}\left(S^{(U)} \text{ hits } -k \text{ first at time } n\right).$$

Note that in particular, this last event may only occur if $S_n^{(U)} = S_n = -k$. We will now condition on S to apply the cyclic lemma. More precisely, by the cyclic lemma, we have

$$\mathbb{P}\left(S^{(U)} \text{ hits } -k \text{ first at time } n | S\right) = \frac{k}{n}\mathbb{1}_{S_n=-k}.$$

By taking the expectation, we obtain Kemperman's formula. □

Example 29. For the simple random walk, Kemperman's formula gives, for all $n \geq 0$:

$$\mathbb{P}\left(T_1^< = 2n+1\right) = \frac{1}{2n+1}\mathbb{P}\left(S_{2n+1} = -1\right) = \frac{1}{2n+1} \times \frac{1}{2^{2n+1}}\binom{2n+1}{n}.$$

This formula is not new, and can also be obtained by the reflexion principle. However, in order to compute this law using the reflexion principle, we needed to use the fact that the simple random walk on $\mathbb{Z}$ is both skip-free and symetric. This means that it is basically the only model where the law of $T_1^<$ can be computed using the reflexion principle. On the other hand, the cyclic lemma approach is able to handle way more models.

We can now study the probability for a random walk to stay positive for a certain amount of time.

Theorem 9 (Ballot theorem)**.** Let S be a *positive* skip-free random walk on $\mathbb{Z}$. Then for all $k, n \geq 1$ such that $\mathbb{P}(S_n = k) > 0$, we have

$$\mathbb{P}(\forall 1 \leq i \leq n,\, S_i > 0 | S_n = k) = \frac{k}{n}.$$

Note that the assumption is now that S is *positive* skip-free and not negative.

Proof. We want to turn the probability of hitting the *terminal* value of the walk before the end into the probability of hitting the *initial value* after the beginning. Therefore, it makes sense to use the duality property (Lemma 25). By duality, we can write

$$\begin{aligned}
\mathbb{P}(\forall 1 \leq i \leq n,\, S_n > 0 \text{ and } S_n = k) &= \mathbb{P}(n \text{ is an ascending ladder time and } S_n = k) \\
&= \mathbb{P}\left(T_k^> = n \text{ and } S_n = k\right) \\
&= \mathbb{P}\left(T_k^> = n\right) \\
&= \frac{k}{n}\mathbb{P}(S_n = k),
\end{aligned}$$

where the second equality uses the skip-free assumption, and the last one comes from the Kemperman formula applied to $-S$ (which is negative skip-free). □

The name "ballot theorem" comes from the following interpretation. Assume that on an election, two candidates A and B get respectively a and b votes, with $a > b$. We also assume that the votes are counted in a uniform random order. We denote by S_i the difference between the number of votes for A and the number of votes for B after exactly i votes have been counted. Then $(S_i)_{0 \leq i \leq n}$ is a simple random walk on $\mathbb{Z}$ conditioned on $S_{a+b} = a - b$. Therefore, the ballot theorem means that the probabilty that A stayed ahead of B during all the counting is equal to $\frac{a-b}{a+b}$.

We can now pass to the limit to obtain the probability that a random walk stays poitive forever.

Theorem 10. Let S be a *positive* skip-free random walk on $\mathbb{Z}$ with $\mathbb{E}[S_1] > 0$. Then

$$\mathbb{P}(\forall i \geq 1,\, S_i > 0) = \mathbb{E}[S_1] > 0.$$

Proof. By the ballot theorem, we can write

$$\begin{aligned}
\mathbb{P}(\forall i \geq 1,\, S_i > 0) &= \lim_{n \to +\infty} \mathbb{P}(\forall 1 \leq i \leq n,\, S_i > 0) \\
&= \lim_{n \to +\infty} \mathbb{E}\left[\mathbb{P}(\forall i \geq 1,\, S_i > 0 | S_n)\right] \\
&= \lim_{n \to +\infty} \mathbb{E}\left[\frac{S_n}{n} \mathbb{1}_{S_n > 0}\right].
\end{aligned}$$

By the law of large number (and the assumption $\mathbb{E}[S_1] > 0$), we know that $\frac{S_n}{n}\mathbb{1}_{S_n>0}$ converges almost surely to $\mathbb{E}[S_1]$. Moreover, this random variable is nonnegative (because of the indicator) and bounded by 1 (the walk is positive skip-free, so $S_n \leq n$). Therefore, by the dominated convergence theorem, the last limit is equal to $\mathbb{E}[S_1]$. □

We finish this section by computing the probability to stay positive forever for a *negative* skip-free walk. Although the result looks very similar, the method of the proof will be very different (and perhaps more natural, without the cyclic lemma trick) than the proof of Theorem 10.

Theorem 11. Let S be a negative skip-free random walk on $\mathbb{Z}$ with step distribution μ. Then the probability $\mathbb{P}(\forall n \geq 0,\, S_n \geq 0)$ is equal to $1-\alpha$, where α is the smallest solution in $[0,1]$ to the equation

$$\alpha = \sum_{k=-1}^{+\infty} \mu(k)\alpha^{k+1}. \tag{1.10}$$

Proof. If α is the smallest solution to (1.10) in $[0,1]$, we want to prove $\mathbb{P}(T_1^< < +\infty) = \alpha$. For this, the idea of the proof will be to condition on the first step S_1. By the weak Markov property, we know that conditionally on S_1, the law of $(S_n)_{n\geq 1}$ is the law of (S_n) translated by S_1.

We now define the stopping times $(\tau_i)_{i\geq 0}$ by

$$\tau_i = \inf\{n \geq 1 | S_n = S_1 - i\}.$$

In other words, the $\tau_i - 1$ are the descending ladder times for the random walk $(S_{n+1} - S_1)_{n\geq 1}$. Therefore, the weak Markov property (applied to time 1) implies that $(\tau_i)_{i\geq 0}$ is independent from S_1, and the strong Markov property (applied to the times τ_i) implies that for any $i \geq 0$, we have

$$\mathbb{P}(\tau_i < +\infty) = \mathbb{P}\left(T_1^< < +_i nfty\right)^i.$$

On the other hand, our walk hits the value -1 if and only if τ_{S_1+1} is finite. By summing over all possible values of S_1, we obtain It follows that

$$\begin{aligned}\mathbb{P}\left(T_1^< < +\infty\right) &= \sum_{k=-1}^{+\infty} \mathbb{P}(S_1 = k \text{ and } \tau_{k+1} < +\infty) \\ &= \sum_{k=-1}^{+\infty} \mu(k)\mathbb{P}\left(T_1^< < +\infty\right)^{k+1},\end{aligned}$$

so $\mathbb{P}(T_1^< < +\infty)$ is a solution to (1.10).

We now want to ensure that it is the smallest one. For this, we will show that there can only be very few solutions to (1.10) in $[0,1]$. Indeed, the function $F(x) = \sum_{k\geq -1} \mu(k)x^{k+1}$ is convex, with $F(1) = 1$ and $F'(1) = \mathbb{E}[S_1] + 1 \in [1, +\infty]$, so the equation $F(x) = x$ has at ost two solutions (unless $F(x) = x$ on a nontrivial interval). More precisely, we can distinguish the following cases.

- If $\mathbb{E}[S_1] > 0$, then F is strictly convex because there is $k \geq 1$ such that $\mu(k) > 0$ and we have $F(0) \geq 0$ and $F'(1) > 1$, so $F(1-\varepsilon) < 1-\varepsilon$ for $\varepsilon > 0$ small enough (see the left part of Figure 1.3). Int follows that $F(\alpha) = \alpha$ has exactly two solutions in $[0,1]$, one of which is equal to 1 (the second comes frm the intermediate value theorem). We know that $\mathbb{P}(T_1^< < +\infty) < 1$ by the law of large numbers, so it is the smallest of the two solutions.

- If $S_1 = 0$ almost surely, then $F(x) = x$, so the smallest solution is 0, but we also have $\mathbb{P}(T_1^< < +\infty) = 0$.

- If $\mathbb{E}[S_1] \leq 0$ and $\mathbb{P}(X = 0) < 1$, then F is strictly convex with $F'(1) \leq 1$, so 1 is the only solution to (1.10) (see the right part of Figure 1.3). On the other hand S oscillates or drifts to $-\infty$ so $\mathbb{P}(T_1^< < +\infty) = 1$.

□

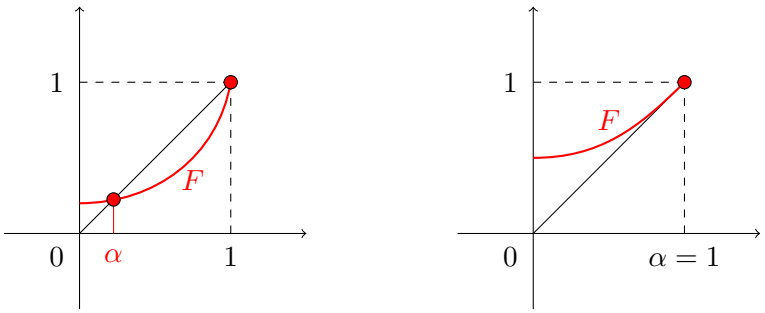

Figure 1.3: Study of the fixed points of F: on the left we have $F'(1) > 1$ and F has two fixed points 1 and α. On the right $F'(1) \leq 1$ and the only fixed point is 1.

Remark 30. The last proof may remind the reader of the computation of the extinction probability of a Galton–Watson process, with the function F playing the same role as the generating function of the offspring distribution. This is not a coincidence: we will see in the next chapter that Galton–Watson trees can be conveniently encoded by negative skip-free random walks, in such a way that extinction of the process corresponds to the random walk hitting a negative value.

1.4 Applications to random trees

Our goal is now to apply the study of random walks to random combinatorial structures. In this chapter, we focus on the simplest such structure: trees. Random trees are widely used in a lot of different contexts: population dynamics in biology, rumor propagation on social networks... They also serve as a building block for the study of more complicated random combinatorial structures such as random graphs.

To study random trees, our main tool will be the encoding of trees by random walks. We will see too such encodings: the contour function and the Lukasiewicz path.

1.4.1 Definitions

We start by defining formally what we mean by a tree. Our definition will be slightly more involved than just defining a tree as a specific kind of graph. The reason for this is that for our walk encodings to behavve nicely, it will be important that the children of each vertex are ordered from "left to right".

We write $\mathcal{U} = \bigcup_{n\geq 0}(\mathbb{N}^*)^n$, with the convention $(\mathbb{N}^*)^0 = \{\emptyset\}$ (this is the set of finite words on the alphabet $\mathbb{N}$). We will write $|u|$ for the length of a word u, and uv for the concatenation of two words u and v. If $v = uj$ with $j \in \mathbb{N}^*$, we call u the *parent* of v and say that v is a *child* of u.

With these "genealogical" conventions, the set $\mathcal{U}$ can be represented by a tree as on Figure 1.4 and is sometimes also called the *Ulam tree*. The trees that we will be interested in will be described as "locally finite" subsets of the Ulam tree.

Definition 31. A (locally finite) *plane tree* is a subset t of $\mathcal{U}$ such that:

- $\emptyset \in t$,

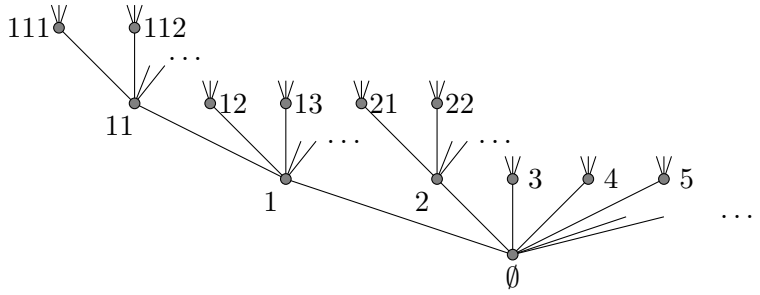

Figure 1.4: The Ulam tree.

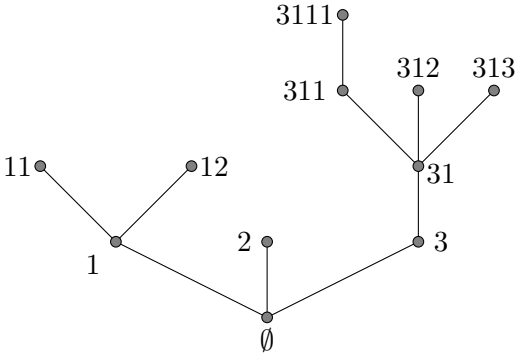

Figure 1.5: An example of plane tree t, with for example $k_{31}(t) = 3$. Note that the image of this plane tree under a reflexion in a vertical mirror gives the plane tree $\{\emptyset, 1, 2, 3, 11, 31, 32, 111, 112, 113, 1131\}$. This is a different plane tree, although the underlying graph is the same.

- if $v \in t \backslash \{\emptyset\}$, the parent of u also belongs to t,
- for all $u \in t$, there is $k_u(t) \in \mathbb{N}$ such that for any $j \in \mathbb{N}$, we have $uj \in t$ if and only if $j \leq k_u(t)$.

We can think of $\emptyset$ as the *root* of t, and of $k_u(t)$ as the number of children of u in the tree t. The name *plane tree* comes from the fact that t can be seen as a graph with an "embedding" in the plane, which orders the children of any vertex from left to right. See Figure 1.5 for an example.

Finally, it will be convenient for us to consider families of trees, which we call *forests*.

Definition 32. A *plane forest* is a finite sequence $(t_i)_{1 \leq i \leq m}$ of plane trees.

1.4.2 The contour function

We will now define the contour function, which is a first way to encode a tree using a walk trajectory. let t be a plane tree with $n+1$ vertices (and therefore n edges). A *corner* of t is a sector around a vertex of t delimited by two consecutive edges. We write $V(c)$ for the vertex incident to the corner c. Let $(c_i)_{0 \leq i \leq 2n}$ be the enumeration of the corners of t in clockwise order, with $c_0 = c_{2n}$ being the corner below the root $\emptyset$. This sequence is called the *contour process*

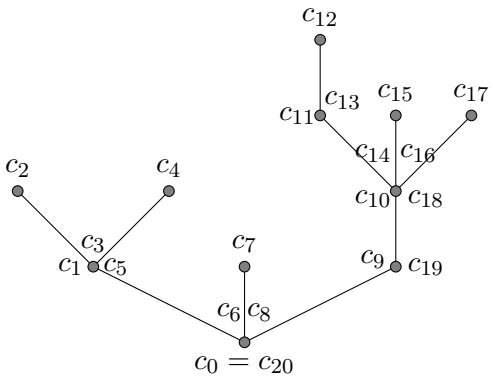

Figure 1.6: The contour process of a finite plane tree with $n = 10$.

of t (see Figure 1.6 for an example). Moreover, the *contour process* of a forest is obtained by concatenating the contour process of all its trees. Note that for each of the trees, the corner below the root appears twice in the contour process. Therefore, if a forest f has m trees and n edges in total, its contour process is indexed by integers from 0 to $2n + m - 1$.

The contour function of a tree simply consists of following the contour process of a tree and recording the height as a function of time. For technical reason (in particular, application of the cyclic lemma later), it will be convenient to extend it by an additional step of -1 in the end. The contour function of the forest will be the concatenation of the contour functions of its trees (including a -1 step in the end of each tree, which makes the formal definition slightly more complicated). See Figures 1.7 and 1.8 for examples.

Definition 33.

- Let t be a tree with n edges, and let $(c_i)_{0\leq i\leq 2n}$ be its contour process. The *contour function* of t is the function $\mathcal{C}(t) : \{0, 1, \ldots, 2n + 1\}$ defined by

$$\mathcal{C}_i(t) = \begin{cases} |V(c_i)| & \text{for } 0 \leq i \leq 2n, \\ -1 & \text{for } i = 2n + 1. \end{cases}$$

- Let $f = (t_1, \ldots, t_m)$ be a forest of m trees with n edges in total, and let $(c_i)_{0\leq i\leq 2n+m-1}$ be its contour process. The *contour function* of f is the function $\mathcal{C}(f) : \{0, 1, \ldots, 2n + m\}$ defined by

$$\mathcal{C}_i(f) = \begin{cases} |V(c_i)| - (j-1) & \text{for } \sum_{k=1}^{j-1}(2|t_k| + 1) \leq i < \sum_{k=1}^{j}(2|t_k| + 1),, \\ -m & \text{for } i = 2n + m. \end{cases}$$

Another interpretation of the contour function is that $+1$ steps correspond to "discovering that a vertex has one more child", and -1 steps correspond to "discovering that a vertex has no more child". In particular, the -1 step in the end of each tree of a forest correspond to learning that the root of that tree has no other child, and therefore that the exploration of that tree is over. This definition was tailored to make it a bijection with a nice set of paths.

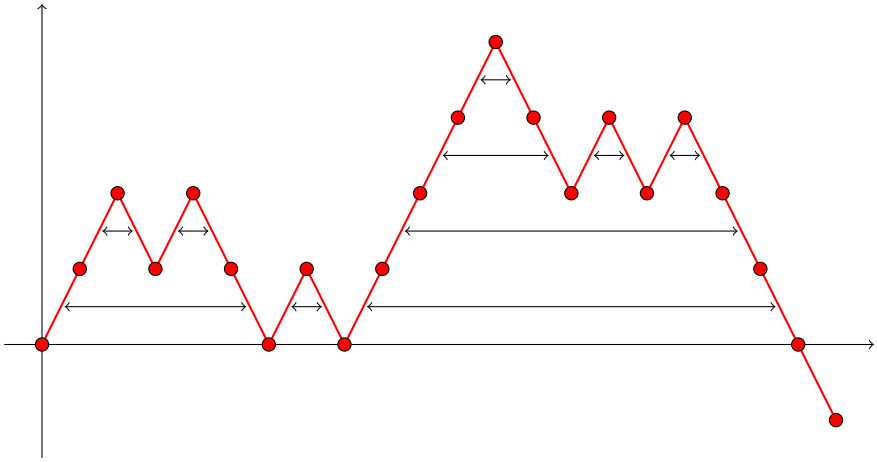

Figure 1.7: The contour function of the tree of Figure 1.5. Arrows indicate the pairs of segments which should get glued together to obtain back the tree.

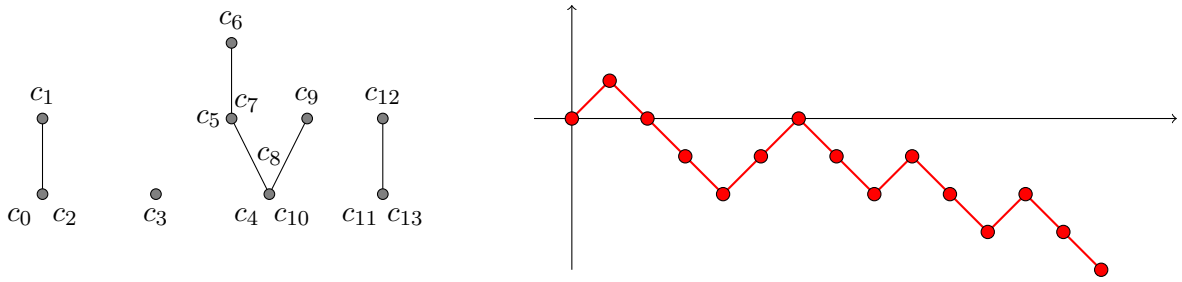

Figure 1.8: On the left, a forest with 4 trees and 5 edges in total, together with its contour process. On the right, its contour function.

Proposition 34. • For all $n \geq 0$, the contour function is a bijection between plane trees with n edges and paths $\mathcal{C} : \{0, 1, \ldots, 2n+1\} \to \mathbb{Z}$ such that $\mathcal{C}_0 = 0$ and $|\mathcal{C}_{i+1} - \mathcal{C}_i| = 1$ for all $0 \leq i \leq 2n$ and

$$\min\{i \geq 0 | \mathcal{C}_i = -1\} = 2n + 1.$$

Such paths are also called *Dyck paths* of length $2n + 1$.

- For all $n \geq 0$ and $m \geq 1$, the contour function is a bijection between plane forests with m trees and n edges in total and paths $\mathcal{C} : \{0, 1, \ldots, 2n+1\} \to \mathbb{Z}$ such that $\mathcal{C}_0 = 0$ and $|\mathcal{C}_{i+1} - \mathcal{C}_i| = 1$ for all $0 \leq i \leq 2n$ and

$$\min\{i \geq 0 | \mathcal{C}_i = -m\} = 2n + m.$$

Proof. We first prove the result for trees. As the height stays nonnegative, it is immediate that the contour function of a tree is a Dyck paths. We now describe the reverse bijection: given a Dyck path, we draw ist graph. We notice that segments of the graph (except the one corresponding to the final -1 step) can be grouped in pair, where two segments of the same pair face each other under the graph. If we glue together the two segments of any such pair, we obtain a tree. We let the reader convince themselves that this is the inverse operation to the contour function.

For forests, note that for $1 \leq j \leq m$, the j-th tree of the forest is encoded by the walk in the time interval $[\sigma_{j-1}, \sigma_j]$, where σ_j is the hitting time of $-j$. Therefore, from the path, we can recover each of the tree and the whole forest. □

Our goal will be to use this bijection to understand probabilistic properties of random trees. Before doing so, we first study combinatorial consequences on the counting of trees and forests.

Corollary 35. • For all $n \geq 0$, the number of plane trees with n edges is

$$\frac{1}{2n+1}\binom{2n+1}{n} = \frac{1}{n+1}\binom{2n}{n}. \tag{1.11}$$

- For all $n \geq 0$ and $m \geq 1$, the number of plane forests with m trees and n edges in total is $\frac{1}{2n+m}\binom{2n+m}{n}$.

The sequence (1.11) appears a lot in combinatorics and is known as *Catalan numbers*.

Proof. • We start with the proof for trees. By Proposition 34, it is sufficient to count Dyck paths of length $2n + 1$. In order to have $\mathcal{C}_{2n+1} = -1$, we need exactly $n + 1$ steps $+1$ and n steps -1, so there are $\binom{2n+1}{n}$ paths $\mathcal{C}$ with $\mathcal{C}_{2n+1} = -1$. Moreover, let us say that two such paths are equivalent if they are cyclic shifts of each other. Then each equivalence class contains $2n+1$ paths, and the cyclic lemma shows that each equivalence class contains exactly one Dyck paths, so Dyck paths represent a proportion $\frac{1}{2n+1}$ of paths $\mathcal{C}$ with $\mathcal{C}_{2n+1} = -1$, which proves the formula.

- The proof for forests is the same, and relies on the cyclic lemma in the same way. The only difference is that it is now possible to obtain the same path after a nontrivial cyclic shift. As a consequence, some equivalence classes may have a size which is not $2n + m$, but a divisor of $2n + m$. However, the proportion of Dyck paths in each equivalence class will always be $\frac{m}{2n+m}$.

□

If we are more interested in probability than in combinatorics, the following corollary of Proposition 34 is also important.

Corollary 36. Let T_n be a uniform plane tree with n edges. then its contour function has the law of a simple random walk S on $\mathbb{Z}$, conditioned on

$$\min\{i \geq 0 | S_i = -1\} = 2n + 1.$$

Scaling limits of random trees? This is very useful because many natural quantities for the tree can be read directly on the contour function. For example, there is a direct correspondence between the height of a tree and the maximum of its contour function. It is even possible to go even further: similarly to the convergence (after proper scaling) of a simple random walk to a Brownian motion, it can be proved that the conditioned random walk converges to a variant of Brownian motion called the *Brownian excursion*. This can be seen as a Brownian motion B conditioned on $B_1 = 0$ and on staying positive on the interval $(0, 1)$. We can then apply to the Brownian excursion a continuous version of the gluing operation of the proof of Proposition 34. This allows to prove that uniform random trees "converge" to a random compact metric space called the *Brownian tree*. Giving a proper meaning to all this is beyond the scope of this course, but we refer to [2] (in French) or [4] for more about random trees and their scaling limits.

Typical height in a random tree. Our next goal will be much more modest: we will study the height of a typical vertex (i.e. a vertex picked uniformly at random) in a random tree. Since a random walk of length n is typically of order $\sqrt{n}$, it should not be surprising that the height is also of order $\sqrt{n}$. However, our result will be much more precise than that: we compute exactly the distribution of this typical height.

Theorem 12. We fix $n \geq 0$. Let T_n be a uniform plane tree with n edges. Conditionally on T_n; let v_n be a uniform vertex of T_n, and let H_n be the height of v_n in T_n. Then for all $h \geq 0$, we have

$$\mathbb{P}(H_n = h) = \frac{2h+1}{2n+1} \frac{\binom{2n+1}{n-h}}{\binom{2n}{n}}. \tag{1.12}$$

In particular, we have the convergence in distribution

$$\frac{H_n}{\sqrt{n}} \xrightarrow[n\to+\infty]{} R,$$

where R is a random variable with density $2xe^{-x^2}\mathbb{1}_{x>0}$ with respect to the Lebesgue measure (this is the *Rayleigh* distribution).

Proof. There are usually two natural approaches to study random combinatorial objects when we are able to count them. The first one is to use combinatorial results as black boxes, and express the probability of events we are interested in as quotients of combinatorial quantities. The second one is to use the proof of those results, when it is explicit enough (here, it would mean using the relation with random walks). In this case, we will rely on the first approach.

The pair (T_n, v_n) is uniform among all pairs (t, v) consisting of a plane tree with n edges, and one of its vertices. The number of such pairs is equal to

$$(n+1) \times \frac{1}{n+1}\binom{2n}{n} = \binom{2n}{n}.$$

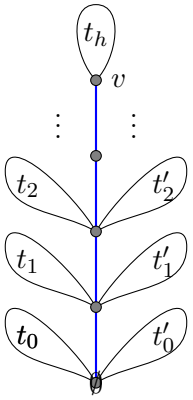

Figure 1.9: A tree with a vertex v at height h. The tree can be decomposed into a spine with length h (in blue), and $2h+1$ trees (h grafted on the left of the spine, h on the right and one on top of v).

On the other hand, let us count pairs where v is at height h. For any such pair, we can draw t as a spine of length h from the root to v, and $2h+1$ trees $t_0, t'_0, \dots, t_{h-1}, t'_{h-1}, t_h$ grafted on this spine as on Figure **??**. This construction is a bijection with forests of $2h+1$ trees with $n-h$ edges in total (n edges in t, minus the h edges in the spine). By Corollary 35, the number of such forests is

$$\frac{2h+1}{2(n-h)+(2h+1)}\binom{2(n-h)+(2h+1)}{n-h} = \frac{2h+1}{2n+1}\binom{2n+1}{n-h}.$$

which proves (1.12).

Finally, if we take $h \sim x\sqrt{n}$ with $x > 0$ and apply the Stirling formula, we find

$$\mathbb{P}(H_n = h) \underset{n\to+\infty}{\sim} \frac{2xe^{-x^2}}{\sqrt{n}},$$

which proves the convergence in distribution to a Rayleigh variable. □

The contour function is a very powerful to understand uniform tree with given size. However, if we try to study different models (for example trees with constraints on the number of children of vertices), it can quickly introduce complex dependences in the contour function and make things untractable. We will now introduce a different encoding of trees by random walks that is useful for a wider class of models.

1.4.3 The Lukasiewicz path

We first define the lexicographical order on a tree. If $u = u_1 \dots u_k$ and $v = v_1 \dots v_k$ are at the same height in a tree, let i_0 be the first index i such that $u_i \neq v_i$. We write $u < v$ if $u_{i_0} < v_{i_0}$. We can now define the *breadth-first order* on a tree: we write $u \preceq v$ if $|u| < |v|$ or if $|u| = |v|$ and $u < v$ for the lexicographical order (recall that $|u|$ is the length of the word u, or its height in the tree). See the left part of Figure 1.10 for an example. We also recall that $k_u(t)$ stands for the number of children of a vertex u in a tree t.

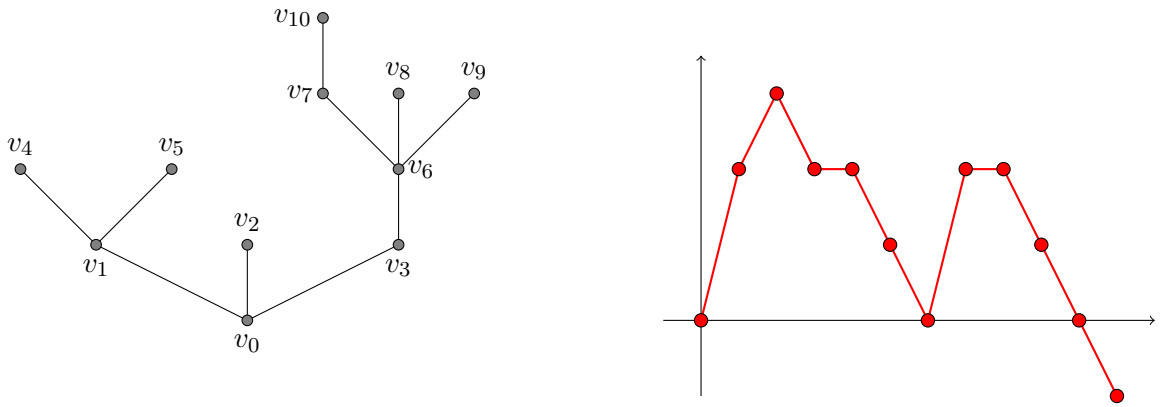

Figure 1.10: On the left, a tree with its vertices listed in the breadth-first order. On the right, its Lukasiewicz path.

Definition 37. Let t be a (finite or infinite) plane tree with $n+1$ vertices (n may be infinite), and let $(v_i)_{0\leq i\leq n}$ be its vertices listed in breadth-first order. The *Lukasiewicz path* $(\mathcal{L}_i(t))_{0\leq i\leq n+1}$ associated to t is defined by induction by

$$\mathcal{L}_i(t) = \begin{cases} \mathcal{L}_0(t) = 0, \\ \mathcal{L}_{i+1}(t) = \mathcal{L}_i(t) + k_{v_i}(t) - 1 \quad \text{for } 0 \leq i \leq n. \end{cases}$$

Here is a more intuitive interpretation of the Lukasiewicz path: we reveal in the breadth-first order how many children each vertex has. When the number of children of a vertex is revealed, we consider these children as "explored". We call a vertex "active" if it has been explored, but its children have not yet been revealed. Then $\mathcal{L}_i(t)$ is the number of active vertices after i steps, minus 1. Indeed, the i-th step removes one vertex v_i from active vertices, and adds $k_{v_i}(t)$ new active vertices. More formally, we can write

$$\mathcal{L}_i(t) = \#\{j > i | \text{the parent of } v_j \text{ is } v_{j'} \text{ for some } j' \leq i\} - 1.$$

In particular, it is clear that $\mathcal{L}_i(t) \geq -1$ with equality if and only if there is no active vertex left, which means that the exploration of the tree is over. Therefore:

- if t is infinite, then $\mathcal{L}_i(t) \geq 0$ for all $i \geq 0$,
- if t has n vertices, then $\mathcal{L}_{n+1} = -1$ and $\mathcal{L}_i(t) \geq 0$ for $0 \leq i \leq n$.

Note also that $\mathcal{L}(t)$ is always a negative skip-free path, and that the construction of the Lukasiewicz path is reversible. It follows that $\mathcal{L}$ is a bijection between plane trees and negative skip-free paths on $\mathbb{Z}$ started from 0 and killed when they hit -1. As for the contour function, this has interesting combinatorial consequences.

Proposition 38. Let $d_0, d_1, \dots \in \mathbb{N}$ such that $\sum_i d_i = 1 + \sum_i i d_i = n + 1$. Then the number of plane trees where for all i, exactly d_i vertices have i children is given by

$$\frac{n!}{d_0! d_1! d_2! \dots}.$$

Note that the condition $\sum_i d_i = 1 + \sum_i i d_i$ is necessarily for such trees to exist (it expresses that each vertex has one parent, except for the root).

Proof. By the Lukasiewicz path bijection, it is sufficient to count paths of length $n+1$ with d_i steps $i-1$ for all $i \geq 0$ (such paths must end at -1 by our condition on the d_i), and hit -1 first at time $n+1$. Without the condition to stay nonnegative, these paths are just counted by the multinomial coefficient

$$\frac{(n+1)!}{d_0!d_1!d_2!\dots}.$$

By the cyclic lemma, the proportion of these paths which hit -1 first at time $n+1$ is $\frac{1}{n+1}$, which concludes the proof. □

1.4.4 Galton–Watson trees

We now introduce a class of models of random trees which is much more general than "uniform trees with a fixed number of edges". Roughly speaking, Galton–Watson trees are trees where the number of children of vertices are given by i.i.d. variables. These are very natural models for population dynamics under "ideal" assumptions (for example, neglecting that individuals compete for resources).

More precisely, let us fix a probability measure μ on the set $\mathbb{N}$ of nonnegative integers. We remind that the Ulam tree is $\mathcal{U} = \bigcup_{n\geq 0}(\mathbb{N}^*)^n$.

Definition 39. Let $(K_u)_{u\in\mathcal{U}}$ be i.i.d. with law μ, and let

$$T_\mu = \{u = u_1 \dots u_k \in \mathcal{U} \,|\, \forall 1 \leq i \leq k,\, u_i \leq K_{u_1\dots u_{i-1}}\}.$$

A *Galton–Watson tree with offspring distribution* μ is a random plane tree with the law of T_μ.

This means that if t is finite,then

$$\mathbb{P}\left(T_\mu = t\right) = \prod_{v\in t} \mu\left(k_v(t)\right).$$

However, because of the possibility of infinite trees, this formula is not sufficient to describe completely the law of T_μ.

The reason why the Lukasiewicz walk is so useful is the following result.

Proposition 40. Let T be a Galton–Watson tree with offspring distribution μ, and let (S_n) be a random walk whose increment distribution is given by $\mathbb{P}\left(S_1 = k\right) = \mu(k+1)$ for all $k \geq -1$. Let $T_1^< \in \mathbb{N}^* \cup \{+\infty\}$ be the first time where S hits -1. Then

$$\left(\mathcal{L}_0(T), \mathcal{L}_1(T), \dots, \mathcal{L}_{|T|}(T)\right) \overset{(d)}{=} \left(S_0, S_1, \dots, S_{T_1^<}\right). \tag{1.13}$$

Proof. Let us fix $\ell_0 = 0, \ell_1, \dots, \ell_n$ such that for all $0 \leq i \leq n-1$, we have $\ell_i \geq 0$ and $\ell_{i+1} - \ell_i \geq -1$ (in other words ℓ is the possible beginning of a trajectory for a Lukasiewicz path). Let also $v_0, \dots, v_{n-1}$ be the n first vertices of T in breadth-first order. Then we have

$$(\mathcal{L}_0(T), \mathcal{L}_1(T), \dots, \mathcal{L}_n(T)) = (\ell_0, \ell_1, \dots, \ell_n)$$

if and only if, for all $0 \leq i \leq n-1$, the vertex v_i has exactly $\ell_{i+1} - \ell_i + 1$ children. This occurs with probability

$$\mathbb{P}\left(\forall 0 \leq i \leq n-1,\, K_{v_i} = \ell_{i+1} - \ell_i + 1\right) = \prod_{i=0}^{n-1} \mu\left(\ell_{i+1} - \ell_i + 1\right) = \mathbb{P}\left((S_0, \ldots, S_n) = (\ell_1, \ldots, \ell_n)\right).$$

This proves that the first n steps of the left and right-hand sides of (1.13) have the same distribution for all n, which proves the result. □

Extinction of a Galton–Watson tree. One of the most natural questions that can be asked about a random tree is if it is finite or not. We now compute the probability for our tree to survive and be infinite.

Theorem 13. Let μ be a probability measure on $\mathbb{N}$ with mean m, and let T be a Galton–Watson tree with offspring distribution μ. Then $\mathbb{P}(T \text{ is finite})$ is the smallest solution α in $[0,1]$ to the equation

$$\alpha = \sum_{k \geq 0} \mu(k) \alpha^k.$$

In particular, the tree T is almost surely finite if and only if $m \leq 1$ and $\mu \neq \delta_1$.

We give two proofs of this result.

First proof. By Proposition 40, we have

$$\mathbb{P}(T \text{ is finite}) = \mathbb{P}\left(T_1^< < +\infty\right),$$

where $T_1^<$ is the hitting time of -1 for a negative skip-free random walk. We have shown in the previous chapter that this probability is α. □

Second proof. We write Z_n for the number of vertices at height n in T, and define the *generating function* g of μ as

$$g(x) = \sum_{k \geq 0} \mu(k) x^k = \mathbb{E}[x^{Z_1}].$$

Then the law of the process Z can be described by $Z_0 = 1$ and

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{n,i}$$

for $n \geq 0$, where the variables $\xi_{n,i}$ for $n \geq 0, i \geq 1$ are i.i.d. with law μ. We can then write

$$\mathbb{E}\left[x^{Z_{n+1}}\right] = \mathbb{E}\left[\mathbb{E}\left[x^{\sum_{i=1}^{Z_n} \xi_{n,i}} | Z_n\right]\right] = \mathbb{E}\left[\prod_{i=1}^{Z_n} \mathbb{E}\left[x^{\xi_{n,i}}\right]\right] = \mathbb{E}\left[g(x)^{Z_n}\right].$$

By induction on n, it follows that

$$\mathbb{E}\left[x^{Z_n}\right] = g \circ g \circ \cdots \circ g(x),$$

where g appears n times. In particular, we have

$$\mathbb{P}(Z_n = 0) = \mathbb{E}\left[0^{Z_n}\right] = g \circ \cdots \circ g(0).$$

From here, using that g is nondecreasing, we easily obtain that this sequence converges to the smallest fixed point of g. □

Size of a Galton–Watson tree. If we want to be a bit more precise, we can study the law of the total size of a Galton–Watson tree.

Proposition 41. Assume that $\sum_i i\mu(i) = 1$ and $\mathrm{Var}(\mu) = \sigma^2 \in (0, +\infty)$. If T is a Galton–Watson tree with offspring distribution μ, then

$$\mathbb{P}(|T| = n) \underset{n\to+\infty}{\sim} \frac{1}{\sqrt{2\pi\sigma^2}} n^{-3/2}.$$

Proof. Using the Lukasiewicz walk, we can write

$$\begin{aligned}\mathbb{P}(|T| = n) &= \mathbb{P}(\mathcal{L}(T) \text{ hits } -1 \text{ first at time } n)\\ &= \frac{1}{n}\mathbb{P}(\mathcal{L}_n(T) = -1)\\ &= \frac{1}{n}\left(\frac{1}{\sqrt{2\pi\sigma^2}} n^{-1/2} + o\left(n^{-1/2}\right)\right),\end{aligned}$$

where the second line uses the cyclic lemma and the third line uses the local CLT. □

We conclude this section with a few informal comments on the power of the Lukasiewicz path. It may seem at first sight that it easily encodes the size of the tree, but is not as good as the contour function at encoding finer geometric properties of the tree (for example, the maximal height of the tree is not a direct, simple function of the Lukasiewicz path). However, the power of this tool is its flexibility. We have chosen to present it when the vertices are explored in the breadth-first order, but most of what we have done still applies for any order in which we order the vertices[2]. In particular, if the tree is finite and we choose the *depth-first order* (which corresponds to the "true" alphabetical order on the Ulam tree), it can be shown that the Lukasiewicz path becomes a good approximation to the contour function. Therefore, it can be used (with some additional work) to recover information about heights in random trees in a much more general context than what we did with the contour function. We refer to [4] for more details.

1.5 Range and intersections of random walks

In this section, we will only work with random walks on $\mathbb{Z}^d$.

1.5.1 Linear growth of the range

Let S be a random walk on $\mathbb{Z}^d$. We will be interested in the number of distinct points visited by S during its first n steps.

Definition 42. For $n \geq 0$, we write R_n for the cardinal of the set

$$\{S_0, S_1, \ldots, S_n\}.$$

We call R_n the *range* of S until time n.

[2]The only advantage of the breadth-first order is that it ensures that we discover all the vertices, even if the tree is infinite.

It is clear that we always have $1 \leq R_n \leq n+1$. The first question that we will study is whether the upper bound has the right order of magnitude.

Theorem 14 (Kesten–Spitzer)**.** We have the convergence in probability

$$\frac{R_n}{n} \xrightarrow[n\to+\infty]{(P)} \mathbb{P}\left(\forall n > 0,\, S_n \neq 0\right).$$

In particular, if S is transient, the constant appearing in the right-hand side is positive, and the range grows linearly in n. On the other hand, if the walk is recurrent, the range is $o(n)$.

Proof. The proof simly consists of performing first and second moment computations on the random variable S_n. As usual (think of the proof of the weak law of large numbers for L^2 variables), such computations are convenient to perform when the variable R_n is decomposed as a sum of "small" variables. For this, we will write

$$R_n = \sum_{i=0}^{n} \mathbb{1}_{\forall 0 \leq k < i,\, S_k \neq S_i}.$$

That is, the number of points visited by S is equal to the number of times where S visits a new point for the first time. To simplify notation, we also write

$$\alpha := \mathbb{P}\left(\forall n > 0,\, S_n \neq 0\right).$$

We can now perform the first moment computation:

$$\mathbb{E}\left[R_n\right] = \mathbb{E}\left[\sum_{i=0}^{n} \mathbb{1}_{\forall 0 \leq k < i,\, S_k \neq S_i}\right] = \sum_{i=0}^{n} \mathbb{P}\left(\forall 0 \leq k < i,\, S_k \neq S_i\right) = \sum_{i=0}^{n} \mathbb{P}\left(\forall 0 < k \leq i,\, S_k \neq 0\right), \tag{1.14}$$

where the last equality comes from duality. The terms in the right-hand side converge to α as i goes to $+\infty$. Therefore, by summing, we get $\mathbb{E}[R_n] = (\alpha + o(1))n$ as $n \to +\infty$. If $\alpha = 0$, this implies the result immediately by the Markov inequality.

If $\alpha > 0$, we need a second moment computation to establish concentration. To lighten notation we write, for all $i \geq 0$,

$$\eta_i = \mathbb{1}_{\forall 0 \leq k < i,\, S_k \neq S_i}.$$

We then have

$$\mathbb{E}\left[R_n^2\right] = \sum_{0 \leq i,j \leq n} \mathbb{E}\left[\eta_i \eta_j\right] = O(n) + 2 \sum_{0 \leq i < j \leq n} \mathbb{E}[\eta_i \eta_j]. \tag{1.15}$$

Let us now fix $i < j$ and study the corresponding term. Note that we are interested in an *upper* bound on the variance. Using duality, we obtain

$$\begin{aligned}
\mathbb{E}\left[\eta_i \eta_j\right] &= \mathbb{P}\left(\forall 0 \leq k < j,\, S_k \neq S_j \text{ and } \forall 0 \leq k < i,\, S_k \neq S_i\right) \\
&= \mathbb{P}\left(\forall 0 < k \leq j,\, S_k \neq 0 \text{ and } \forall j - i < k \leq j,\, S_k \neq S_{j-i}\right) \\
&\leq \mathbb{P}\left(\forall 0 < k \leq j - i,\, S_k \neq 0 \text{ and } \forall j - i < k \leq j,\, S_k \neq S_{j-i}\right) \\
&= \mathbb{P}\left(\forall 0 < k \leq j - i,\, S_k \neq 0\right) \mathbb{P}\left(\forall 0 < k \leq i,\, S_k \neq S_0\right),
\end{aligned}$$

where the last equality comes from the weak Markov property at time $j - i$. Note that the tricky part here was to write down the right inequality to make the two events "independent". Now

each of these two factors go to α as respectively $j-i$ and i go to infinity. In other words, let us fix $\varepsilon > 0$. Then there is A such that if both i and $j-i$ are at least than A, then

$$\mathbb{E}\left[\eta_i \eta_j\right] \leq \alpha^2 + \varepsilon.$$

By bounding by 1 the terms for which $\min(i, j-i) < A$ and summing over i, j, we find

$$\mathbb{E}[R_n^2] \leq (\alpha^2+\varepsilon)n^2 + O(n) + 2 \sum_{\substack{1 \leq i < j \leq n \\ i < A}} 1 + 2 \sum_{\substack{1 \leq i < j \leq n \\ j-i < A}} 1.$$

In the last two sums, the number of terms is $O(n)$, so we have proved

$$\mathbb{E}[R_n^2] \leq (\alpha^2+\varepsilon)n^2 + O(n).$$

Since this is true for any $\varepsilon > 0$, we have shown

$$\mathrm{Var}(R_n) = \mathbb{E}[R_n^2] - \mathbb{E}[R_n]^2 \leq (\alpha^2 + o(1))n^2 - (\alpha^2 + o(1))n^2 = o(n^2).$$

Finally, the Bienaymé–Chebychev inequality shows that for any $\varepsilon > 0$, we have for n large enough

$$\mathbb{P}\left(\left|\frac{R_n}{n} - \alpha\right| \geq \varepsilon\right) \leq \mathbb{P}\left(|R_n - \mathbb{E}[R_n]| \geq \frac{\varepsilon}{2} n\right) \leq \frac{\mathrm{Var}(R_n)}{(\varepsilon/2)^2 n^2} \xrightarrow[n \to +\infty]{} 0.$$

□

Remarks 43.

- Using tools from ergodic theory, it can be proved that the convergence in probability is actually almost sure.
- In the particular case of the simple random walk on $\mathbb{Z}^d$, the result shows that the range is linear for $d \geq 3$ and $o(n)$ in dimension 1 or 2. For $d = 1$, it is actually easy to get something much more precise than the bound $R_n = o(n)$. Indeed, the set of values visited by the walk is exactly the interval between the maximal and the minimal value, so

  $$R_n = 1 + \max_{0 \leq i \leq n} S_i - \min_{0 \leq i \leq n} S_i.$$

  In particular, it is of order $\sqrt{n}$. We can even deduce from the Donsker theorem that

  $$\frac{R_n}{\sqrt{n}} \xrightarrow[n \to +\infty]{(d)} \sup_{[0,1]} B - \inf_{[0,1]} B,$$

  where B is a Brownian motion.

1.5.2 The range in dimension 2

From the previous subsection, the only dimension where we do not understand well the order of magnitude of the range of the simple random walk is $d = 2$. Therefore, in this subsection S_n will stand for the simple random walk in 2 dimensions.

Theorem 15 (Dvoretzky–Erdös)**.** For the simple random walk on $\mathbb{Z}^2$, we have

$$\frac{\log n}{n} R_n \xrightarrow[n \to +\infty]{(P)} \pi.$$

The large-scale strategy of the proof will be the same as before. In the previous proof, we decomposed the expectation of R_n as a sum of small variables η_i and argued using duality that $\mathbb{E}[\eta_i]$ was roughly the same as $\mathbb{P}(\tau_0 > i)$, where τ_0 is the first return time to 0. We will do the same here, but we will need an estimate more accurate than $\mathbb{P}(\tau_0 > i) \to 0$ as $i \to +\infty$.

Lemma 44. We have $\mathbb{P}(\tau_0 > n) \sim \frac{\pi}{\log n}$ as $n \to +\infty$.

Sketch of proof. We will first compute something much easier: we will show that for all $n \geq 0$, we have

$$\mathbb{P}(S_{2n} = 0) = \frac{1}{4^{2n}} \binom{2n}{n}^2.$$

Note that S_{2n+1} is impossible because the parity of the sum of coordinates changes at every step. For this, we will use a trick specific to dimension 2 consisting of rotating the square lattice by 45 degrees. More precisely, let S'_n be the random walk on $\mathbb{Z}^2$ whose step distribution is the uniform distribution on $\{(-1,-1),(-1,1),(1,-1),(1,1)\}$. Then S'_n is the image of S_n under the linear automorphism of R^2 corresponding the the matrix $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. On the other hand, the two coordinates of S' are independent simple random walks on $\mathbb{Z}$, so

$$\mathbb{P}(S_{2n} = 0) = \mathbb{P}\left(S'_{2n} = 0\right) = \mathbb{P}\left(S_{2n}^{\dim=1}\right)^2 = \left(\frac{1}{2^{2n}}\binom{2n}{n}\right)^2.$$

By the Stirling formula, we deduce

$$\mathbb{P}(S_{2n} = 0) \underset{n\to+\infty}{\sim} \frac{1}{\pi n}. \tag{1.16}$$

In dimension 1, once this is computed, we would deduce the computation of $\mathbb{P}(\tau_0 = n)$ using the cyclic lemma. Unfortunately, no variant of the cyclic lemma is available in dimension 2, so we will need a completely different method. We will rely on a more analytic argument using generating functions. More precisely, we define the generating functions

$$F(x) = \sum_{n\geq 0} x^n \mathbb{P}(S_{2n} = 0), G(x) \qquad = \sum_{n\geq 1} x^n \mathbb{P}(\tau_0 = 2n), H(x) = \sum_{n\geq 0} x^n \mathbb{P}(\tau_0 > 2n).$$

The idea of the proof will be to establish relations between these generating functions. More precisely, we will want to use the fact that the asymptotics of the coefficients of a generating function are linked to the behaviour of that function in the neighbourhood of its radius of convergence. Using (1.16) the radius of convergence of F is 1 and elementary analysis shows

$$F(x) \underset{x\to 1^-}{\sim} \frac{1}{\pi} \log \frac{1}{1-x}.$$

We will now find a link between F and G. It is more convenient here to think of the question as "how can we express f in terms of G?", even if what we want to do in the end is the other way around. For $k \geq 0$, let $\tau_0^{(k)}$ be the k-th times where S hits 0, with $\tau_0^{(0)} = 0$. Then for any

$|x| < 1$, writing n_i for the increments between the times $\tau_0^{(i)}$, we have

$$\begin{aligned}
F(x) &= \sum_{n\geq 0}\sum_{k\geq 0} x^n \mathbb{P}\left(\tau_0^{(k)} = 2n\right) \\
&= \sum_{k\geq 0}\sum_{n_1,\dots,n_k\geq 1} x^{n_1+\cdots+n_k}\mathbb{P}\left(\tau_0^{(1)} = 2n_1, \tau_0^{(2)} = 2n_1+2n_2, \dots, \tau_0^{(k)} = 2n_1+2n_2+\cdots+2n_k\right) \\
&= \sum_{k\geq 0}\sum_{n_1,\dots,n_k\geq 1} x^{n_1+\cdots+n_k}\mathbb{P}\left(\tau_0 = 2n_1\right)\times\cdots\times\mathbb{P}\left(\tau_0 = 2n_k\right) \\
&= \sum_{k\geq 0} G(x)^k \\
&= \frac{1}{1-G(x)},
\end{aligned}$$

where in the end we applied the strong Markov property to the times $\tau_0^{(k)}$. We now find a link between G and H, which is simpler. For $|x| < 1$, we have

$$\begin{aligned}
G(x) &= \sum_{n\geq 1} x^n \left(\mathbb{P}\left(\tau_0 > 2n-2\right) - \mathbb{P}\left(\tau > 2n\right)\right) \\
&= \sum_{n\geq 0} x^{n+1}\mathbb{P}\left(\tau_0 > 2n\right) - \sum_{n\geq 1} x^n \mathbb{P}\left(\tau_0 > 2n\right) \\
&= xH(x) - (H(x)-1).
\end{aligned}$$

By combining the last two equations that we found, we obtain, for $|x| < 1$:

$$H(x) = \frac{1-G(x)}{1-x} = \frac{1}{(1-x)F(x)} \underset{x\to 1^-}{\sim} \frac{\pi}{(1-x)\log\frac{1}{1-x}}. \tag{1.17}$$

The conclusion now comes from a Tauberian theorem: by the Hardy–Littlewood–Karamata Tauberian (see for example [1, Theorem VI.13], with $\alpha = 1$ and $\Lambda(y) = \frac{1}{\log y}$), this implies

$$\sum_{i=0}^{n} \mathbb{P}\left(\tau_0 > 2i\right) \underset{n\to+\infty}{\sim} \pi\frac{n}{\log n}.$$

In general, it is not possible to deduce a better estimate using only (1.17). However, we also know that $\mathbb{P}\left(\tau_0 > 2n\right)$ is decreasing in n, which will allow us to conclude. More precisely, fix $\varepsilon > 0$. By using monotonicity followed by the last display at times $(1-\varepsilon)n$ and n, we find

$$\varepsilon n\mathbb{P}\left(\tau_0 > 2n\right) \leq \sum_{i=(1-\varepsilon)n}^{n} \mathbb{P}\left(\tau_0 > 2i\right) \leq (1+\varepsilon)\pi\varepsilon\frac{n}{\log n}$$

for n large enough, so $\mathbb{P}\left(\tau_0 > 2n\right) \leq \frac{(1+\varepsilon)\pi}{\log n}$. The same argument in the time interval $[n, (1+\varepsilon)n]$ shows $\mathbb{P}\left(\tau_0 > 2n\right) \geq \frac{(1-\varepsilon)\pi}{\log n}$ for n large enough, which concludes the proof. □

Proof of the Erdös–Dvoretzky Theorem. The proof is now very similar to the proof of the Kesten–Spitzer Theorem. For the same reason as in (1.14), we can write

$$\mathbb{E}[R_n] = \sum_{i=0}^{n} \mathbb{P}\left(\tau_0 > i\right) \underset{n\to+\infty}{\sim} \sum_{i=2}^{n} \frac{\pi}{\log i} \underset{n\to+\infty}{\sim} \pi\frac{n}{\log n}.$$

Let us now bound the second moment. As in (1.15), we have

$$\mathbb{E}\left[R_n^2\right] \leq O(n) + 2 \sum_{0 \leq i < j \leq n} \mathbb{P}\left(\tau_0 > i\right) \mathbb{P}\left(\tau_0 > j - i\right).$$

Now fix $\varepsilon > 0$. Then the number of terms where $i \leq n^{1-\varepsilon}$ or $j - i \leq n^{1-\varepsilon}$ is $O\left(n^{2-\varepsilon}\right)$, so their total contribution is $O\left(n^{2-\varepsilon}\right)$. On the other hand, for all the other terms, if n is large enough, we have

$$\mathbb{P}\left(\tau_0 > i\right) \leq (1+\varepsilon)\frac{\pi}{\log(n^{1-\varepsilon})} = \frac{1+\varepsilon}{1-\varepsilon}\frac{\pi}{\log n},$$

and similarly for $\mathbb{P}\left(\tau_0 > j - i\right)$. We finally obtain

$$\mathbb{E}[R_n^2] \leq O\left(n^{2-\varepsilon}\right) + \left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 \left(\frac{\pi}{\log n}\right)^2 n^2.$$

Since this is true for any $\varepsilon > 0$, this shows $\operatorname{var}(R_n) = o\left(\mathbb{E}[R_n]^2\right)$. Together with the Bienaymé–Chebychev inequality, this concludes the proof. □

1.5.3 Intersections of random walks

In this subsection, we will try to answer the following equation: do the trajectories of two independent random walks intersect a lot? More precisely, we fix two parameters d (dimension) and k (number of random walks). Let μ be a probability distribution on $\mathbb{Z}^d$. Let $S^{(1)}, \dots, S^{(k)}$ be k independent random walks on $\mathbb{Z}^d$ with step distribution μ. Finally, let $\mathcal{I}$ be the number of points visited by $S^{(i)}$ for all $1 \leq i \leq k$. We want to understand if $\mathcal{I}$ is finite or not. We start with a 0-1 law.

Lemma 45. For any $d, k \geq 1$ and μ as above, we have

$$\mathbb{P}\left(\mathcal{I} = +\infty\right) \in \{0, 1\}.$$

Proof. The main tool of the proof is the Hewitt–Savage 0-1 law. Indeed, the event $\{\mathcal{I} = +\infty\}$ is not invariant under changing the value of one increment of one of the random walks, but it is invariant under permutation of the increments in one fixed walk. Let $S^{(1)}, \dots, S^{(k)}, \widetilde{S}^{(1)}, \dots, \widetilde{S}^{(k)}$ be $2k$ independent random walks with step distribution μ. For any set $A \subset \mathbb{Z}^d$, the Hewitt–Savage 0-1 law implies that

$$\mathbb{P}\left(\#S^{(1)} \cap A = +\infty\right) \in \{0, 1\}.$$

Therefore, almost surely, we have

$$\mathbb{1}_{\#S^{(1)} \cap A = +\infty} = \mathbb{P}\left(\#S^{(1)} \cap A = +\infty\right) = \mathbb{1}_{\#\widetilde{S}^{(1)} \cap A = +\infty}.$$

We now apply this to the (random) set A of points visited by the $k-1$ walks $S^{(2)}, \dots, S^{(k)}$. We obtain that almost surely

$$\mathbb{1}_{\mathcal{I}\left(S^{(1)}, S^{(2)}, \dots, S^{(k)}\right) = +\infty} = \mathbb{1}_{\mathcal{I}\left(\widetilde{S}^{(1)}, S^{(2)}, \dots, S^{(k)}\right) = +\infty}.$$

By applying the same reasoning repeatedly, we can replace $S^{(i)}$ by $\widetilde{S}^{(i)}$ for all i to obtain almost surely

$$\mathbb{1}_{\mathcal{I}\left(S^{(1)}, S^{(2)}, \dots, S^{(k)}\right) = +\infty} = \mathbb{1}_{\mathcal{I}\left(\widetilde{S}^{(1)}, \widetilde{S}^{(2)}, \dots, \widetilde{S}^{(k)}\right) = +\infty}.$$

Therefore, we have proved that two i.i.d. copies of $\mathbb{1}_{\mathcal{I}\left(S^{(1)}, S^{(2)}, \dots, S^{(k)}\right) = +\infty}$ are almost surely equal, so the corresponding event has probability 0 or 1. □

We will now give a characterization of when this intersection is infinite for "nice" enough step distributions μ.

Theorem 16 (Erdös–Taylor)**.** We assume that μ is centered, aperiodic on $\mathbb{Z}^d$ and has bounded support.

- If $d \geq 5$ and $k \geq 2$, then $\mathcal{I} < +\infty$ a.s..
- If $d = 4$, then we have $\mathcal{I} = +\infty$ a.s. for $k = 2$ but $\mathcal{I} < +\infty$ a.s. for $k \geq 3$.
- If $d = 3$, then we have $\mathcal{I} = +\infty$ a.s. for $k \leq 3$ but $\mathcal{I} < +\infty$ a.s. for $k \geq 4$.
- If $d \in \{1, 2\}$, then $\mathcal{I} = +\infty$ a.s. for any $k \geq 1$.

Note that the result for $d \in \{1, 2\}$ is immediate since then the walk is recurrent, so all points are visited by all random walks. Note also that the result only depends on d and k, but not on μ (as long as μ statisfies the assumptions).

A way to see these last results "intuitively" is to think of the trajectory of a random walk as a 2-dimensional object. This can be justified by the fact that a random walk in $d \geq 3$ will typically be at distance $\sqrt{n}$ from 0 at time n, so it will spend a time roughly r^2 in the ball of radius r, so the intersection of its range with the ball should be of order r^2, just like for a 2-dimensional subspace. We can then compare these properties with intersection properties of generic 2-dimensional subspaces of $\mathbb{R}^d$.

Main tool: the Green function Since the result is immediate for $d \in \{1, 2\}$, we assume $d \geq 3$, which in particular implies that S is transient. As before, the proof will mostly consist of first and second moment computations. We will see that the first moment estimate involves the quantity

$$q(x) := \mathbb{P}\left(S \text{ vists } x\right)$$

for $x \in \mathbb{Z}^d$. We also define the *Green function*, which is more convenient to estimate: for all $x \in \mathbb{Z}^d$, we write

$$G(x) = \mathbb{E}\left[\#\text{visits of } x \text{ by } S\right] = \sum_{n \geq 0} \mathbb{P}\left(S_n = x\right), \tag{1.18}$$

which is finite by transience. In particular, we have

$$q(x) \leq G(x) \leq q(x)G(0),$$

where the first bound is immediate from the definition and the second comes from the strong Markov property applied to the first hitting time of x. Therefore, estimating $G(x)$ will be enough to estimate $q(x)$ up to multiplicative constants.

Lemma 46. Assume $d \geq 3$. There are constants $c_1, c_2 > 0$ (depending on μ) such that for all $x \in \mathbb{Z}^d \backslash \{0\}$, we have

$$\frac{c_1}{|x|^{d-2}} \leq G(x) \leq \frac{c_2}{|x|^{d-2}}.$$

Sketch of proof. The idea of the proof is to start from (1.18) and to estimate each term using the local CLT (this is the reason why G is nicer to work with than q). We start with the lower bound. We write

$$\begin{aligned} G(x) &\geq \sum_{n=|x|^2}^{2|x|^2} \mathbb{P}\left(S_n = x\right) \\ &= \sum_{n=|x|^2}^{2|x|^2} n^{-d/2} \gamma_Q\left(\frac{x}{\sqrt{n}}\right) + o\left(n^{-d/2}\right), \end{aligned}$$

where γ_Q is the density of a centered Gaussian vector with the same covariance matrix Q as μ. In particular, this is bounded away from 0 in the region of interest, so there is a constant c_0 such that all x, we have

$$G(x) \geq c_0 |x|^2 \times \left(2|x|^2\right)^{-d/2} = c_1 |x|^{2-d}.$$

The proof of the upper bound is more annoying, as we need to take into account all values of n, including those which do not fall into the range $|x| = O(\sqrt{n})$. In particular, for $n < |x|^2$, it is possible that the error term $o\left(n^{-d/2}\right)$ appearing in the local CLT is larger than $|x|^{-d}$. This should not be a problem as for a random walk with bounded increments, it is very unlikely to have $|S_n|$ much larger than $\sqrt{n}$, but we need a version of the local CLT with explicit error terms when $|x| >> \sqrt{n}$. We refer to [3, Section 4.3] for the full proof. □

Let us now move on to the proof of Theorem 16. We start with a first moment estimate. It will be convenient to write $u_n \asymp v_n$ if we have both $u_n = O(v_n)$ and $v_n = O(u_n)$. We have

$$\mathbb{E}\left[\mathcal{I}\right] = \sum_{x \in \mathbb{Z}^d} \mathbb{P}\left(S^{(1)}, \dots S^{(k)} \text{ visit } x\right) = \sum_{x \in \mathbb{Z}^d} q(x)^k \asymp \sum_{x \in \mathbb{Z}^d} G(x)^k \asymp \sum_{x \in \mathbb{Z}^d} |x|^{(2-d)k}.$$

We now assume that $|\cdot|$ stands for the L^∞ norm (this was not important so far, since all our estimates were up to multiplicative constants). Then $|x|$ is always an integer and the number of points $x \in \mathbb{Z}^d$ for which $|x| = n$ is of order n^{d-1}, so we can write

$$\mathbb{E}\left[\mathcal{I}\right] \asymp \sum_{n \geq 1} n^{(2-d)k} n^{d-1}.$$

From here, it is easy to check that $\mathbb{E}[\mathcal{I}]$ is finite for $d \geq 5, k \geq 2$ or $d = 4, k \geq 3$ or $d = 3, k \geq 4$, and that it is infinite for $(d,k) = (4,2)$ or $(3,2)$ or $(3,3)$. In particular, in the cases where $\mathbb{E}[\mathcal{I}]$ is finite, this is sufficient to conclude that $\mathcal{I}$ is a.s. finite.

However, in the cases where $\mathbb{E}[\mathcal{I}] = +\infty$, this is not enough to conclude. We need a second moment estimate, but this estimate will not be helpful is the first and second moment are both infinite. Therefore, we will study a *truncated* version of $\mathcal{I}$. For all $n \geq 0$, we denote by $\mathcal{I}_n$ the number of points $x \in \mathbb{Z}^d$ such that for any $1 \leq i \leq k$, the walk S hits x before time n. We also write

$$q_n(x) = \mathbb{P}\left(\tau_x \leq n\right),$$

where τ_x is the hitting time of x. In particular, the same computation as above shows

$$\mathbb{E}\left[\mathcal{I}_n\right] = \sum_{x \in \mathbb{Z}^d} q_n(x)^k.$$

Let us now bound the second moment:

$$\mathbb{E}\left[\mathcal{I}_n^2\right] = \sum_{x,y\in\mathbb{Z}^d} \mathbb{P}\left(x, y \text{ are visited by each } S^{(i)} \text{ before time } n\right) = \sum_{x,y\in\mathbb{Z}^d} \mathbb{P}\left(\tau_x, \tau_y \le n\right)^k.$$

By distinguishing whether τ_x or τ_y occur first and using the strong Markov property on the first of the two to occur, we get the bound

$$\mathbb{E}\left[\mathcal{I}_n^2\right] \le \sum_{x,y\in\mathbb{Z}^d} \left(q_n(x)q_n(y-x) + q_n(y)q_n(x-y)\right)^k.$$

Using the inequality $(a+b)^k \le (2\max(a,b))^k = 2^k \max(a^k, b^k) \le 2^k(a^k+b^k)$, this becomes

$$\mathbb{E}\left[\mathcal{I}_n^2\right] \le 2^{k+1} \sum_{x,y\in\mathbb{Z}^d} q_n(x)^k q_n(y-x)^k = 2^{k+1} \left(\sum_{x\in\mathbb{Z}^d} q_n(x)^k\right)^2 = 2^{k+1}\mathbb{E}\left[\mathcal{I}_n\right]^2. \tag{1.19}$$

This is not as good as the second moment estimates that we obtained for range questions, and it does not for example exclude the possibility that $\mathcal{I}_n$ is zero with probability $\frac{1}{2}$. To make this useful, we will need to use the following lemma instead of the Bienaymé–Chebychev inequality.

Lemma 47. Let $C > 0$, and let X be a real variable X with value in $\mathbb{R}^+$. If $0 < \mathbb{E}[X^2] \le C\mathbb{E}[X]^2 < +\infty$, then

$$\mathbb{P}\left(X \ge \frac{\mathbb{E}[X]}{2}\right) \ge \frac{1}{4C}.$$

Proof. We decompose the expectation of X as follows:

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}\left[X\mathbb{1}_{X<\mathbb{E}[X]/2}\right] + \mathbb{E}\left[X\mathbb{1}_{X\ge\mathbb{E}[X]/2}\right] \\ &\le \frac{\mathbb{E}[X]}{2} + \mathbb{E}[X^2]^{1/2}\mathbb{P}\left(X \ge \frac{\mathbb{E}[X]}{2}\right)^{1/2}\end{aligned}$$

after using the Cauchy–Schwarz inequality for the second term. it follows that

$$\mathbb{P}\left(X \ge \frac{\mathbb{E}[X]}{2}\right) \ge \frac{1}{\mathbb{E}[X^2]}\left(\frac{\mathbb{E}[X]}{2}\right)^2 \ge \frac{1}{4C}.$$

□

End of the proof of Theorem 16. Assume that we are in one of the cases where $\mathbb{E}[\mathcal{I}] = +\infty$. Then (1.19) together with the last lemma shows that for all n, we have

$$\mathbb{P}\left(\mathcal{I}_n \ge \frac{1}{2}\mathbb{E}[\mathcal{I}_n]\right) \ge \frac{1}{2^{k+3}}.$$

Therefore, let $A > 0$. By taking n large enough to have $\mathbb{E}[\mathcal{I}_n] > 2A$, we have

$$\mathbb{P}\left(\mathcal{I} \ge A\right) \ge \mathbb{P}\left(\mathcal{I}_n \ge A\right) \ge \frac{1}{2^{k+3}}.$$

this is true for all A, so $\mathbb{P}\left(\mathcal{I} = +\infty\right) \ge \frac{1}{2^{k+3}}$, and we conclude by the 0-1 law of Lemma 45. □

Chapter 2

Random walks on general graphs

2.1 Electric networks

In this chapter, we introduce the theory of electric networks. On the one hand, it draws a very nice parallel between the behaviour of electric circuits in physics and properties of random walks on general graphs. On the other hand, we will see that it is a very powerful tool to estimate certain quantities related to random walks, and in particular to find out if the random walk on a given graph is recurrent or transient.

To illustrate why these techniques can be way more powerful than what we have seen so far, let us consider the graph consisting of $\mathbb{Z}^2$ plus an additional edge linking $(0,0)$ to $(1,1)$: this graph should be reccurrent since it is very similar to $\mathbb{Z}^2$. However, the approaches that we already know to swho the recurrence of $\mathbb{Z}^2$ (counting of pathhs via binomial coefficients or Fourier transform) seem to fail miserably in this case. The electric networks approach is way more robust: once we will be able to show the recurrence of $\mathbb{Z}^2$, the exact same proof will work if we add or remove a "reasonnable" number of edges.

2.1.1 Networks, voltages and flows

We will start with (a lot of) definitions.

Definition 48.

- A *network* is a graph $G = (V, E)$ (where V is the set of *vertices* and E the set of *edges*) equipped with a function $c : E \to [0, +\infty]$. We call $c(e)$ the *conductance* of e and $r(e) = \frac{1}{c(e)}$ its *resistance*.
- We say that a network is *connected* if the graph (G, E^+) is connected, where $E^+ \subset E$ is the set of edges e such that $c(e) > 0$.

We will always assume that the networks that we work with are connected. If the edge e joins the vertices u and v, we will also write $c(u, v)$ for $c(e)$.

Random walk on a network. For $v \in V$, we write

$$\pi(v) = \sum_{\{u,v\} \in E} c(u, v).$$

If $\pi(v) < +\infty$ for all v, we can define the random walk on our network as the Markov chain $(S_n)_{n_geq0}$ with transitions

$$\mathbb{P}\left(S_{n+1} = v | S_n = u\right) = \frac{c(u,v)}{\pi(u)}.$$

Remarks 49.

- If all the conductances $c(e)$ are equal to 1, this is just the simple random walk on the graph G.
- From the definition, it is clear that π is a reversible leasure for the Markov chain S. Moreover, a Markov chain is the random walk associated to a certain network if and only if it has a reversible measure π (and in this case, we can take $c(u,v) = \pi(u)p(u,v)$).

We will first be interested in the electric potential in our network. Potentials in physics are often given by harmonic functions.

Definition 50.

- A function $f : V \to \mathbb{R}$ is *harmonic* at $v \in V$ if

$$h(v) = \frac{1}{\pi(v)} \sum_{\{u,v\} \in E} c(u,v)h(u).$$

- Given two distinct vertices a and z, a *voltage from a to z* is a function $h : V \to \mathbb{R}$ that is harmonic at every point of $V \backslash \{a, z\}$.

We should think of a and z as the two points of our electric network that are plugged to the current source. As often in mathematics, let us first prove that the physical solution exists and is unique.

Proposition 51. In a finite network, for any $\alpha, \beta \in \mathbb{R}$, there is a unique voltage from a to z such that $h(a) = \alpha$ and $h(z) = \beta$.

Proof.

- As usual, uniqueness will come from the maximum principle. Let h_1, h_2 be two solutions. Then $\widetilde{h} = h_1 - h_2$ is a voltage with $\widetilde{h}(a) = \widetilde{h}(z) = 0$. Let $M \geq 0$ be its maximal value, and let A be the set of vertices x where $\widetilde{h}(x) = M$. If $v \in A$ with $v \neq a, z$, then

$$M = \widetilde{h}(v) = \frac{1}{\pi(v)} \sum_{\{u,v\} \in E} \widetilde{h}(u) \leq \frac{1}{\pi(v)} \sum_{\{u,v\} \in E} \widetilde{h}(v) = h(v)$$

  by maximality of $h(v)$. In particular, we must have equality in each term, so for all u such that $c(u,v) > 0$, we have $h(u) = h(v) = M$ so $u \in A$. By connectedness, there is a path from v to a or to z, so $a \in A$ or $z \in A$, so $M = 0$. The same argument shows $\min \widetilde{h} = 0$, so $h_1 = h_2$.

- For the existence, we can just take

$$h(v) = \alpha + (\beta - \alpha)\mathbb{P}_v\left(\tau_z < \tau_a\right),$$

  where $\tau_u = \min\{n \geq 0 | S_n = u\}$ and $\mathbb{P}_v$ is the law of the random walk on our network started from v. This is a voltage by applying the weak Markov property at time 1.

□

Flows. In physics, an electric circuit can be understood either from the study of the electric potential, or from the study of the current intensity in each piece of the circuit. It will be important for us to be able to hadle both approaches, so we now define flows, which will describe the "current intensities" in our networks.

We define $\vec{E}$ as the set of oriented edges of our graph. That is, for each edge $e = \{u, v\}$ of E, the set $\vec{E}$ contains two oriented edges (u, v) and (v, u).

Definition 52. A *flow* from a to z is a function $\theta : \vec{E} \to \mathbb{R}$ satisfying

1. for any $\{u, v\} \in E$, we have $\theta(vu) = -\theta(uv)$,
2. for any $v \in V \backslash \{a, z\}$, we have $\sum_{\{u,v\}\in E} \theta(uv) = 0$.

The second condition says that each vertex has as much current coming in as going out. It is known as Kichhoff's current law ("loi des noeuds" in french). Note that these conditions are not sufficient to describe the current flow on an electric network: there are many possible flows that do not correspond to the "physical" reality.

In physics, the electric potential and the current intensity are linked by the relation $U = RI$. We can also use this relation to define a flow associated to a voltage.

Definition 53. Given a voltage h, the current flow θ_h associated to h is given by

$$\theta(uv) = c(u, v) \left(h(u) - h(v)\right). \tag{2.1}$$

Note that this is indeed a flow: for each $u \in V \backslash \{a, z\}$, we have

$$\sum_{\{u,v\}\in E} \theta(uv) = \sum_{\{u,v\}\in E} c(u, v) \left(h(u) - h(v)\right) = \pi(u)h(u) - \sum_{\{u,v\}\in E} c(u, v)h(v) = 0$$

by harmonicity. Moreover, the flow θ_h satisfies the *cycle law* ("loi des mailles" in french): for any cycle $(x_0, x_1, \ldots, x_m = x_0)$ in the graph, we have

$$\sum_{i=1}^{m} r(x_{i-1}, x_i)\theta(x_{i-1}x_i) = \sum_{i=1}^{m} \left(h(x_{i-1}) - h(x_i)\right) = 0.$$

This property turns out to be very important, as it characterizes the flows that actually occurs in physics.

Definition 54. If a flow θ satisfies the cycle law, then there is a voltage h_θ such that θ is the current flow associated to h_θ. Moreover, this voltage is unique up to an additive constant.

Note that the sign convention here is consistent with the physical one: the current flows from the endpoint with the higher potential to the point with the lower potential.

Proof. • We start with uniqueness: if both h and h' work, then (2.1) shows that for any two neighbours u and v with $c(u, v) > 0$, we have

$$h(u) - h(v) = h'(u) - h'(v) = r(u, v)\theta(uv),$$

so $h(u) - h'(u) = h(v) - h'(v)$. By connectedness, this proves that $h - h'$ is constant, so h is unique up to an additive constant.

- For the existence, we note again that if we know $h(u)$ for some u, then (2.1) fixes the value of h on its neighbours. Therefore, to find $h(v)$, we can always fix $h(a) = 0$ and use the path from a to v. The cycle law will then guarantee that the value we obtain does not depend on the choice of the path.

  More precisely, let $v \in V$ and let $(a = x_0, x_1, \dots, x_k = v)$ be a path from a to v with $c(x_{i-1}x_i) > 0$ for all i. We write

  $$h(v) = -\sum_{i=1}^{k} r(x_{i-1}, x_i)\theta(x_{i-1}x_i). \tag{2.2}$$

  If $\mathbf{x}$ and $\mathbf{y}$ are two such paths, then the concatenation of $\mathbf{x}$ forwards and $\mathbf{y}$ backwards is a cycle, and the cycle law applied to this cycle proves that the value of $h(v)$ provided by (2.2) is the same using $\mathbf{x}$ or $\mathbf{y}$. Therefore h is well defined and does not depend on arbitrary choices.

  We know check that $\theta_h = \theta$: if $\{u, v\} \in E$, we can apply (2.2) using the path $(x_0, \dots, x_k = v, x_{k+1} = u)$ from a to u. We find

  $$h(u) = h(v) - r(v, u)\theta(vu), \tag{2.3}$$

  so $\theta(uv) = c(u, v)\,(h(u) - h(v))$. Finally, let us check that h is indeed a voltage. For any $v \in V\backslash\{a, z\}$, using (2.3), we have

  $$\begin{aligned}\frac{1}{\pi(v)} \sum_{\{u,v\}\in E} c(u, v)h(u) &= h(v) + \frac{1}{\pi(v)} \sum_{\{u,v\}\in E} c(u, v)r(u, v)\theta(vu) \\ &= h(v) + \frac{1}{\pi(v)} \sum_{\{u,v\}\in E} \theta(vu) = h(v)\end{aligned}$$

  because θ is a flow, which concludes the proof.

□

We now define the strength of a flow, which corresponds to the total quantity of current flowing in and out of an electric circuit.

Definition 55. The *strength* of a flow θ from a to z is the quantity

$$\|\theta\| = \sum_{\{a,v\}\in E} \theta(av) = \sum_{\{v,z\}\in E} \theta(vz).$$

Note that despite the notation, the quantity $\|\theta\|$ does not have to be positive!

Proof of the last formula. By summing the antisymmetry of a flow over all the edges, we find

$$0 = \sum_{\vec{e}\in\vec{E}} \theta(\vec{e}) = \sum_{u\in V} \sum_{\{u,v\}\in E} \theta(uv).$$

But we know that this quantity is 0 for all $u \neq a, z$, so

$$0 = \sum_{\{a,v\}\in E} \theta(av) + \sum_{\{z,v\}\in E} \theta(zv),$$

which shows that both definitions are equivalent. □

Finally, we check that injecting an electric current in a circuit indeed defines uniquely a current flow.

Proposition 56. For any $I \in \mathbb{R}$, there is a unique flow θ which satisfies the cycle law and $\|\theta\| = I$. We will call such a flow a *current flow*.

Proof.
- Let us start with uniqueness. If θ_1 and θ_2 both work, then $\theta = \theta_1 - \theta_2$ is a flow satisfying the cycle law, so there is a voltage h with $\theta = \theta_h$. Moreover, we have $\|\theta\| = \|\theta_1\| - \|\theta_2\| = 0$. Rewriting this in terms of h means that h is harmonic at a and z, so h is harmonic on all the network. By connectedness and the maximum principle, this means that h is constant, so $\theta = 0$ and $\theta_1 = \theta_2$.

- To prove the existence, let h be the voltage with $h(a) = 1$ and $h(z) = 0$. We want to prove $\|\theta_h\| > 0$. For this, let $(x_0 = a, x_1, \dots, x_k = z)$ be an injective path from a to z (such a path exists by connectedness). We know that for all v, we have

$$h(v) = \mathbb{P}_v(\tau_a < \tau_z) \leq 1 = h(a).$$

  But since the random walk has a positive probability of using exactly our path (x_i), we also have

$$h(x_1) = \mathbb{P}_{x_1}(\tau_a < \tau_z) < 1 = h(a).$$

  Therefore, we can write

$$\|\theta_h\| = \sum_{\{a,v\} \in E} c(a,v)\,(h(a) - h(v)) \geq c(a,x_1)\,(h(a) - h(x_1)) > 0.$$

  Therefore, there is a multiple of $\|\theta_h\|$ with strength exactly I.

□

2.1.2 Effective resistances

In this section, we will always consider finite, connected networks with two distinguished vertices a and z. Our next goal is to define the *effective resistance*: no matter how complicated a network is, it will always behave like a single electric resistance (whose value depends on the network of course).

Definition 57. Let h be a non-constant voltage and let θ_h be the current flow associated to h. The ratio $\frac{h(a)-h(z)}{\|\theta_h\|}$ does not depend on the choice of h, and is called the *effective resistance* between a and z. We write it $R_{\text{eff}}(a \leftrightarrow a)$.

Proof. We know that $\|\theta_h\|$ is a linear function of h, so it is a linear function of the pair $(h(a), h(z))$. We also know that if $h(a) = h(z)$, then h is constant so $\|\theta_h\| = 0$, so there is a constant $r \in \mathbb{R}$ such that

$$\|\theta_h\| = r\,(h(z) - h(a)).$$

Finally, the existence proof of Proposition 56 shows that if $h(a) = 0$ and $h(z) = 1$, then $\|\theta_h\| > 0$, so $r > 0$. □

Let us now (finally!) do some probability again: we will provide a probabilistic interpretation of the effective resistance. We will write

$$\tau_a^+ = \min\{n \geq 1 | S_n = a\}.$$

Proposition 58. We have

$$R_{\text{eff}}(a \leftrightarrow z) = \frac{1}{\pi(a)\mathbb{P}_a\left(\tau_z < \tau_a^+\right)}.$$

Proof. Let h be the voltage from a to z with $h(a) = 0$ and $h(z) = 1$. Then for any $v \neq a, z$, we have $h(v) = \mathbb{P}_v(\tau_z < \tau_a)$. hence, using the weak Markov property at time 1, we can write

$$\begin{aligned}
\mathbb{P}_a\left(\tau_z < \tau_a^+\right) &= \frac{1}{\pi(a)} \sum_{\{a,v\} \in E} c(a,v)\mathbb{P}_v\left(\tau_z < \tau_a\right) \\
&= \frac{1}{\pi(a)} \sum_{\{a,v\} \in E} c(a,v)\left(h(v) - h(a)\right) \\
&= \frac{1}{\pi(a)} \sum_{\{a,v\} \in E} \theta(va) \\
&= -\frac{\|\theta\|}{\pi(a)} \\
&= \frac{1}{\pi(a) R_{\text{eff}}(a \leftrightarrow z)}.
\end{aligned}$$

□

Roughly speaking, the reason why this last result is very important is that when we will work with infinite networks, by letting z go to "infinity", the recurrence or transience of S will be equivalent to whether $R_{\text{eff}}(a \leftrightarrow \infty)$ is finite or not.

Some computation rules. We can now state some of the usual computation rules for effective resistances: the series and parallel laws. Note that these two rules are far from being sufficient to compute the effective resistance in any network!

- Parallel Law: If there are to edges with conductances c_1 and c_2 linking the same pair $\{u, v\}$ of vertices, then replacing them by one single edge with conductance $c_1 + c_2$ does not change the value of $R_{\text{eff}}(a \leftrightarrow z)$. This is immediate, since any voltage before the transformation is still a voltage after the transformation.
- Series Law: If a vertex v is only linked to two vertices u (by a resistance r_1) and w (by a resistance r_2), then removing v and linking u to w by a resistance $r_1 + r_2$ does not change $R_{\text{eff}}(a \leftrightarrow z)$. Indeed, a current flow before the transformation can easily be turned into a current flow after the transformation, without affecting the strength of the flow or the associated voltage.

We now provide a second probabilistic interpretation of the effective resistance.

Proposition 59 (Commute-time identity)**.** We have

$$\mathbb{E}_a[\tau_z] + \mathbb{E}_z[\tau_a] = 2R_{\text{eff}}(a \leftrightarrow z) \sum_{e \in E} c(e).$$

Proof. For any u, we define

$$G_1(u) = \mathbb{E}_a\left[\#\text{visits at } u \text{ strictly before } \tau_z\right].$$

In particular, we have

$$\mathbb{E}_a\left[\tau_z\right] = \sum_{u\in V} G_1(u).$$

Then we claim that the function $\widetilde{G}_1(u) = \frac{G_1(u)}{\pi(u)}$ is harmonic on $V\backslash\{a,z\}$. Indeed, we can compute, for any $v \neq a, z$:

$$\begin{aligned}
\frac{1}{\pi(v)}\sum_{\{u,v\}\in E} c(u,v)\frac{G_1(u)}{\pi(u)} &= \frac{1}{\pi(v)}\sum_{\{u,v\}\in E}\sum_{n\geq 0}\frac{c(u,v)}{\pi(u)}\mathbb{P}\left(S_n = u,\, \tau_z > n\right)\\
&= \frac{1}{\pi(v)}\sum_{\{u,v\}\in E}\sum_{n\geq 0}\mathbb{P}\left(S_n = u,\, \tau_z > n,\, S_{n+1} = v\right)\\
&= \frac{1}{\pi(v)}\sum_{n\geq 0}\sum_{\{u,v\}\in E}\mathbb{P}\left(S_n = u,\, \tau_z > n+1,\, S_{n+1} = v\right)\\
&= \frac{1}{\pi(v)}\sum_{n\geq 0}\mathbb{P}\left(\tau_z > n+1,\, S_{n+1} = v\right)\\
&= \frac{1}{\pi(v)}\sum_{n\geq 1}\mathbb{P}\left(\tau_z > n,\, S_n = v\right)\\
&= \frac{G_1(v)}{\pi(v)}.
\end{aligned}$$

Note that we have used $v \neq z$ to replace $\tau_z > n$ by $\tau_z > n+1$, and $v \neq a$ in the very end to replace the sum over $n \geq 1$ by a sum over $n \geq 0$. Moreover, we know that $\widetilde{G}_1(z) = 0$ and that the number of visits of a before time τ_z is a geometric random variable with success probability $\mathbb{P}\left(\tau_z < \tau_a^+\right)$, so

$$\widetilde{G}_1(a) = \frac{1}{\pi(a)\mathbb{P}_a\left(\tau_z < \tau_a^+\right)} = R_{\text{eff}}\left(a \leftrightarrow z\right).$$

We can now define symetrically $G_2(u)$ as the expected number of vists at u before time a for the walk started from z and $\widetilde{G}_2(u) = \frac{G_2(u)}{\pi(u)}$. Then $\widetilde{G}_2$ is a voltage with $\widetilde{G}_2(a) = 0$ and $\widetilde{G}_2(z) = R_{\text{eff}}\left(a \leftrightarrow z\right)$. Therefore $\widetilde{G}_1 + \widetilde{G}_2$ is a voltage taking the same value at a and z, so it is constant and equal to $R_{\text{eff}}\left(a \leftrightarrow z\right)$. We can finally conclude

$$\mathbb{E}_a\left[\tau_z\right] + \mathbb{E}_z\left[\tau_a\right] = \sum_{u\in V}\pi(u)\widetilde{G}_1(u) + \pi(u)\widetilde{G}_2(u) = \sum_{u\in V}\pi(u)R_{\text{eff}}\left(a \leftrightarrow z\right) = 2\left(\sum_{e\in E} c(e)\right)R_{\text{eff}}\left(a \leftrightarrow z\right).$$

□

2.1.3 Energy and variational principles

We will now give another description of the effective resistance as the minimizer for some notions of energy. This description is not only physically relevant (physics love configurations minimizing energy), but it is also the reason why electric networks are such a powerful tool to study random walks. Indeed, if the effective resistance is the minimal energy of a flow in some class, then any

flow in that class provides an upper bound on the resistance. In particular, it becomes possible to prove reasonable estimates without computing exactly the resistances, just from guessing what a "close to harmonic function" should look like! As usual, our notion of energy will be directly inspired from physics.

Definition 60. The *energy* of a flow from a to z is the quantity

$$\mathcal{E}(\theta) = \frac{1}{2} \sum_{\vec{e} \in \vec{E}} r(\vec{e}) \theta(\vec{e})^2.$$

Note that $\theta(e)$ is not only defined up to its sign for a non-oriented edge e, but its square is well-defined.

Theorem 17. We have

$$R_{\text{eff}}(a \leftrightarrow z) = \min \{ \mathcal{E}(\theta) \,|\, \theta \text{ flow from } a \text{ to } z \text{ with } \|\theta\| = 1 \}.$$

Moreover, the unit current flow satisfying the cycle law is the unique minimizer.

Proof. Let I be the unit current flow from a to z, and h the corresponding voltage (which is defined up to an additive constant). We first check that $\mathcal{E}(\mathcal{I}) = R_{\text{eff}}(a \leftrightarrow z)$:

$$\begin{aligned}
\mathcal{E}(I) &= \frac{1}{2} \sum_u \sum_{\{u,v\} \in E} r(u,v) I(uv)^2 \\
&= \frac{1}{2} \sum_u \sum_{\{u,v\} \in E} (h(u) - h(v)) \, I(uv) \\
&= \frac{1}{2} \sum_u h(u) \sum_{\{u,v\} \in E} I(uv) - \frac{1}{2} \sum_v h(v) \sum_{\{u,v\} \in E} I(uv) \\
&= \frac{1}{2} (h(a) - h(z)) - \frac{1}{2} (h(z) - h(a)) \\
&= h(a) - h(z) = R_{\text{eff}}(a \leftrightarrow z),
\end{aligned}$$

where we have used that the total current going out of a vertex u is 1 for $u = a$, is -1 for $u = z$ and is 0 else.

We now check that any unit flow has energy larger than I. Let J be a flow from a to z with $\|J\| = 1$, and let $\theta = J - I$, so that $\|\theta\| = 0$. By expanding the square, we can write

$$\begin{aligned}
\mathcal{E}(J) &= \frac{1}{2} \sum_u \sum_{\{u,v\} \in E} r(u,v) \left(I(uv) + \theta(uv) \right)^2 \\
&= \mathcal{E}(I) + \mathcal{E}(\theta) + \sum_u \sum_{\{u,v\} \in E} r(u,v) I(uv) \theta(uv) \\
&= \mathcal{E}(I) + \mathcal{E}(\theta) + \sum_u \sum_{\{u,v\} \in E} (h(u) - h(v)) \, \theta(uv).
\end{aligned}$$

The same computation as in the first part of the proof shows that the big sum is equal to

$$\|\theta\| (h(z) - h(a)) = 0,$$

so

$$\mathcal{E}(J) = \mathcal{E}(I) + \mathcal{E}(\theta) \geq \mathcal{E}(I),$$

with equality if and only if $\theta = 0$, i.e. $J = I$. This shows that I is the unique minimizer and completes the proof. □

We now state a few direct consequences. Note that although the following results may seem intuitively clear, their proofs would not be immediate without the last result.

Corollary 61 (Rayleigh monotonicity law)**.** Let $(r_1(e))_{e\in E}$ and $(r_2(e))_{e\in E}$ be such that for any e, we have $r_1(e) \leq r_2(e)$. Then

$$\mathbb{R}_{\text{eff}}\left(a \leftrightarrow z \text{ in } (G, r_1)\right) \leq \mathbb{R}_{\text{eff}}\left(a \leftrightarrow z \text{ in } (G, r_2)\right).$$

Proof. Let θ be any unit flow on G. We have

$$\sum_{e\in E} r_1(e)\theta(e)^2 \leq \sum_{e\in E} r_2(e)\theta(e)^2.$$

and we conclude by taking the infimum over θ. The important point here is that the notion of flow is independent from the conductances, so we consider the same set of flows for both sides. □

Corollary 62. Gluing two vertices together or adding a new edge never increase effective resistances.

Intuitively speaking, these results can be thought of as special cases of the monotonicity law, with the convention that merging two vertices means having a resistance 0 between them, and no edge is equivalent to an edge with resistance ∞. Let us still proved a full proof.

Proof.

- Let I be the unit current flow on G, and let G' be obtained from G by merging two vertices. Then I can still be seen as a flow on G' (if $u' \in G'$ was obtained by merging u_1 and u_2 from G, then the outgoing current from u' is equal to the sum of the outgoing currents from u_1 and u_2, which is 0 unless one of them is a or z). Therefore, we can write

  $$\mathbb{R}_{\text{eff}}\left(a \leftrightarrow z \text{ in } G'\right) \leq \mathcal{E}_{G'}(I) \leq \mathcal{E}_G(I) = \mathbb{R}_{\text{eff}}\left(a \leftrightarrow z \text{ in } G\right),$$

  where the last equality follows from the definition of I. Note that the $\mathcal{E}_{G'}(I) \leq \mathcal{E}_G(I)$ does not have to be an equality, because it is possible that an edge between the two vertices that were merged was destroyed in the process.

- The proof is exactly the same: a flow on a network G is still a flow after new edges have been added.

□

Theorem 17 is a very useful tool to obtain upper bounds on resistances. We now state an analog result for lower bounds. For this, we will work not on flows, but on functions on vertices.

Definition 63. The *energy* of a function $f : V \to \mathbb{R}$ is the quantity

$$\mathcal{E}(f) = \sum_{\{u,v\}\in E} c(u,v)\left(h(u) - h(v)\right)^2.$$

Theorem 18. We have

$$\frac{1}{R_{\text{eff}}\,(a \leftrightarrow z)} = \min\{\mathcal{E}(f) \,|\, f : V \to \mathbb{R},\, f(a) = 1,\, f(z) = 0\}.$$

Proof. The energy $\mathcal{E}$ is a continuous function from the finite-dimensional space $\mathbb{R}^{V\backslash\{a,z\}}$ to $\mathbb{R}$. Moreover, we have $\mathcal{E}(f) \to +\infty$ as $f \to \infty$. Indeed, if at least one coordinate of f is very large, then since $f(z) = 0$ there is at least one edge of G along which the increment of f is very large, so $\mathcal{E}(f)$ is very large as well. This implies that $\mathcal{E}$ has a minimizer f_0. Since $\mathcal{E}$ is a smooth function, its differential must vanish at f_0. that is, for any vertex $u \in V\backslash\{a, z\}$, we must have

$$0 = \frac{\partial \mathcal{E}(f)}{\partial f(u)}(f_0) = \sum_{\{u,v\}\in E} \frac{\partial}{\partial f(u)} \left(f(u) - f(v)\right)^2 (f_0) = \sum_{\{u,v\}\in E} 2c(u, v)\left(f(u) - f(v)\right).$$

This last equation precisely means that f_0 is harmonic at u, so f_0 is the voltage h on G with $h(a) = 1$ and $h(z) = 0$. We have proved that h is the unique minimizer of $\mathcal{E}$.

To conclude, we compute

$$\begin{aligned}
\mathcal{E}(h) &= \sum_{\{u,v\}\in E} c(u, v)\left(h(u) - h(v)\right)^2 \\
&= \sum_{\{u,v\}\in E} r(u, v)\theta_h(uv)^2 \\
&= \mathcal{E}(\theta_h) \\
&= \|\theta_h\|^2 \mathcal{E}\left(\frac{\theta_h}{\|\theta_h\|}\right).
\end{aligned}$$

On the one hand, by definition of the effective resistance, we have $\|\theta\| = \frac{1}{R_{\text{eff}}(a\leftrightarrow z)}$. On the other hand $\frac{\theta_h}{\|\theta_h\|}$ is the unit current flow from a to z, so it has energy $R_{\text{eff}}\,(a \leftrightarrow z)$ by Theorem 17 and the result follows. □

2.1.4 Resistances in infinite networks

Now that we have tools to estimate effective resistances, we want to adapt the theory to the setting of infinite networks, in order to apply it to recurrence and transience of random walks. In this section, we will assume that G is infinite but locally finite (i.e. every vertex is incident to finitely many edges) and connected.

We fix a vertex $a \in G$. Our goal is to define the effective resistance from a to infinity. For this, we will rely on finite aproximations. Let $(G_n)_{n\geq 1}$ be a sequence of subsets of V such that for any n, we have $a \in G_n \subset G_{n+1}$ and

$$\bigcup_{n\geq 1} G_n = V.$$

Let also z_n be the vertex obtained by merging together all the vertices of $G\backslash G_n$. Then $G_n \cup \{z_n\}$ is a finite graph, whose edges correspond to all the edges of G with at least one endpoint in G_n.

Proposition 64. The limit

$$\lim_{n\to+\infty} R_{\text{eff}}\,(a \leftrightarrow z_n \text{ in } G_n \cup \{z_n\})$$

exists in $(0, +\infty]$, and does not depend on the choice of (G_n) (as long as it satisfies the assumptions above). We write it $R_{\text{eff}}\,(a \leftrightarrow \infty)$.

Proof. The graph $G_n \cup \{z_n\}$ can be built from $G_{n+1} \cup \{z_{n+1}\}$ by merging z_{n+1} with all the vertices of $G_{n+1} \backslash G_n$. By Corollary 62, this implies that $\mathbb{R}_{\text{eff}}(a \leftrightarrow z_n)$ is nondecreasing in n, so it converges in $(0, +\infty]$. Moreover, let (G'_n) be another choice of approximations. By induction, we can build two sequences (i_k) and (j_k) of indices such that for any $k \geq 1$, we have

$$G_{i_k} \subset G'_{j_k} \subset G_{i_{k+1}}.$$

The limit of effective resistances along this subsequence exists, so the limit along $(G_{i,k})$ and along $\left(G'_{j_k}\right)$ must be the same. It follows that

$$\lim_{n \to +\infty} R_{\text{eff}}(a \leftrightarrow z_n \text{ in } G_n \cup \{z_n\}) = \lim_{n \to +\infty} R_{\text{eff}}\left(a \leftrightarrow z'_n \text{ in } G'_n \cup \{z'_n\}\right).$$

□

Resistances and recurrence. With the last definition in our hands, we can now discuss the link between electric resistances and recurrence of random walks. Given an infinite electric network, we can define the associated random walk in the exact same way as in the finite case: this is the Markov chain with transitions $p(u, v) = \frac{c(u,v)}{\pi(u)}$. By Proposition 58, for any n, we can write

$$R_{\text{eff}}(a \leftrightarrow z_n \text{ in } G_n \cup \{z_n\}) = \frac{1}{\pi(a)\mathbb{P}_a\left(\tau_{G \backslash G_n} < \tau_a^+\right)}.$$

Indeed, the hitting time of z_n for the random walk in the finite graph $G_n \cup \{z_n\}$ corresponds to the hitting time of $G \backslash G_n$ for the random walk on G. When we let n go to $+\infty$, both sides converge and we get

$$R_{\text{eff}}(a \leftrightarrow \infty) = \frac{1}{\pi(a)\mathbb{P}_a\left(\tau_a^+ = +\infty\right)}.$$

In particular, noting that the denominator on the right is 0 if and only if S is recurrent, we have proved the following result.

Theorem 19. The walk S is recurrent if and only if $R_{\text{eff}}(a \leftrightarrow \infty) = +\infty$, and this property does not depend on a.

By combining this with Corollary 62, we can obtain (for the first time in this chapter!) a result whose statement does not involve any electric networks!

Corollary 65. If a network G is transient, then any network obtained from G by adding edges and merging vertices (in such a way that the graph remains locally finite) is still transient.

Flows on infinite graphs. Sometimes, it will be convenient to use flows directly on an infinite graphs, rather than on finite approximations. We can define flows on infinite graphs in a very similar way as in the finite case. The only difference is that we will work with flows from a vertex a to ∞: for any vertex $u \neq a$, we must have

$$\sum_{\{u,v\} \in E} \theta(uv) = 0$$

and we define the strength as $\|\theta\| = \sum_{\{a,v\} \in E} \theta(av)$.

Theorem 20. We have

$$R_{\text{eff}}(a \leftrightarrow \infty) = \min\{\mathcal{E}(\theta) \,|\, \theta \text{ flow from } a \text{ to } \infty,\ \|\theta\| = 1\}. \tag{2.4}$$

Proof. We start with the easy direction. Let θ be any unit flow on G from a to ∞. Then θ induces a flow on $G_n \cup \{z_n\}$ from a to z_n with $\|\theta_n\| = \|\theta\| = 1$. We can then write

$$R_{\text{eff}}(a \leftrightarrow z_n) \leq \mathcal{E}(\theta_n) \leq \mathcal{E}(\theta).$$

By letting $n \to +\infty$, we find $R_{\text{eff}}(a \leftrightarrow \infty) \leq \mathcal{E}(\theta)$, so

$$R_{\text{eff}}(a \leftrightarrow \infty) \leq \inf\{\mathcal{E}(\theta) \,|\, \theta \text{ flow from } a \text{ to } \infty,\ \|\theta\| = 1\}.$$

To conclude the proof, it is sufficient to build a unit flow on G with energy $R_{\text{eff}}(a \leftrightarrow \infty)$. We build it by extracting a converging subsequence from current flows on finite approximations. In order to have the compactness we need, we start with the following lemma.

Lemma 66. Let I be the unit current flow on a finite network G_f. Then for any oriented edge $\vec{e}_0$, we have $|I(\vec{e}_0)| \leq 1$.

Proof. Let h be the associated voltage with $h(a) > h(z) = 0$. We write $\vec{e}_0 = (x, y)$ and assume $h(x) > h(y)$ without loss of generality (if $h(x) = h(y)$, then $I(\vec{e}_0) = 0$). Let A be the set of vertices u such that $h(u) \geq h(x)$. Finally, let ∂A be the set of oriented edges with their starting point inside A and their terminal point outside. In particular, we have $\vec{e}_0 \in \partial A$ and for any $\vec{e} \in \partial A$, we have $I(\vec{e}) \geq 0$. Moreover, let $\widetilde{G}$ be the network obtained by merging together all the vertices in the complement of A. Then the quantity $\sum_{\vec{e} \in \partial A} I(\vec{e})$ corresponds to the total current coming in z in $\widetilde{G}$. Since I is still a flow (but maybe not a current flow) in $\widetilde{G}$, this quantity is the strength of I in $\widetilde{G}$, and therefore also in G. We have therefore proved

$$I(\vec{e}_0) \leq \sum_{\vec{e} \in \partial A} I(\vec{e}) = \|I\| = 1.$$

□

Now, for all $n \geq 1$, let $\theta^{(n)}$ be the unit current flow from a to z_n on $G_n \cup \{z_n\}$. By the last lemma, for any edge e, we have $|\theta^{(n)}| \leq 1$ for n large enough. Therefore, by a diagonal extraction argument, we can find a function $\theta : \vec{E} \to \mathbb{R}$ and a subsequence (n_k) such that for any $\vec{e} \in \vec{E}$, we have

$$\theta(\vec{e}) = \lim_{k \to +\infty} \theta^{(n_k)}(\vec{e}).$$

It is then easy to check that θ is a unit flow in G from a to ∞. To conclude, we only need to check that $\mathcal{E}(\theta) \leq R_{\text{eff}}(a \leftrightarrow \infty)$. For this, we enumerate the edges of G as $(e_i)_{i \geq 1}$. Now fix $m \geq 1$. For k large enough, we can write

$$\sum_{i=1}^{m} r(e_i)\theta^{(n_k)}(e_i)^2 \leq \mathcal{E}(\theta^{(n_k)}) = R_{\text{eff}}(a \leftrightarrow z_{n_k} \text{ in } G_{n_k} \cup \{z_{n_k}\}).$$

By letting $k \to +\infty$, we obtain

$$\sum_{i=1}^{m} r(e_i)\theta(e_i)^2 \leq R_{\text{eff}}(a \leftrightarrow \infty \text{ in } G).$$

Finally, letting $m \to +\infty$ gives $\mathcal{E}(\theta) \leq R_{\text{eff}}(a \leftrightarrow \infty \text{ in } G)$ and concludes the proof. □

In particular, by looking whether (2.4) is finite or not, we immediately obtain the following result.

Corollary 67. The random walk S is transient if and only if the network admits a unit flow from a to ∞ with finite energy.

2.1.5 Recurrence and transience criteria and applications

Our goal is now to turn the criterion given by Theorem 19 into more explicit criteria that can be directly be applied to examples. We start with a simple sufficient criterion for recurrence, that relies on the notion of *cutset*.

Definition 68. A *cutset* separating a from z (or from ∞) is a set of edges such that every path from a to z (or to ∞) uses at least one edge of Γ.

Roughly speaking, small cutsets mean that the graph is "not very connected", and that the random walk should have difficulties escaping to infinity, so they should provide *lower* bounds on resistances.

Theorem 21 (Nash–Williams inequality)**.** Let $(\Gamma_i)_{1\leq i\leq n}$ be disjoint cutsets separating a from z in a finite network. Then

$$R_{\text{eff}}\left(a \leftrightarrow z\right) \geq \sum_{i=1}^{n} \left(\sum_{e\in\Gamma_i} c(e) \right)^{-1}. \tag{2.5}$$

Proof in a particular case. We first give a proof which does not cover the general case, but which makes it clearer where (2.5) comes from. We make the additional assumption that the cutsets Γ_i are "nested" in the sense that Γ_i separates $a, \Gamma_1, \ldots, \Gamma_{i-1}$ from $\Gamma_{i+1}, \ldots, \Gamma_n, z$. In this case, for $0 \leq i \leq n$, we define V_i as the set of vertices lying between Γ_i and Γ_{i+1} (with the convention that $\Gamma_0 = \{a\}$, $\Gamma_{n+1} = \{z\}$ and that $a \in V_0$, $z \in V_n$). Let also $\widetilde{G}$ be the network obtained by merging each of $V_0, \ldots, V_n$ into one vertex $\widetilde{v}_0, \ldots, \widetilde{v}_n$. Then $\widetilde{G}$ just has these $n+1$ vertices and, for any $1 \leq i \leq n$, there is one edge from $\widetilde{v}_{i-1}$ to $\widetilde{v}_i$ corresponding to each edge of Γ_i. By monotonicity, we can write

$$\begin{aligned} R_{\text{eff}}\left(a \leftrightarrow z \text{ in } G\right) &\geq R_{\text{eff}}\left(\widetilde{v}_0 \leftrightarrow \widetilde{v}_n \text{ in } \widetilde{G}\right) \\ &= \sum_{i=1}^{n} R_{\text{eff}}\left(\widetilde{v}_{i-1} \leftrightarrow \widetilde{v}_i\right) \\ &= \sum_{i=1}^{n} \left(\sum_{e\in\Gamma_i} c(e) \right)^{-1}, \end{aligned}$$

where the computation of $R_{\text{eff}}\left(\widetilde{v}_0 \leftrightarrow \widetilde{v}_n\right)$ is done using just the series and parallel rules. □

Note that as shown on Figure 2.1, there are families of disjoint cutsets to which the above "proof" does not apply. We now move on to the "true" proof of the Nash–Williams inequality. We start with a lemma bounding from below the flow through a cutset.

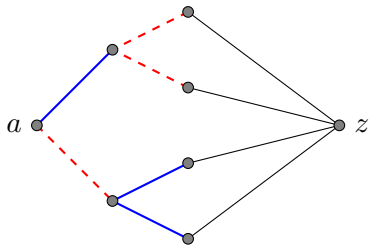

Figure 2.1: In red and blue, two cutsets which are not "one after the other".

Lemma 69. Let θ be a flow from a to z with $\|\theta \geq 0\|$, and let Γ be a cutset separating a from z. Then we have

$$\sum_{e \in \Gamma} |\theta(e)| \geq \|\theta\|.$$

Proof. The proof is very similar to the proof of Lemma 66 above. Let Z be the set of vertices separated from a by Γ, and let $\widetilde{G}$ be the network obtained by merging all the vertices of Z into one vertex $\widetilde{z}$. Then θ induces a flow $\widetilde{\theta}$ on $\widetilde{G}$ with $\|\widetilde{\theta}\| = \|\theta\|$, since the edges going out of a are not affected by going from G to $\widetilde{G}$. Therefore, we have

$$\|\theta\| = \|\widetilde{\theta}\| = \sum_{\{u,\widetilde{z}\} \in \widetilde{E}} \widetilde{\theta}(u\widetilde{z}) = \sum_{e \in \Gamma} \theta(e) \leq \sum_{e \in \Gamma} |\theta(e)|,$$

since each edge $u\widetilde{z}$ comes from an edge of Γ. □

We can now prove the Nash–Williams inequality for real.

True proof of the Nash–Williams inequality. Let θ be any unit flow from a to z. We can write

$$\mathcal{E}(\theta) \geq \sum_{i=1}^{n} \left(\sum_{e \in \Gamma_i} r(e)\theta(e)^2 \right).$$

By the Cauchy–Schwarz inequality and Lemma 69, for any $1 \leq i \leq n$ we can write

$$\left(\sum_{e \in \Gamma_i} r(e)\theta(e)^2 \right) \left(\sum_{e \in \Gamma_i} c(e) \right) \geq \left(\sum_{e \in \Gamma_i} \sqrt{c(e)r(e)\theta(e)^2} \right)^2 = \left(\sum_{e \in \Gamma_i} |\theta(e)| \right)^2 \geq \|\theta\|^2 = 1,$$

so $\sum_{e \in \Gamma_i} r(e)\theta(e)^2 \geq \left(\sum_{e \in \Gamma_i} c(e) \right)^{-1}$. We conclude by summing over i and taking the minimum over θ. □

From here, we immediately deduce a robust criterion for recurrence.

Corollary 70 (Nash–Williams criterion)**.** If an infinite network admits disjoint cutsets $(\Gamma_i)_{i \geq 1}$ such that

$$\sum_{i \geq 1} \left(\sum_{e \in \Gamma_i} c(e) \right)^{-1} = +\infty,$$

then G is recurrent.

Remark 71. For the simple random walk on a graph, i.e. if all conductances are equal to 1, the condition becomes

$$\sum_{i\geq 1} \frac{1}{|\Gamma_i|} = +\infty.$$

Applications in dimensions 1 and 2.

- A first immediate application of the Nash–Williams criterion is the simple random walk on $\mathbb{Z}$. For any $i \geq 1$, we can take

  $$\Gamma_i = \{\{i, i+1\}, \{-i, -i-1\}\}.$$

  These are disjoint cutsets separating 0 from ∞ with $|\Gamma_i| = 2$, so in particular $\sum_{i\geq 1} \frac{1}{\Gamma_i} = +\infty$.

- Let us now move on to $\mathbb{Z}^2$. We define Γ_i as the set of edges between the ball of radius i (for the infinite norm) and its complement. That is, the edges of Γ_i are

  - the edges $\{(x, i), (x, i+1)\}$ for $-i \leq x \leq i$,
  - the edges $\{(x, -i), (x, -i-1)\}$ for $-i \leq x \leq i$,
  - the edges $\{(i, x), (i+1, x)\}$ for $-i \leq x \leq i$,
  - the edges $\{(-i, x), (-i-1, x)\}$ for $-i \leq x \leq i$.

  Then $|\Gamma_i| = 4(2i+1)$, so $\sum_{i\geq 1} \frac{1}{\Gamma_i} = +\infty$ and $\mathbb{Z}^2$ is recurrent. Reproving the recurrence of $\mathbb{Z}^2$ may not look spectacular, but note that this proof is much more robust than proofs using path counting or the Fourier transform. In particular, the argument still works with almost no change if:

  1. we add conductances on the edges (as long as those conductances are bounded by some constant),
  2. we remove edges,
  3. we add new edges, as long as for all i, we add $O(i)$ edges between the "inside" and the "outside" of Γ_i.

Random paths. We will now describe a useful criterion to prove that a network is transient. This criterion uses *random paths*. Note that we do not make any assumption on the random paths that we consider: we do not require them to be random walks or to satisfy some kind of Markov property.

Lemma 72. Let P be a random path from a to z in a finite network. For all $\vec{e} = (x, y) \in \vec{E}$, let

$$\theta_P(\vec{e}) = \mathbb{E}\left[\#\{i \geq 0 \,|\, (P_i, P_{i+1}) = (x, y)\}\right] - \mathbb{E}\left[\#\{i \geq 0 \,|\, (P_i, P_{i+1}) = (y, x)\}\right].$$

Then θ_P is a flow from a to z with $\|\theta\| = 1$.

Proof. For any deterministic path p from a to z, the quantity

$$\#\{i \geq 0 \,|\, (p_i, p_{i+1}) = (x, y) - \#\{i \geq 0 \,|\, (p_i, p_{i+1}) = (y, x)\}$$

is a unit flow from a to z. By linearity, so is its expectation when p is made random. □

Proposition 73. Let P be a random injective path from a to ∞ in an infinite network G. If

$$\sum_{e \in E} r(e) \mathbb{P}\left(P \text{ uses } e\right)^2 < +\infty, \tag{2.6}$$

then G is transient.

Proof. If P is injective, then in particular it uses each nonoriented edge at most once, so for any edge $\vec{e}$ we have

$$|\theta_P(\vec{e})| = \mathbb{P}\left(P \text{ uses } e\right).$$

It follows that

$$\mathcal{E}\left(\theta_P\right) = \sum_{e \in E} r(e) \mathbb{P}\left(P \text{ uses } e\right)^2 < +\infty,$$

so G is transient by Corollary 67. □

Remark 74. When we deal with actual examples, we want to build a random path that makes the sum (2.6) as small as possible. For this, because of the square in the equation, we want the distribution of the random path to be as "well spread" as possible between the edges of the graph. This is good to keep in mind when looking at the applications below.

Example 75. We consider the complete binary tree $\mathbb{T}_2$, i.e. the infinite tree where every vertex, including the root a, has exactly two children. We denote by P the nonbacktracking random walk, i.e. the random path started from a such that, for all $i \geq 0$, the vertex P_{i+1} is chosen uniformly at random among the two children of P_i.

In particular, for all $i \geq 0$, the path P uses exactly one of the 2^{i+1} edges between height i and height $i+1$, and this edge is chosen uniformly at random. It follows that

$$\sum_{e \in E} \mathbb{P}\left(P \text{ uses } e\right)^2 = \sum_{i \geq 0} 2^{i+1} \times \left(\frac{1}{2^{i+1}}\right)^2 = 1,$$

so $\mathbb{T}_2$ is transient.

Again, this example may not seem spectacular: the transience of $\mathbb{T}_2$ can also be proved by arguing that the height of the random walk evolves like a random walk on $\mathbb{Z}$ with positive bias (at each step, it has probability $\frac{2}{3}$ to increase by 1 and probability $\frac{1}{3}$ to decrease by 1).

Let us move on to another example where the construction of the random path is less simple: the three-dimensional square lattice.

Theorem 22. The lattice $\mathbb{Z}^3$ is transient.

Proof. We will build a random path that "follows a straight line" in a uniformly chosen direction. More precisely, let $\vec{u}$ be a vector chosen uniformly at random in the unit sphere of $\mathbb{R}^3$. We define the region $\mathcal{R}_{\vec{u}}$ of points which are "approximately in direction $\vec{u}$ as

$$\mathcal{R}_{\vec{u}} = \{x \in \mathbb{R}^3 \mid d\left(x, \|x\|_2 \vec{u}\right) \leq 10\},$$

where we use the Euclidean distance. In particular, there is an injective path P from 0 to ∞ that stays inside $\mathbb{Z}^3 \cap \mathcal{R}_{\vec{u}}$. We choose P to be an arbitrary such path. Now let e be an edge in

$\mathbb{Z}^3$ started from a vertex x, and let $r = \|x\|_\infty$. Then

$$\begin{aligned}\mathbb{P}(P \text{ uses } e) &\leq \mathbb{P}(d(x, \|x\|_2 \vec{u}) \leq 10) \\ &= \mathbb{P}\left(d\left(\frac{x}{\|x\|_2}, \vec{u}\right) \leq 10\right) \\ &= O\left(\frac{1}{\|x\|^2}\right) \\ &= O\left(\frac{1}{r^2}\right).\end{aligned}$$

Moreover, the number of edges e yielding a certain value of r is $O(r^2)$ (this is where it is more convenient to work with the infinite norm), so

$$\sum_{e \in E} \mathbb{P}(P \text{ uses } e)^2 = O\left(\sum_{r \geq 1} r^2 \times \left(\frac{1}{r^2}\right)^2\right) < +\infty.$$

□

2.1.6 Invariance under quasi-isometries

We have already argued that electric networks are a "robust" method, in the sense that it is not affected by adding or removing edges, as long as these additions do not change the geometry of the graph completely. We conclude this chapter by turning this into a precise and general statement. The notion of "not changing the geometry too much" will be given by *quasi-isometries*. We first note that any graph can be seen as a metric space by considering the *graph distance*, i.e. the length of the shortest path between two vertices.

Definition 76. Let (X_1, d_1) and (X_2, d_2) be two metric spaces, and let $f : X_1 \to X_2$.

- We say that f is a *quasi-isometry* if there are two constants $A, B > 0$ such that:

  (i) for all $x, y \in X_1$, we have

  $$\frac{1}{A} d_1(x, y) - B \leq d_2(f(x), f(y)) \leq A d_1(x, y) + B; \tag{2.7}$$

  (ii) for all $z \in X_2$, there is $x \in X_1$ such that $d_2(f(x), z) \leq B$.

- We say that X_1 and X_2 are *quasi-isometric* if there is a quasi-isometry from X_1 to X_2.

The first condition can be interpreted as saying that f "almost preserves distances", and the second as saying that f is "almost surjective" (a form of almost-injectivity already follows from the first condition).

Remark 77. The relation "being quasi-isometric" is an equivalence relation. Indeed, the identity from (X, d) to itself is a quasi-isometry, and transitivity comes from the fact that the composite of two quasi-isometries is still a quasi-isometry. Finally, if $f : X_1 \to X_2$ is a quasi-isometry and, for any $z \in X_2$, we choose $g(z) \in X_1$ such that $d_2(f(g(z)), z) \leq B$, then we can check that $g : X_2 \to X_1$ is a quasi-isometry.

Example 78. 1. For any $d \geq 1$, the spaces $\mathbb{Z}^d$ and $\mathbb{R}^d$ are quasi-isometric, even though one is discrete and the other continuous (the identity provides a quasi-isometry, for any choice of the norm on $\mathbb{R}^d$).

2. The square and triangular lattices are isometric (for example because they are both quasi-isometric to $\mathbb{R}^2$).

3. For any $\mathrm{d}_1 \neq d_2$, the graphs $\mathbb{Z}^{d_1}$ and $\mathbb{Z}^{d_2}$ are not quasi-isometric. Indeed, if for example there was a quasi-isometry $f : \mathbb{Z}^3 \to \mathbb{Z}^2$, we could assume $f(0) = 0$. Then any point of the ball $B_r(0)$ in $\mathbb{Z}^3$ would be sent to a point of $B_{Ar+B}(0)$ in $\mathbb{Z}^2$. Since the first ball has volume $\approx r^3$ and the second has volume $(Ar + B)^2 \approx r^2$, we could find $\approx r$ points of $\mathbb{Z}^3$ with the same image. Since a ball of radius s has volume of order s^3, we could find two points $x, y \in \mathbb{Z}^3$ with $d(x, y)$ of order at least $r^{1/3}$ such that $f(x) = f(y)$, which violates (2.7).

Theorem 23. Let (G, c) and (G', c') be two electric networks with $G = (V, e)$ and $G' = (V', E')$. We assume that:

1. G and G' are quasi-isometric,

2. there is a constant $C > 0$ such that

$$\forall e \in E,\ K^{-1} \leq c(e) \leq K \quad \text{and} \quad \forall e' \in E,\ K^{-1} \leq c'(e') \leq K,$$

3. there is a constant K such that

$$\forall v \in V,\ \deg(v) \leq K \quad \text{and} \quad \forall v' \in V',\ \deg(v') \leq K.$$

Then (G, c) is transient if and only if (G', c') is.

Note that the assumption on conductances does not cost much, as changing conductances within a bounded window only affects effective resistances by constant factors.

Proof. Since the assumptions are symetric (recall that quasi-isometry is an equivalence relation!), it is sufficient to prove that if (G, c) is transient, then so is (G', c'). For this, we use the characterization using flows with finite energy: we will argue that if G admits a flow with finite energy, so does G'.

Now fix $a \in V$. Let θ be a unit flow from a to ∞ in G, and let $f : G \to G'$ be a quasi-isometry. If $\vec{e} = (x, y)$ is an edge of G, then (2.7) implies $d(f(x), f(y)) \leq A + B$, so let $\gamma_{\vec{e}}$ be an injective path from $f(x)$ to $f(y)$ in G' with length bounded by $A + B$. We choose the paths $\gamma_{\vec{e}}$ in such a way that for any edge $\vec{e} = (x, y)$, the path $\gamma_{(y,x)}$ is the reversal of the path $\gamma_{(x,y)}$. We now define a function θ' on the set $\vec{E}'$ of oriented edges of G' by

$$\theta'(\vec{e}') = \sum_{\vec{e} \in \vec{E}} \theta(\vec{e}) \mathbb{1}_{\vec{e}' \in \gamma_{\vec{e}}}.$$

Roughly speaking, the idea behind this formula is to replace, for every edge $\vec{e} \in \vec{E}$, the current $\theta(\vec{e})$ through $\vec{e}$ by a current $\theta(\vec{e})$ through each edge of $\gamma_{\vec{e}}$.

We now check that θ is a unit flow from $f(a)$ to ∞ in G'. Let us fix $u' \in V'$. We can write

$$\sum_{(u',v') \in \vec{E}'} \theta(u'v') = \sum_{(u,v) \in \vec{E}} \theta(uv) \sum_{(u',v') \in \vec{E}'} \mathbb{1}_{(u',v') \in \gamma_{(u,v)}} = \sum_{(u,v) \in \vec{E}} \theta(uv) \mathbb{1}_{u' \in \gamma_{(u,v)}} \mathbb{1}_{u' \neq f(v)}.$$

We now notice that the contributions of the oriented edge (v,u) is almost the opposite to that of (u,v), except that the condition $u' \neq f(v)$ becomes $u' \neq f(u)$. We obtain

$$\begin{aligned}
2\sum_{(u',v')\in\vec{E}'}\theta(u'v') &= \sum_{(u,v)\in\vec{E}}\theta(uv)\mathbb{1}_{u'\in\gamma_{(u,v)}}\mathbb{1}_{u'\neq f(v)} + \sum_{(u,v)\in\vec{E}}\theta(vu)\mathbb{1}_{u'\in\gamma_{(v,u)}}\mathbb{1}_{u'\neq f(u)} \\
&= \sum_{(u,v)\in\vec{E}}\theta(uv)\mathbb{1}_{u'\in\gamma_{(u,v)}}\left(\mathbb{1}_{u'=f(u)} - \mathbb{1}_{u'=f(v)}\right) \\
&= \sum_{u\in f^{-1}(u')}\sum_{(u,v)\in\vec{E}}\theta(uv) - \sum_{v\in f^{-1}(u')}\sum_{(u,v)\in\vec{E}}\theta(uv).
\end{aligned}$$

If $u' \neq f(a)$, then all the vertices u in the first sum and v in the second are distinct from a, so the sum is 0 because θ is a flow. This means that θ is indeed a flow from $f(a)$ to ∞. Finally, for $u' = f(a)$, the contributions from $u \neq a$ in the first sum and from $v \neq a$ in the second are also 0, so we get

$$2|\theta'\| = 2\sum_{(f(a),v')\in\vec{E}'}\theta(f(a)v') = \sum_{(a,v)\in\vec{E}}\theta(av) - \sum_{(u,a)\in\vec{E}}\theta(ua) = 2\|\theta\| = 2,$$

so θ is indeed a unit flow.

We now estimate the energy of θ'. The difficulty here will be to bound the number of paths $\gamma_{\vec{e}}$ that can use the same edge $\vec{e}'$ of G'. For this, we fix $\vec{e}' \in \vec{E}'$, and a vertex x_0 of G such that $d(f(x_0),\vec{e}') \leq B$ (such a vertex exists by definition of f). If $\vec{e}' \in \gamma_{\vec{e}}$ for some $\vec{e} = (x_1,y_1) \in \vec{E}$, then $\vec{e}'$ belongs to a path of length at most $A+B$ started from $f(x_1)$, so $d\left(f(x_1),\vec{e}'\right) \leq A+B$. By the triangle inequality, we get

$$A + 2B \geq d\left(f(x_0),f(x_1)\right) \geq \frac{1}{A}d\left(x_0,x_1\right) - B,$$

where the second inequality comes from (2.7), so $d(x_0,x_1) \leq A(A+3B)$. Moreover, since vertex degrees in G are bounded by K, any ball of radius r contains at most K^r vertices. Therefore, there are at most $K^{A(A+3B)}$ possible choices for x_1. Since each possible x_1 is incident to at most K edges, there are at most $K' := K^{A(A+3B)+1}$ oriented edges $\vec{e}$ such that $\vec{e}' \in \gamma_{\vec{e}}$. We can now finish the energy computation. By bounding all the resistances $r'(e')$ by C, we find

$$\begin{aligned}
\mathcal{E}(\theta') &\leq C\sum_{\vec{e}'\in\vec{E}'}\theta'(\vec{e}')^2 \\
&= C\sum_{\vec{e}'\in\vec{E}'}\left(\sum_{\vec{e}\in\vec{E}}\theta(\vec{e})\mathbb{1}_{\vec{e}'\in\gamma_{\vec{e}}}\times\mathbb{1}_{\vec{e}'\in\gamma_{\vec{e}}}\right)^2 \\
&\leq C\sum_{\vec{e}'\in\vec{E}'}\left(\sum_{\vec{e}\in\vec{E}}\theta(\vec{e})^2\mathbb{1}_{\vec{e}'\in\gamma_{\vec{e}}}\right)\left(\sum_{\vec{e}\in\vec{E}}\mathbb{1}_{\vec{e}'\in\gamma_{\vec{e}}}\right) \\
&\leq CK'\sum_{\vec{e}'\in\vec{E}'}\sum_{\vec{e}\in\vec{E}}\mathbb{1}_{\vec{e}'\in\gamma_{\vec{e}}}\theta(\vec{e})^2 \\
&\leq CK'\times(A+B)\sum_{\vec{e}\in\vec{E}}\theta(\vec{e})^2 \\
&\leq 2C^2K'(A+B)\mathcal{E}(\theta),
\end{aligned}$$

where we have used the Cauchy–Schwarz inequality, and in the end that the resistances $r(e)$ are bounded from below by C^{-1}. In particular, if θ has finite energy, so does θ'. Therefore, if G is transient, then G' is transient as well. □

2.2 Spectral radius and isoperimetric inequalities

Let $G = (V, E)$ be an infinite, locally finite connected graph. We denote by (S_n) the simple random walk on G, i.e. the Markov chain with transitions

$$p(x, y) = \mathbb{P}_x (S_1 = y) = \frac{1}{\deg(x)} \mathbb{1}_{\{x,y\}\in E}.$$

Finally, for $n \geq 0$, we write

$$p_n(x, y) = \mathbb{P}_x (S_n = y) .$$

Let $a \in V$. Our goal in this section is to describe a link between geometric properties of G and the asymptotic behaviour of $p_n(a, a)$ as $n \to +\infty$. We will now define the geometric property that we will be interested in.

Cheeger constant. For any finite subset S of V, we define its *volume*

$$|S| = \sum_{x\in S} \deg(x).$$

We also define its *perimeter*

$$|\partial S| = \#\{\text{edges with one endpoint inside } S \text{ and one endpoint outside}\} .$$

Definition 79. The *Cheeger constant* of G is given by

$$h(G) = \inf \left\{ \frac{|\partial S|}{|S|} \mid S \subset V \text{ finite} \right\} .$$

In other words, this is the largest constant h such that $|\partial S| \geq h|S|$ for every finite $S \subset V$.

Example 80.

- For any $d \geq 1$, we have $h(\mathbb{Z}^d) = 0$. Indeed, if we choose a ball of radius r for S, then $|S|$ is of order r^d but $|\partial S|$ is of order r^{d-1}, so the ratio $\frac{|\partial S|}{|S|}$ can be made arbitrarily small.

- The simplest example of a graph with positive Cheeger constant is probably the complete binary tree $\mathbb{T}$ where every vertex has 3 neighbours. Indeed, if V is connected and has n vertices, then we immediately have $|S| = 3n$, and we can check by induction on n that $|\partial S|$ only depends on n and is equal to $n + 2$ (indeed, adding one vertex removes 1 edge from the boundary but creates two new ones). Moreover, disconnecting S only increases $|\partial S|$, so

$$h(\mathbb{T}) = \inf_{n\geq 1} \frac{n + 2}{3n} = \frac{1}{3} > 0.$$

  Similarly, the infinite d-regular tree has Cheeger constant $\frac{d-2}{d}$.

- Another example of "regular" graph with positive Cheeger constant is the 7-regular hyperbolic triangulation of the plane, which can be seen as the analog in hyperbolic geometry of the 6-regular triangular lattice (see Figure 2.2).

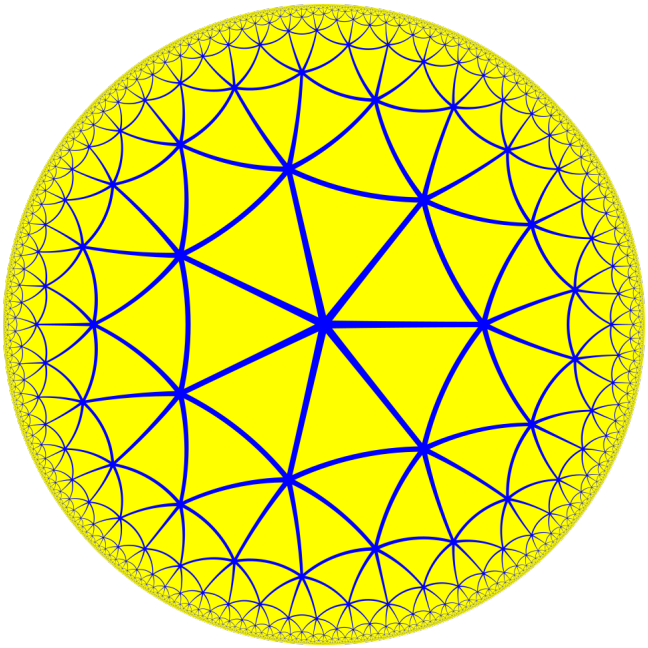

Figure 2.2: The 7-regular triangulation of the plane.

Bibliography

[1] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[2] I. Kortchemski. Arbres et marches aléatoires. *Journées X-UPS*, 2016.

[3] G. F. Lawler and V. Limic. *Random Walk: A Modern Introduction*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2010.

[4] J.-F. Le Gall. Random trees and applications. *Probability Surveys [electronic only]*, 2:245–311, 2005.

[5] L. A. Shepp. Recurrent random walks with arbitrarily large steps. *Bulletin of the American Mathematical Society*, 70(4):540 – 542, 1964.

[6] F. Spitzer. *Principles of Random Walk*. Graduate texts in Mathematics. Springer, 1964.