Kac's Random Walk and Coupon Collector's Process on Posets

by

Sergiy Sidenko

B.S., Moscow Institute of Physics and Technology, 2004

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2008

Author ....
Department of Mathematics
April 24, 2008

Certified by ....
Igor Pak
Associate Professor of Applied Mathematics
Thesis Supervisor

Accepted by ....
Alar Toomre
Chairman, Applied Mathematics Committee

Accepted by ....
David Jerison
Chairman, Department Committee on Graduate Students

Kac's Random Walk and Coupon Collector's

Process on Posets

by

Sergiy Sidenko

Submitted to the Department of Mathematics
on April 24, 2008, in partial fulfillment of the
requirements for the degree of
DOCTOR OF PHILOSOPHY

Abstract

In the first part of this work, we study a long standing open problem on the mixing time of Kac's random walk on $SO(n, \mathbb{R})$ by random rotations. We obtain an upper bound mix $= O\left(n^{2.5}\log n\right)$ for the weak convergence which is close to the trivial lower bound $\Omega\left(n^2\right)$. This improves the upper bound $O\left(n^4 \log n\right)$ by Diaconis and Saloff-Coste [13]. The proof is a variation on the coupling technique we develop to bound the mixing time for compact Markov chains, which is of independent interest.

In the second part, we consider a generalization of the coupon collector's problem in which coupons are allowed to be collected according to a partial order. Along with the discrete process, we also study the Poisson version which admits a tractable parametrization. This allows us to prove convexity of the expected completion time $\mathbb{E}\,\tau$ with respect to sample probabilities, which has been an open question for the standard coupon collector's problem. Since the exact computation of $\mathbb{E}\,\tau$ is formidable, we use convexity to establish the upper and the lower bound (these bounds differ by a log factor). We refine these bounds for special classes of posets. For instance, we show the cut-off phenomenon for shallow posets that are closely connected to the classical Dixie Cup problem. We also prove the linear growth of the expectation for posets whose number of chains grows at most exponentially with respect to the maximal length of a chain. Examples of these posets are d-dimensional grids, for which the Poisson process is usually referred as the last passage percolation problem. In addition, the coupon collector's process on a poset can be used to generate a random linear extension. We show that for forests of rooted directed trees it is possible to assign sample probabilities such that the induced distribution over all linear extensions will be uniform. Finally, we show the connection of the process with some combinatorial properties of posets.

Thesis Supervisor: Igor Pak
Title: Associate Professor of Applied Mathematics

Acknowledgments

I am grateful to Richard Dudley, Alex Postnikov, Peter Shor, Richard Stanley, and Dan Stroock for the insightful remarks and helpful conversations on the subject. Special thanks to Ravi Montenegro and Yuval Peres for finding a critical error in the early version of this paper, and to Igor Pak for having made my work as interesting and unforgettable as was possible.

Contents

Chapter 1

Introduction

1.1 Kac's random walk

The MCMC (Monte Carlo Markov Chain) method has proved extremely powerful and led to remarkable advances in both theory and practice. Despite a large body of literature, finding sharp bounds on the mixing time of finite Markov chains remains technical and exceedingly difficult (see e.g. [5, 26, 36, 43, 44]).

In this paper we study the classical Kac's random walk on $SO(n,\mathbb{R})$ by random rotations in the basis 2-dimensional planes. Our main result is the $O(n^{2.5}\log n)$ mixing time upper bound. This is sharper than previous results and within striking distance from the trivial $\Omega(n^2)$ lower bound. This random walk arose in Kac's effort to simplify Boltzmann's proof of the H-theorem [29] (see also [35]) and Hastings's simulations of random rotations [20]. Most recently, this walk has appeared in the work of Ailon and Chazelle [3] in connection with generating random projections onto subspaces.

Kac's random walk was first rigorously studied by Diaconis and Saloff-Coste [13] who viewed it as a natural example of a Glauber dynamics on $SO(n,\mathbb{R})$. They used a modified comparison technique and proved $O(n^4\log n)$ upper bound on the mixing time, by reducing the problem to a problem of a random walk with *all* rotations, which was solved earlier by using the character estimates in [42] (see also [40]).

In the wake of the pioneer work [13], there has been a flurry of activity in the subject, aimed especially at finding sharp bounds for the eigenvalues [10, 25, 49].

In [34] Maslin was able to explicitly compute the eigenvalues, but due to the large multiplicity of the highest eigenvalue this work does not improve the mixing time of Kac's random walk, in effect showing the limitations of this approach.

It is worth mentioning that there are several notions of the mixing time in this case, and in contrast with the discrete case, the connections between them are yet to be completely understood. We use here the mixing time in terms of the Lipschitz metric for weak convergence, the same as used by Diaconis and Saloff-Coste in [13].

Our approach is based on the coupling technique, a probabilistic approach which goes back to Doeblin [14] (see also [4, 30, 39]). In recent years, this technique has been further adapted to finite Markov chains, largely on combinatorial objects (see "path coupling" technique in [9, 16]). In this paper we adapt the coupling technique to compact Markov chains. While coupling on compact groups has been studied earlier (see e.g. [30] and references therein), our approach is more general, as we define the stopping time to be the first time when two random walks are at a certain distance from each other.

Let us emphasize that while natural in the context, the continuous coupling we consider does not seem to have known analogues in the context of finite Markov chains (cf. [5, 44]). In conclusion, let us mention that a related approach is used in [21], which develops a "distance-decreasing" (with good probability) coupling technique. While the particular coupling construction we employ can also be viewed as distance-decreasing in a certain precise sense (for the Frobenius distance on matrices), our setting is more general.

1.2 Coupon collector's process on posets

The classical coupon collector's problem (CCP) is defined as a discrete process of sampling n distinct coupons. At each step we draw a coupon, and the probability of drawing coupon i is p_i (usually $p_i = \frac{1}{n}$). One of the most important quantities of this process is the number τ of coupons needed to be drawn in order to complete the whole collection, i.e. to sample each coupon at least once. The CCP has a huge literature. It is a natural framework for many combinatorial questions (e.g. [18, 28]), randomized algorithms ([37]) and other numerous applications ([7]).

We consider a generalization of the coupon collector's problem when we are allowed to collect coupons, that is to add a coupon to the collection, according to some partial order on them. In other words, the collection can be represented as a partially ordered set $\mathcal{P}$ on n elements, and we are allowed to collect an element i only if all elements smaller than i in $\mathcal{P}$ have been already collected.

Recall that a partially ordered set $\mathcal{P} = \{P, \prec\}$ (or *poset*, for short) is a formal pair of a set P and a binary relation $\preceq$ (or $\preceq_{\mathcal{P}}$ to avoid confusion) on elements of P which satisfies the following three properties:

1. *reflexivity*: for all $x \in P$, $x \preceq x$,
2. *antisymmetry*: if $x \preceq y$ and $y \preceq x$, then $x = y$,
3. *transitivity*: if $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

We will also write $x \succeq y$ to mean $y \preceq x$, $x \prec y$ to mean $x \preceq y$ and $x \neq y$, and $x \succ y$ to mean $y \prec x$. We will say that two elements x and y are *comparable* if either $x \preceq y$ or $y \preceq x$, and *incomparable* otherwise. As a slight abuse of notation, we will not distinguish a poset $\mathcal{P}$ from its set of elements P, and we will use $x \in \mathcal{P}$ to mean $x \in P$.

Example 1.2.1. Let n be a positive integer. Then the set $1, 2, \ldots, n$ with its usual order forms a poset. We will denote this poset by $\mathcal{S}_n$. Note that any two elements in $\mathcal{S}_n$ are comparable. We will say that $\mathcal{S}_n$ is *totally ordered.* An opposite poset to $\mathcal{S}_n$ is a poset on n incomparable elements. We will denote it by $\mathcal{I}_n$.

Obviously, in the case $\mathcal{P} = \mathcal{I}_n$, we have the classical CCP, since any coupon can be collected when it is drawn.

Along with this random process, we also study its continuous version in which each coupon i arrives as a Poisson process with rate p_i, and all these Poisson processes are independent. We will refer to these two versions as *discrete* and *continuous* CCP on posets. For the classical problem, connections between both processes were studied in [22]. The following lemma shows that the two versions for the same poset are also closely related (all statements on the CCP given in this section will be proved in Chapter 3).

Lemma 1.2.2. *Let τ_C and τ_D denote stopping times of collecting all elements in poset $\mathcal{P}$ for the continuous and discrete versions respectively. Then*

$$e^t \mathbf{P}(\tau_C > t) = \sum_{m=0}^{\infty} \frac{t^m}{m!} \mathbf{P}(\tau_D > m). \tag{1.1}$$

In particular,

$$\mathbb{E}\,\tau_C = \mathbb{E}\,\tau_D. \tag{1.2}$$

For relations between the higher moments, see Corollary 3.4.1.

The usual assumption for the CCP is $\sum_{i=1}^n p_i = 1$. However, it is more convenient to assume that

$$\sum_{i=1}^n p_i = p \leq 1,$$

and if $p < 1$, at each step we do nothing with probability $1 - p$.

We will consider a few simple general properties of the discrete and continuous processes in Section 3.2. Namely, by comparing processes on different posets and for different sample probabilities vectors, we will establish stochastic dominance results (Propositions 3.2.1 and 3.2.3), which, for instance, yield that $\mathbb{E}\,\tau$ is a decreasing function of sample probabilities. We will also provide a recursion for $\mathbb{E}\,\tau$ (Theorem 3.2.6). This gives a representation of $\mathbb{E}\,\tau$ in terms of rational functions (Corollary 3.2.7).

Intuitively, the tail probability $\mathbf{P}(\tau > t)$ of the coupon collector's process should decay exponentially. As the following theorem shows, the rate of the decay depends only on the minimal sample probability.

Theorem 1.2.3. *Let $p_{\min}$ denote the minimal sample probability, i.e.*

$$p_{\min} = \min_{1 \leq i \leq n} p_i.$$

Then

$$\lim_{m\to\infty} \frac{1}{m} \log \mathbf{P}(\tau_D > m) = \log(1 - p_{\min}), \tag{1.3}$$

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbf{P}(\tau_C > t) = -p_{\min}, \tag{1.4}$$

for any poset $\mathcal{P}$ on n elements.

A similar result was proved in [33, 45] for the classical coupon collector problem. The proof of Theorem 1.2.3 involves the reduction of the discrete coupon collector's process to a Markov chain (for more details, see Corollary 3.3.2). However, it seems difficult to use this approach to prove other properties of CCP.

Since the expectation $\mathbb{E}\,\tau$ is a function of a sample probability vector

$$\mathbf{p} = (p_1, \ldots, p_n)^T$$

and the poset $\mathcal{P}$, this function might be used for exploring the structure of the poset $\mathcal{P}$ itself. Unfortunately, we cannot reconstruct the poset exactly by using the corresponding $\mathbb{E}\,\tau$. Indeed, if $\mathcal{P} = \mathcal{S}_n$ with sample probabilities $p_1, \ldots, p_n$, then

$$\mathbb{E}\,\tau = \frac{1}{p_1} + \ldots + \frac{1}{p_n}$$

for any permutation of the n elements. However, knowing $\mathbb{E}\,\tau$ we can determine for any two elements whether they are comparable in the poset.

We need a few definitions to state the result. For a poset $\mathcal{P}$ we will call a set s of elements $i_1, i_2, \ldots, i_k$ a *chain*, if

$$i_1 \prec i_2 \prec \ldots \prec i_k.$$

If any two of these elements are incomparable, we will call s an *antichain*. Denote by $\mathrm{Ch}(\mathcal{P})$ and $\mathrm{ACh}(\mathcal{P})$ be families of chains and antichains of $\mathcal{P}$ respectively. For any subset s of elements of $\mathcal{P}$, let

$$\boldsymbol{\chi}(s) = (\chi_1, \ldots, \chi_n)^T$$

be the *indicator vector* of s such that $\chi_i = 1$ if element i belongs to s and $\chi_i = 0$ otherwise. For a poset $\mathcal{P}$ on n elements, consider the polytope $\mathcal{C}(\mathcal{P})$ in $\mathbb{R}^n$ given by the following inequalities:

$$x_i \geq 0 \quad \text{for all} \quad 1 \leq i \leq n, \tag{1.5}$$

$$\chi(c) \cdot \mathbf{x} \leq 1 \quad \text{for all chains} \quad c \in \mathrm{Ch}(\mathcal{P}), \tag{1.6}$$

where $\mathbf{x} = (x_1, \ldots, x_n)^T$, and "$\cdot$" stands for the standard scalar product of two vectors. Such a polytope was first considered in [48], and it is usually called the *chain polytope* of $\mathcal{P}$.

Theorem 1.2.4. *Let $\mathcal{P}$ and $\mathcal{Q}$ be two posets on the same set of n elements, and let k be positive integer. Then*

$$\mathbb{E}\,\tau^k(\mathbf{p}, \mathcal{P}) = \mathbb{E}\,\tau^k(\mathbf{p}, \mathcal{Q}) \tag{1.7}$$

for any $\mathbf{p}$ *if and only if*

$$\mathcal{C}(\mathcal{P}) = \mathcal{C}(\mathcal{Q}).$$

Thus, the function $\mathbb{E}\,\tau^k(\mathbf{p})$ completely determines the random process for τ.

Stanley ([48]) proved that the vertices of the chain polytope are the indicator vectors of the poset's antichains. Hence, relation (1.7) implies that posets $\mathcal{P}$ and $\mathcal{Q}$ have the same antichains.

Let us call a poset $\mathcal{Q}$ the reverse of $\mathcal{P}$, if the poset $\mathcal{Q}$ is defined on the same set of elements, and for any two elements $i, j \in \mathcal{P}$ we have

$$i \prec j \text{ in } \mathcal{Q} \quad \text{if and only if} \quad j \prec i \text{ in } \mathcal{P}.$$

Obviously, if c is a chain in $\mathcal{P}$, then it is a chain in $\mathcal{Q}$. Hence, a poset and its reverse have the same chains, and therefore, their chain polytopes coincide. Thus, we obtain the following result.

Corollary 1.2.5. *If $\mathcal{Q}$ is the reverse of $\mathcal{P}$, then*

$$\mathbb{E}\,\tau^k(\mathbf{p}, \mathcal{P}) = \mathbb{E}\,\tau^k(\mathbf{p}, \mathcal{Q})$$

for any positive integer k.

Along with the sample probabilities, it is convenient to consider inverse sample probabilities $r_i = \frac{1}{p_i}$, and with some abuse of notation we will write

$$\mathbf{r} = (r_1, \ldots, r_n)^T = \mathbf{p}^{-1}.$$

In [11], for the classical CCP it has been suggested that if $\sum_i p_i = 1$ the expectation $\mathbb{E}\,\tau$ is minimized by the uniform case (all p_i are equal). Holst ([23]) stated without proof that for two sample probability vectors $\mathbf{p}$ and $\mathbf{q}$ we have

$$\mathbb{E}\,\tau(\mathbf{p}) \geq \mathbb{E}\,\tau(\mathbf{q})$$

if $\mathbf{p}$ *majorizes* $\mathbf{q}$, that is the vectors $\mathbf{p}$ and $\mathbf{q}$ satisfy the following conditions:

$$\begin{aligned} p_1 &\geq q_1, \\ p_1 + p_2 &\geq q_1 + q_2, \\ &\vdots \\ p_1 + \ldots + p_{n-1} &\geq q_1 + \ldots + q_{n-1}, \\ p_1 + \ldots + p_n &= q_1 + \ldots + q_n. \end{aligned}$$

Such a property of $\mathbb{E}\,\tau$ is called *Schur-convexity*, and it is much weaker than convexity of symmetric functions. The complete proof of Schur-convexity for the classical problem is given, for example, in [45]. In [11] it was checked that $\mathbb{E}\,\tau$ is a convex function of sample probabilities when $n \leq 6$. It turns out that $\mathbb{E}\,\tau$ is convex for any poset. More rigorously, we have the following result.

Theorem 1.2.6. *For any poset $\mathcal{P}$ on n elements, the expected stopping time $\mathbb{E}\,\tau$*

to complete the collection is a convex function of $\mathbf{p}$. *Also,* $\mathbb{E}\,\tau$ *is a convex function of* $\mathbf{r}$. *In addition, for the continuous version, all moments of* τ, *i.e.* $\mathbb{E}\,\tau_C^k$ *for positive integer* k, *are convex functions both of* $\mathbf{p}$ *and* $\mathbf{r}$.

Theorem 1.2.6 can be used to prove a useful upper bound for $\mathbb{E}\,\tau$. Denote by H_m the m-th harmonic number, i.e.

$$H_m = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{m}. \tag{1.8}$$

Theorem 1.2.7. *Let*

$$L := L(\mathbf{p}, \mathcal{P}) = \max_{c \in \mathrm{Ch}(\mathcal{P})} \chi(c) \cdot \mathbf{r}, \tag{1.9}$$

and M *be the maximum size of antichains in* $\mathcal{P}$. *Then*

$$L \le \mathbb{E}\,\tau(\mathbf{p}, \mathcal{P}) \le L \cdot H_M. \tag{1.10}$$

In Theorem 3.9.3 we will give a generalization of (1.10) for higher moments.

It is not difficult to show that the average number of steps in order to collect all elements from a chain $c = (i_1 \prec \ldots \prec i_k)$ is

$$\frac{1}{p_1} + \ldots + \frac{1}{p_k} = r_1 + \ldots + r_k = \chi(c) \cdot \mathbf{r}.$$

Thus, we can say that L is the expected number of steps to collect all elements from the chain c that maximizes $\chi(c) \cdot \mathbf{r}$. Since we have to complete every chain c from the poset in order to collect all coupons, this explains the lower bound in (1.10). The upper bound is similar to the classical CCP with equal sample probabilities. In this case $L = \frac{1}{p}$, where p is the sample probability, and $\mathbb{E}\,\tau = \frac{1}{p} \cdot H_n$.

Since the maximum size of an antichain in a poset does not exceed the number of elements in the poset, the bounds in (1.10) differ by an $O(\log n)$ factor. However, for some posets we can improve the upper bound.

Let

$$\mathcal{P}_i = \{j \in \mathcal{P} \mid j \preceq i \text{ or } i \prec j\}$$

be the set of all elements of $\mathcal{P}$ that are comparable with or equal to i, and let $\#\mathcal{P}_i$ denote the size of $\mathcal{P}_i$. We will denote by $\mathbf{1}$ the vector whose entries are equal to 1.

Theorem 1.2.8. *Let $\mathcal{P}$ be a poset on n elements, such that $\#\mathcal{P}_i \le k$ for all $i \in \mathcal{P}$ and some k. If all sample probabilities are equal to p, and $k < \log n$, then*

$$\frac{1}{p}\log n < \mathbb{E}\,\tau\,(p \cdot \mathbf{1}, \mathcal{P}) < \frac{1}{p}\,(\log n + k \log\log n + 3k)\,. \tag{1.11}$$

Moreover, if we have a sequence of posets $\mathcal{P}^1, \mathcal{P}^2, \ldots$ for which

$$k = o\left(\frac{\log n}{\log\log n}\right),$$

then

$$\sqrt{\operatorname{Var}\tau} = o\,(\mathbb{E}\,\tau)\,, \tag{1.12}$$

and we have a sharp cut-off for this sequence.

The formal definition of a *sharp cut-off* (or a *cut-off*) is the following. We will say that a sequence $\xi_1, \xi_2, \ldots$ of non-negative random variables has a sharp cut-off if for any $\varepsilon > 0$ there is a sequence of intervals

$$(a_1^\varepsilon, b_1^\varepsilon)\,, (a_2^\varepsilon, b_2^\varepsilon)\,, \ldots$$

such that for any integer $k > 0$ we have

$$\mathbf{P}\,(\xi_k \in (a_k^\varepsilon, b_k^\varepsilon)) > 1 - \varepsilon, \quad \text{and} \quad \lim_{k\to\infty} \frac{b_k^\varepsilon - a_k^\varepsilon}{\mathbb{E}\,\xi_k} = 0.$$

Roughly speaking, with high probability random variable ξ_k assumes its values on an interval whose length is small in comparison with the magnitude of $\mathbb{E}\,\xi_k$. Chebyshev's inequality (for instance, [15, 18]) implies that condition (1.12) is sufficient for a cut-off.

If $\mathcal{P}$ consists of m incomparable chains of length k, then the coupon collector's process behaves in the same way as the Dixie Cup problem (completing k identical collections of m elements each) considered by Erdös and Rényi [17]. For $p = \frac{1}{m}$, they proved the following fact.

Theorem 1.2.9 (Erdös, Rényi). *Let $\tau_k(m)$ be the stopping time when we complete k identical collections of m elements each when the sample probabilities are all equal to $\frac{1}{m}$. Then:*

$$\mathbb{E}\,\tau_k(m) = m\left(\log m + (k-1)\log\log m + C - \log(k-1)! + o(1)\right), \tag{1.13}$$

$$\lim_{m\to\infty} \mathbf{P}\left(\tau_k(m) > m\left(\log m + (k-1)\log\log m + x\right)\right) = \exp\left(-\frac{e^{-x}}{(k-1)!}\right), \tag{1.14}$$

where C denotes the Euler constant.

Obviously, in this specific case, expression (1.13) agrees with (1.11), and (1.14) implies (1.12).

The continuous version of CCP has a connection to percolation problems. Let

$$\mathcal{R}\left(n_1, n_2, \ldots, n_k\right)$$

be the poset whose elements are integer vectors $\mathbf{x} = (x_1, x_2, \ldots, x_k)^T$ such that

$$1 \leq x_i \leq n_i, \quad \text{for all} \quad 1 \leq i \leq k,$$

and $\mathbf{y} \preceq \mathbf{x}$ if and only if

$$y_i \leq x_i, \quad \text{for all} \quad 1 \leq i \leq k.$$

Let us consider the continuous process on the poset

$$\mathcal{R}_n = \mathcal{R}\left(a_1 n, a_2 n, \ldots, a_k n\right),$$

where $a_1, \ldots, a_n$ are kept constant and all the sample probabilities equal 1. The quantity τ in the case $k = 2$ is commonly known as the last passage and has been thoroughly studied (see, for instance, [27]).

Inequality (1.10) implies that

$$(a_1 + a_2 + \ldots + a_n)\, n \le \mathbb{E}\,\tau\,(\mathbf{1}, \mathcal{R}_n) \le (a_1 + a_2 + \ldots + a_n)\, n \log n.$$

In Section 3.7 (Corollary 3.7.3) we will prove the following bound on $\mathbb{E}\,\tau$ which gives us the existence of

$$\lim_{n\to\infty} \frac{1}{n}\, \mathbb{E}\,\tau\,(\mathbf{1}, \mathcal{R}_n).$$

We will call the chain c of poset $\mathcal{P}$ *maximal* if there is no chain c' of $\mathcal{P}$ which contains c as a subset. Let also $\mathrm{MCh}\,(\mathcal{P})$ denote the set of all maximal chains of the poset $\mathcal{P}$.

Theorem 1.2.10. *Let ℓ stand for the maximal length of a chain in $\mathcal{P}$. Denote by ξ^* the root of the equation*

$$\xi - \log \xi = 1 + \frac{\log \#\,\mathrm{MCh}\,(\mathcal{P})}{\ell}, \quad \xi \ge 1. \tag{1.15}$$

Then

$$\frac{1}{p} \cdot \ell \le \mathbb{E}\,\tau\,(p \cdot \mathbf{1}, \mathcal{P}) \le \frac{1}{p} \cdot \left(\xi^* \ell + \frac{\xi^*}{\xi^* - 1} \right).$$

Recall that a *linear extension* of a poset $\mathcal{P}$ on n elements is an *order preserving* bijection $f\colon \mathcal{P} \to \mathcal{S}_n$, i.e.

$$i \prec j \text{ in } \mathcal{P} \quad \Longrightarrow \quad f\,(i) \prec f\,(j).$$

In other words, a linear extension is a way to complete the partial order to a total order. Note that the order in which we collect the elements of the poset $\mathcal{P}$ always gives a linear extension of $\mathcal{P}$. Moreover, a sample probability vector $\mathbf{p}$ induces some distribution over the set of linear extensions. Therefore, the coupon collector's process can be used to generate a random linear extension of $\mathcal{P}$.

Theorem 1.2.11. *Let $\mathcal{U}\,(i)$ stand for the set of elements in poset $\mathcal{P}$ that succeed or* *equal i, i.e.*

$$\mathcal{U}(i) = \{j \in \mathcal{P} \mid i \preceq j\}.$$

If $\mathcal{P}$ is a forest of rooted directed trees, then there is a sample probability vector $\mathbf{p}$ *such that the induced distribution over all linear extensions is uniform. A possible choice of vector* $\mathbf{p}$ *is*

$$p_i = \lambda \cdot \#\mathcal{U}(i)$$

for all $i \in \mathcal{P}$, where

$$\lambda = \left(\sum_{i \in \mathcal{P}} \#\mathcal{U}(i)\right)^{-1}.$$

Throughout the proofs, we will see that the continuous CCP has a strong connection with the integral

$$I(t, \mathbf{a}) = a_1 \dots a_n \int_{t\mathcal{C}(\mathcal{P})} \exp(\mathbf{a} \cdot \mathbf{x}) \, dx_1 \dots dx_n,$$

where $\mathcal{P}$ is a poset on elements $s_1, s_2, \dots, s_n$, and $\mathbf{a}$ is an n-dimensional real vector:

$$\mathbf{a} = (a_1, a_2, \dots, a_n)^T.$$

For instance, Lemma 3.4.4 states that

$$\mathbf{P}(\tau_C(\mathbf{p}, \mathcal{P}) < t) = (-1)^n \cdot I(t, -\mathbf{p}).$$

It turns out that the coefficients of the power series expansion of $I(t, \mathbf{a})$ have a combinatorial interpretation. Let $\mathbf{m}$ be an n-dimensional vector

$$\mathbf{m} = (m_1, m_2, \dots, m_n)^T$$

with non-negative integer entries, and let

$$|\mathbf{m}| = m_1 + \dots + m_n.$$

For any such a vector $\mathbf{m}$, let us consider maps $f\colon \mathcal{S}_{|\mathbf{m}|} \to \mathcal{P}$ that satisfy the following property

$$f(i) \prec f(j) \text{ in } \mathcal{P} \quad \Longrightarrow \quad i \prec j \text{ in } \mathcal{S}_{|\mathbf{m}|},$$

and map exactly m_i elements of $\mathcal{S}_{|\mathbf{m}|}$ to s_i for all $1 \leq i \leq n$. Denote the number of these maps by

$$e(\mathbf{m}) = e(m_1, m_2, \ldots, m_n).$$

Theorem 1.2.12. *For any finite poset $\mathcal{P}$ we have*

$$a_1 \ldots a_n \int_{t\mathcal{C}(\mathcal{P})} \exp(\mathbf{a} \cdot \mathbf{x})\, dx_1 \ldots dx_n = \sum_{\mathbf{m}} \frac{t^{|\mathbf{m}|}}{|\mathbf{m}|!} e(\mathbf{m})\, a_1^{m_1} a_2^{m_2} \ldots a_n^{m_n}. \tag{1.16}$$

Here the summation is taken over all vectors $\mathbf{m}$ with positive integer entries.

See also Section 3.10 for combinatorial applications of $e(\mathbf{m})$.

Chapter 2

Convergence of Kac's random walk

2.1 Weak convergence of Kac's random walk on group $SO(n,\mathbb{R})$

2.1.1 Main result

We will consider the following discrete time random walk on $SO(n,\mathbb{R})$. At each step we pick a pair of coordinates (i,j) such that $1 \leq i < j \leq n$, and an angle ϕ uniformly distributed on $[0,2\pi)$. Define an *elementary rotation matrix* $R_{i,j}(\phi)$:

$$R_{\phi}^{i,j} = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & \cos\phi & \dots & -\sin\phi & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & \sin\phi & \dots & \cos\phi & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

which differs from the identity matrix only in four entries with coordinates (i,i), (j,i), (i,j) and (j,j). *Kac's random walk* $\{A_k\}$ on $SO(n,\mathbb{R})$ is then defined as follows:

$$A_{k+1} = R_{i,j}(\phi) A_k, \quad \text{for all} \quad k \geq 0,$$

where $A_0 = I$ is the identity matrix. More generally, we can assume that A_0 is chosen from any fixed initial distribution P^0 on $SO(n,\mathbb{R})$ as the upper bound below remains valid in this case.

We endow the group $SO(n,\mathbb{R})$ with the *Frobenius norm* (also called the *Hilbert-Schmidt norm*), denoted $\|\cdot\|_F$ and defined for any real $n \times n$ matrix $M = (m_{ij})$ as follows:

$$\|M\|_F = \sqrt{\sum_{1\leq i,j\leq n} m_{ij}^2} = \sqrt{\operatorname{Tr}(MM^T)} = \sqrt{\sum_{1\leq i\leq n} |\sigma_i|^2}, \tag{2.1}$$

where σ_i are the singular values of M. This defines the *Frobenius distance* $\|A-B\|_F$ between every two $n \times n$ matrices A and B.

Let $\operatorname{Lip}(K)$ be a set of all real-valued functions on $SO(n,\mathbb{R})$ such that

$$\|f\|_L = \sup_{x\neq y} \frac{|f(x) - f(y)|}{\|x-y\|_F} \leq K.$$

We define the distance $\rho(P,Q)$ for two probability laws P and Q on $SO(n,\mathbb{R})$ as follows:

$$\rho(P,Q) = \sup\left\{\left|\int f\, d(P-Q)\right| : f \in \operatorname{Lip}(1)\right\}.$$

It is well known (see e.g. [15, §11]) that ρ metrizes the weak convergence of probability laws.[1]

Theorem 2.1.1 (Main Theorem). *Let P^t be the distribution of the Kac's random walk after t steps, and let U be the uniform distribution on $SO(n,\mathbb{R})$. Then, for every $\varepsilon > 0$ there exists*

$$t = O\left(n^{2.5} \log \frac{n}{\varepsilon}\right)$$

[1]Instead of condition $\|f\|_L < 1$, it is common ([13, 15]) to bound $\|f\|_L + \|f\|_\infty$ in order to metrize the weak convergence. However, in our case all Lipschitz functions are bounded because $SO(n,\mathbb{R})$ is compact in $\|\cdot\|_F$.

such that

$$\rho\left(P^t, U\right) < \varepsilon.$$

Thus, there exists a positive constant c, for which we have

$$\rho\left(P^t, U\right) < n \exp\left(-\frac{t}{cn^{2.5}}\right).$$

The proof of this theorem uses an explicit construction of a "weak coupling" between P^t and U, which enables us to bound $\rho\left(P^t, U\right) \leq \varepsilon$. Our "coupling lemma" is described in the next subsection. The proof of the theorem is presented in Section 2.1, where it is split into a sequence of lemmas. The latter are mostly technical and are proved in Section 2.3. We conclude with final remarks in Section 2.4.

2.1.2 Coupling lemma

The basic goal of the coupling we construct is to obtain a joint distribution on $SO(n, \mathbb{R}) \times SO(n, \mathbb{R})$ such that its marginals are P^t and U, and the distribution is concentrated near the main diagonal. Lemma 2.1.2 makes this precise in a more general setting.

Let $X \subset V$, where V is a metric space with distance $d(\cdot, \cdot)$. Let x_t and y_t, where $t \geq 0$, be two Markov processes on X. A coupling is a joint process (x'_t, y'_t) on X^2 such that x'_t has the same distribution as x_t and y'_t has the same distribution as y_t. By abuse of notation we will use (x_t, y_t) to denote the coupling.

The following lemma is an easy consequence of the Kantorovich-Rubinstein theorem (see [15, §11.8]).

Lemma 2.1.2 (Coupling lemma). *Let x_t and y_t $(t \geq 0)$ be two discrete time Markov processes on X as above, with distributions P^t and Q^t after t steps, respectively. Suppose there exist $\delta > 0$ and a coupling (x_t, y_t), such that for the stopping time*

$$T_\delta = \min\left\{t \mid d(x_t, y_t) < \delta\right\}$$

the distance $d(x_t, y_t)$ is non-increasing for $t \geq T_\delta$. Then we have

$$\rho\left(P^t, Q^t\right) < \delta + \mathbf{P}\left(T_\delta > t\right) \cdot \sup_{x,y \in X} d\left(x, y\right). \tag{2.2}$$

It is of interest to compare Lemma 2.1.2 with the standard coupling lemma for random processes on a finite set. Let us suppose that X is finite, and let the metric d be *discrete*, i.e.

$$d\left(x, y\right) = \begin{cases} 1, & \text{if } x \neq y, \\ 0, & \text{if } x = y. \end{cases}$$

For $\delta = 0$ inequality (2.2) becomes

$$\rho\left(P^t, Q^t\right) < \mathbf{P}\left(T_\delta > t\right), \tag{2.3}$$

which is a classical coupling result ([30, §1.2]).

2.1.3 Variations on the theme

Let us mention that the distribution P^t of Kac's random walk does not have a density with respect to U. To see this, consider the following example.

Example 2.1.3. Let X be the unit square $[0, 1] \times [0, 1]$. A random walk starts off at a point (x_0, y_0) in X, and at step k we choose the coordinate x or y, each with probability $\frac{1}{2}$, and set it to the random variable u_k. All u_k are independent and have the uniform distribution U on $[0, 1]$. Then it is not difficult to see that the distribution P^k after k steps ($k > 0$) is

$$P^k = \frac{1}{2^k}\left(\delta\left(x_0\right) \times U + U \times \delta\left(y_0\right)\right) + \left(1 - \frac{1}{2^{k-1}}\right) U \times U,$$

where $\delta\left(\cdot\right)$ stands for the one-dimensional δ-distribution.

As we can see, since P^k contains $\delta\left(x_0\right)$ and $\delta\left(y_0\right)$ with tiny strictly positive coefficients, it is not absolutely continuous.

Analogously, observe that after t steps with probability $1/\binom{n}{2}^t$ we rotate in the first two coordinates only, so the distribution after t steps is not absolutely continuous. On the other hand, in [13] Diaconis and Saloff-Coste proved a convergence in the total variation distance with super-exponential mixing time (see also Section 2.4). Namely, they showed that for

$$t = (2n)^{n^2/2} \log \frac{1}{\varepsilon}$$

we have

$$\left\| P^t - U \right\|_{TV} < \varepsilon,$$

or, equivalently

$$\left\| P^t - U \right\|_{TV} < \exp\left(-\frac{t}{(2n)^{n^2/2}} \right).$$

In the positive direction, our main theorem is robust enough to establish convergence for a number of other matrix distances, such as operator, spectral or trace norms. For example, we easily obtain

Corollary 2.1.4. *Let $K > 1$ be a fixed constant. Then for every $\varepsilon > 0$ and $f \in \mathrm{Lip}\,(K)$ there exists*

$$t = O\left(n^{2.5} \log \frac{Kn}{\varepsilon} \right)$$

such that

$$\left| \int f \, d\left(P^t - U \right) \right| < \varepsilon.$$

Proof. The corollary follows from Theorem 2.1.1 and the observation that $f/K \in \mathrm{Lip}\,(1)$. ☐

More generally, let us give a convergence result in another norm $\|\cdot\|$ on $n \times n$ real matrices. Since the norms $\|\cdot\|$ and $\|\cdot\|_F$ are both defined on the same finite-dimensional vector space, there exists a constant $C_n > 0$ (it might depend on n) such that

$$\|x - y\| \geq C_n \|x - y\|_F \qquad \text{for all } \; x, y \in M\left(n, \mathbb{R}\right).$$

Corollary 2.1.5. *If $\|f\|_L = K$ with respect to the norm $\|\cdot\|$, then for $\varepsilon > 0$ there exists*

$$t = O\left(n^{2.5} \log \frac{Kn}{C_n \varepsilon} \right)$$

such that

$$\left| \int f \, d\left(P^t - U\right) \right| < \varepsilon.$$

Therefore, for the norm $\|\cdot\|$ *there exists a positive constant* c*, for which we have*

$$\rho\left(P^t, U\right) < \frac{n}{C_n} \exp\left(-\frac{t}{cn^{2.5}}\right).$$

Proof. For all $x \neq y$ we obtain

$$\frac{|f(x) - f(y)|}{\|x - y\|} \leq \frac{1}{C_n} \frac{|f(x) - f(y)|}{\|x - y\|_F}.$$

Hence, $f \in \mathrm{Lip}\,(K/C_n)$ with respect to the Frobenius norm. Now Corollary 2.1.4 implies the result. $\square$

2.2 The coupling process

In this section, we construct a coupling process which decreases the Frobenius distance between two random processes with sufficiently high probability. Formally, we show that there is a probability measure on $SO(n,\mathbb{R}) \times SO(n,\mathbb{R})$ with marginals P^t and U which is concentrated near the main diagonal.

Consider now Kac's random walks $\{A_k\}$ and $\{B_k\}$, the former having the initial distribution P^0, and the latter being initially uniformly distributed. We will prove that at each step we are able to choose rotations so that the quantity $\|A_k - B_k\|_F$ is non-increasing whereas the marginal distributions of A_k and B_k remain the same as we choose the rotations randomly. Define matrices $Q_k = A_k B_k^T$, which will play a crucial role in our construction. The random walk of Q_k is induced by random walks of A_k and B_k. Indeed, if at the k-th step the matrices A_k and B_k are to be rotated by R_A and R_B respectively, then

$$\begin{aligned} Q_{k+1} &= A_{k+1}B_{k+1}^T \\ &= (R_A A_k)(R_B B_k)^T \\ &= R_A \left(A_k B_k^T\right) R_B^T \\ &= R_A Q_k R_B^T. \end{aligned}$$

In the case of orthogonal matrices, the Frobenius distance can be computed by the following simple lemma.

Lemma 2.2.1. *If A and B are orthogonal $n \times n$ matrices, then*

$$\|A - B\|_F = \sqrt{2n - 2\,\mathrm{Tr}\left(AB^T\right)}. \tag{2.4}$$

Proof. From the definition of the Frobenius norm, we have

$$\begin{aligned} \|A - B\|_F^2 &= \mathrm{Tr}\left((A-B)(A-B)^T\right) \\ &= \mathrm{Tr}\left(AA^T - BA^T - AB^T + BB^T\right). \end{aligned}$$

Since A and B are orthogonal, we get

$$A^T A = B^T B = I,$$

thus,

$$\begin{aligned} \|A - B\|_F^2 &= \operatorname{Tr}(2I) - \operatorname{Tr}\left(BA^T + AB^T\right) \\ &= 2n - 2\operatorname{Tr}\left(AB^T\right), \end{aligned}$$

as desired. □

Lemma 2.2.1 implies that minimizing the Frobenius norm $\|A_k - B_k\|_F$ is equivalent to maximizing the trace $\operatorname{Tr}\left(A_k B_k^T\right) = \operatorname{Tr} Q_k$. Therefore, we have to show that we can increase the trace of Q_k by choosing appropriate rotations.

Let $R_A = R_{i,j}(\alpha)$ be the rotation by an angle α for the coordinate pair (i, j). Choose

$$R_B = R_{i,j}(\beta(\alpha)),$$

where the angle $\beta = \beta(\alpha)$ will be determined by an explicit construction in the proof of Lemma 2.2.2. Let

$$(M)_{ij} = \begin{pmatrix} m_{ii} & m_{ij} \\ m_{ji} & m_{jj} \end{pmatrix}$$

denote the 2×2 minor of the matrix M and let

$$(Q_k)_{ij} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (Q_{k+1})_{ij} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}.$$

Since R_A and R_B do not change other diagonal elements, the change in trace

$$\operatorname{Tr} Q_{k+1} - \operatorname{Tr} Q_k = \operatorname{Tr}(Q_{k+1})_{ij} - \operatorname{Tr}(Q_k)_{ij}$$

is determined solely by the traces of the minors $(Q_k)_{ij}$ and $(Q_{k+1})_{ij}$. The following lemma allows us to derive a coupling process.

Lemma 2.2.2. *For every $\alpha \in [0, 2\pi)$ there exists $\beta = \beta(\alpha)$ such that the following inequality holds:*

$$(a' + d') - (a + d) \geq \frac{1}{4}(b - c)^2. \tag{2.5}$$

Moreover, if α is a random variable with the uniform distribution on $[0, 2\pi)$, then $\beta(\alpha)$ is also distributed uniformly on $[0, 2\pi)$.

The construction of the angle $\beta = \beta(\alpha)$ is the centerpiece of the whole coupling process. We conjecture that in fact it gives a nearly optimal coupling (see Section 2.4).

Note that Lemma 2.2.2 shows that the efficiency of choosing the rotations depends solely on $|b - c|$, i.e. the entries of $Q_k - Q_k^T$. However, the norm of $Q_k - Q_k^T$ can be small whereas the matrix Q_k is far from the identity matrix. The following lemma shows that if at the beginning we are already sufficiently close to the identity matrix I, then the coupling reduces $\|Q_k - I\|_F$ exponentially.

Lemma 2.2.3. *Let $Q_0, Q_1, \ldots$ be a sequence obtained from coupling I and Q_0 by choosing rotations at each step as described above. If $\|Q_0 - I\|_F < 2$, then for $\delta, \epsilon > 0$ and*

$$t \geq \frac{n^2}{2} \log \frac{4}{\delta^2 \epsilon},$$

we have

$$\mathbf{P}\left(\|Q_t - I\|_F \geq \delta\right) < \epsilon.$$

The next lemma allows us to avoid the problem of small $|b - c|$ entries by adding intermediate "target matrices".

Lemma 2.2.4. *For every matrix $Q \in SO(n, \mathbb{R})$ there exists a sequence of orthogonal $n \times n$ matrices $M_0, M_1, \ldots, M_\ell$ which satisfies the following conditions:*

1. $Q = M_0$ *and* $I = M_\ell$;

2. $\|M_m - M_{m+1}\|_F < 1$ *for* $0 \leq m < \ell$;

3. $\ell < \pi\sqrt{n} + 2$.

The coupling for (A_k, B_k) is constructed as follows. Set $M_0 = A_0^T B_0$, $M_\ell = I$. By Lemma 2.2.4, we can construct a sequence of matrices $M_0, M_1, \ldots, M_\ell$ of length $O(\sqrt{n})$. First, couple matrices $A_0 M_1$ and B_0. Namely, at each step, choose the same coordinate pair (i, j) for R_B as for R_A, and use the construction from Lemma 2.2.2 to determine $\beta = \beta(\alpha)$. By Lemma 2.2.2, the trace of $Q_k = A_k M_1 B_k^T$ is non-decreasing as k grows. Make τ steps, where the choice of τ will be made later in such a way that the distance between $A_k M_1$ and B_k becomes smaller than some $\delta < 1$ with high probability (the value of δ will also be specified later). For this stage of the coupling, Lemma 2.2.3 is applicable since the initial distance between $A_0 M_1$ and B_0 is at most 1.

After coupling $A_0 M_1$ and B_0, we get a pair (A_τ, B_τ). If the distance between $A_\tau M_1$ and B_τ is less than δ, we couple $A_\tau M_2$ and B_τ, etc. Each time we change the matrix M_m to M_{m+1}, the distance between $A_k M_{m+1}$ and B_k does not exceed $1+\delta < 2$. Therefore, we can use Lemma 2.2.3 to analyze every stage of the coupling process. Since $M_\ell = I$, once we have finished couplings for all matrices M_m, the distance between A_k and B_k becomes less than δ. From this point on, we always choose the same rotations for A_k and B_k, which ensures that these random walks stay at the same distance.

In total, this gives roughly $O(n^{2.5})$ rotations to make two matrices close enough, roughly $O(n^2)$ rotations per matrix M_m. In fact, it is a bit more to account for the probability of failure at each of the ℓ sequence steps. A rigorous analysis will follow. Here we state the main lemma on the coupling process.

Lemma 2.2.5. *Let*

$$t \geq 2n^{2.5} \log \frac{13\sqrt{n}}{\delta^2 \epsilon}$$

for $0 < \delta < 1$ *and* $\epsilon > 0$. *Then*

$$\mathbf{P}\left(T_\delta > t\right) = \mathbf{P}\left(\|A_t - B_t\|_F \geq \delta\right) < \epsilon.$$

Lemma 2.2.5 easily implies the main theorem.

Proof. (Theorem 2.1.1) Note that for any $x, y \in SO(n, \mathbb{R})$ we have

$$\|x - y\|_F \leq 2\sqrt{n}.$$

Now, using Lemmas 2.1.2 and 2.2.5 for $\delta = \frac{\varepsilon}{3}$, $\epsilon = \frac{\varepsilon}{3\sqrt{n}}$, we get

$$\rho\left(P^t, U\right) < \frac{\varepsilon}{3} + 2\sqrt{n} \cdot \frac{\varepsilon}{3\sqrt{n}} = \varepsilon,$$

for

$$\begin{aligned} t &= 2n^{2.5} \log \frac{13\sqrt{n}}{\left(\frac{\varepsilon}{3}\right)^2 \cdot \frac{\varepsilon}{3\sqrt{n}}} \\ &< 6n^{2.5} \log \frac{8n}{\varepsilon} \\ &= O\left(n^{2.5} \log \frac{n}{\varepsilon}\right), \end{aligned}$$

as desired. □

2.3 Proofs of results

Proof of Lemma 2.1.2.

Let M^t be a joint distribution of P^t and Q^t, and $L = \sup_{x,y \in X} d(x,y)$. For any $f \in \mathrm{Lip}(1)$ we have:

$$\begin{aligned}
\left| \int f \, d\left(P^t - Q^t\right) \right| &= \left| \iint \left(f(x) - f(y)\right) dP^t(x) \, dQ^t(y) \right| \\
&= \left| \iint \left(f(x) - f(y)\right) dM^t(x,y) \right| \\
&\leq \iint \left| f(x) - f(y) \right| dM^t(x,y) \\
&\leq \iint\limits_{d(x,y)<\delta} d(x,y) \, dM^t(x,y) + \iint\limits_{d(x,y)\geq\delta} d(x,y) \, dM^t(x,y) \\
&< \delta \, \mathbf{P}\left(d(x_t, y_t) < \delta\right) + L \, \mathbf{P}\left(d(x_t, y_t) \geq \delta\right) \\
&< \delta + L \cdot \mathbf{P}\left(d(x_t, y_t) \geq \delta\right),
\end{aligned}$$

where the probabilities $\mathbf{P}$ are taken with respect to measure M^t. From definition of the stopping time T_δ and assuming that the distances do not increase after T_δ, we obtain

$$\mathbf{P}\big(d(x_t, y_t) \geq \delta\big) = \mathbf{P}\big(T_\delta > t\big),$$

and the lemma follows. □

Proof of Lemma 2.2.2.

We have

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix}.$$

Therefore,

$$\begin{aligned}
a' + d' &= (a\cos\alpha\cos\beta - c\sin\alpha\cos\beta - b\cos\alpha\sin\beta + d\sin\alpha\sin\beta) \\
&+ (a\sin\alpha\sin\beta + c\cos\alpha\sin\beta + b\sin\alpha\cos\beta + d\cos\alpha\cos\beta) \\
&= (a+d)\cos(\alpha-\beta) + (b-c)\sin(\alpha-\beta)\,.
\end{aligned}$$

Let

$$r = \sqrt{(a+d)^2 + (b-c)^2}.$$

If $r = 0$, then

$$a' + d' = a + d = b - c = 0,$$

and the inequality (2.5) holds for all β. In this case we can take any β uniformly distributed on $[0, 2\pi)$ (for instance, we may choose $\beta = \alpha$).

Suppose now that $r > 0$. Let us define an angle θ so that

$$\cos\theta = \frac{a+d}{r} \quad \text{and} \quad \sin\theta = \frac{b-c}{r}. \tag{2.6}$$

Finally, let

$$\beta = \beta(\alpha) = \alpha - \theta.$$

Then we have

$$\begin{aligned}
a' + d' &= (a+d)\cos\theta + (b-c)\sin\theta \\
&= r\cos\theta\cdot\cos\theta + r\sin\theta\cdot\sin\theta \\
&= r.
\end{aligned}$$

Observe that $\theta = \alpha - \beta$ depends only on Q_k. Therefore, if α has the uniform distribution over $[0, 2\pi)$, then so does β.

Without loss of generality we can assume that $a^2 \geq d^2$. Since

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is a minor of an orthogonal matrix, we get that $|b|$ and $|c|$ are bounded from above by $\sqrt{1-a^2}$. Indeed, all rows and columns of Q_k are of length 1, hence $a^2 + b^2 \leq 1$ and $a^2 + c^2 \leq 1$.

Now we get

$$\begin{aligned} r^2 &= \left(a^2+b^2\right) + \left(c^2+d^2\right) + 2\left(ad - bc\right) \\ &\leq 1 + 1 + 2\left(a^2 + \sqrt{1-a^2}\sqrt{1-a^2}\right) \\ &= 4, \end{aligned}$$

which implies that $r \leq 2$. Finally, we have

$$\begin{aligned} (a' + d') - (a + d) &= r - r\cos\theta \\ &\geq r\left(1 - \cos\theta\right) \cdot \frac{1+\cos\theta}{2} \\ &= \frac{1}{2r} \cdot \left(r \sin\theta\right)^2 \\ &\geq \frac{1}{4}\left(b - c\right)^2, \end{aligned}$$

which completes the proof. □

Frobenius distance via the eigenvalues

Define $S_k = Q_k - Q_k^T$, for all $k \geq 1$. In view of Lemma 2.2.2, we need to estimate the entries of S_k in terms of Q_k. The following result expresses the Frobenius norm $\|S_k\|_F$ via the eigenvalues of Q_k.

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of an orthogonal $n \times n$ matrix Q. The analysis differs slightly when n is even or odd. Let $m = \left\lfloor \frac{n}{2} \right\rfloor$. Recall that if λ is an eigenvalue of Q, then $\bar{\lambda}$ (the complex conjugate of λ) is also an eigenvalue of Q. Therefore, we can order $\lambda_1, \ldots, \lambda_n$ so that for all $1 \le i \le m$ we have $\lambda_{2i} = \bar{\lambda}_{2i-1}$, and let $\lambda_n = 1$ if n is odd. Denote by x_i's and y_i's the real and imaginary parts of the eigenvalues, namely:

$$\lambda_{2i-1} = x_i + y_i\sqrt{-1}, \qquad \lambda_{2i} = x_i - y_i\sqrt{-1}, \qquad \text{where } 1 \le i \le m. \tag{2.7}$$

Lemma 2.3.1. *For any orthogonal $n \times n$ matrix Q, the following holds:*

$$\left\|Q - Q^T\right\|_F^2 = 8 \sum_{k \le m} \left(1 - x_k^2\right). \tag{2.8}$$

In particular, if $\|Q - I\|_F < 2$, then

$$\left\|Q - Q^T\right\|_F^2 > 8 \sum_{k \le m} \left(1 - x_k\right). \tag{2.9}$$

Proof. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be the eigenvectors of Q corresponding to eigenvalues λ_1, $\lambda_2, \ldots, \lambda_n$. Since Q is orthogonal, then

$$Q^T Q = QQ^T = I,$$

and all absolute values of λ_i are 1, in particular, they are not equal to 0. Then for any $1 \le i \le n$ we have

$$\mathbf{v}_i = Q^T Q \mathbf{v}_i = \lambda_i Q^T \mathbf{v}_i.$$

Therefore, vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are eigenvectors of Q^T with eigenvalues $\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_n$. Now we obtain

$$\left(Q - Q^T\right) \mathbf{v}_i = \left(\lambda_i - \bar{\lambda}_i\right) \mathbf{v}_i.$$

Hence, the matrix $S = Q - Q^T$ has eigenvalues

$$\pm 2y_1\sqrt{-1},\ \pm 2y_2\sqrt{-1},\ \ldots,\ \pm 2y_m\sqrt{-1}.$$

Thus, the singular values of S are

$$2\,|y_1|\,,\ 2\,|y_1|\,,\ 2\,|y_2|\,,\ 2\,|y_2|\,,\ \ldots,2\,|y_m|\,,2\,|y_m|\,,$$

if $n = 2m$, and

$$2\,|y_1|\,,\ 2\,|y_1|\,,\ 2\,|y_2|\,,\ 2\,|y_2|\,,\ \ldots,2\,|y_m|\,,2\,|y_m|\,,\ 0,$$

if $n = 2m+1$. In both cases the square of the Frobenius norm of S can be represented as follows:

$$\begin{aligned} \|S_k\|_F^2 &= \sum_{1\le i\le m} 2\,|2y_i|^2 \\ &= 8\sum_{1\le i\le m} y_i^2 \\ &= 8\sum_{1\le i\le m} \left(1 - x_i^2\right), \end{aligned}$$

and using (2.7), we conclude with (2.8).

If now $\|Q - I\|_F < 2$ then relation (2.4) implies

$$\sum_{k\le m} (1 - x_k) = \frac{n - \operatorname{Tr} Q}{2} < 1.$$

Therefore, for all $1 \le k \le m$ we get $x_k > 0$. Hence

$$1 - x_k^2 = (1 - x_k)\,(1 + x_k) > 1 - x_k,$$

and (2.9) follows. □

Proof of Lemma 2.2.3.

For any $n \times n$ orthogonal matrix Q, define

$$\begin{aligned} \eta(Q) &:= \sum_{i \le m} (1 - x_i) \\ &\equiv \frac{n - \operatorname{Tr} Q}{2} \\ &\equiv \frac{1}{4} \|I - Q\|_F^2, \end{aligned} \tag{2.10}$$

and let $\eta_t = \eta(Q_t)$.

Applying inequality (2.5) for the coordinate pair (i, j) chosen at step t, we obtain

$$\begin{aligned} \eta_t - \eta_{t+1} &= \frac{1}{2} \left(\operatorname{Tr} Q_{t+1} - \operatorname{Tr} Q_t \right) \\ &\ge \frac{1}{8} \left(Q_{t,ij} - Q_{t,ji} \right)^2, \end{aligned}$$

where $Q_{t,ij}$ is the ij-th entry of Q_t. Since

$$\|Q_t - I\|_F \le \|Q_0 - I\|_F < 2,$$

we can get the following upper bound on the expected value of η_{t+1} with respect to the choice of (i, j):

$$\begin{aligned} \mathbb{E}\, \eta_{t+1} &\le \eta_t - \frac{1}{8} \mathbb{E} \left(Q_{t,ij} - Q_{t,ji} \right)^2 \\ &\le \eta_t - \frac{1}{8} \cdot \frac{2}{n^2} \sum_{1 \le i,j \le n} \left(Q_{t,ij} - Q_{t,ji} \right)^2 \\ &\le \eta_t - \frac{1}{4n^2} \left\| Q - Q^T \right\|_F^2 \\ &< \eta_t - \frac{8\eta_t}{4n^2} \\ &= \eta_t \left(1 - \frac{2}{n^2} \right). \end{aligned}$$

Induction on t and the fact that $\eta_0 = \frac{1}{2}(n - \operatorname{Tr} Q_0) < 1$ give

$$\begin{aligned} \mathbb{E}\,\eta_t &< \eta_0 \left(1 - \frac{2}{n^2}\right)^t \\ &< \left(1 - \frac{2}{n^2}\right)^t \\ &< \exp\left(-\frac{2t}{n^2}\right). \end{aligned}$$

Now, the Markov inequality implies

$$\begin{aligned} \mathbf{P}\left(\|Q_t - I\|_F \geq \delta\right) &= \mathbf{P}\left(\eta_t \geq \frac{\delta^2}{4}\right) \\ &\leq \frac{4}{\delta^2} \cdot \mathbb{E}\,\eta_t \\ &< \frac{4}{\delta^2} \exp\left(-\frac{2t}{n^2}\right). \end{aligned}$$

Therefore, we have

$$\mathbf{P}\left(\|Q_t - I\|_F \geq \delta\right) < \epsilon \quad \text{for } t \geq \frac{n^2}{2} \log \frac{4}{\delta^2 \epsilon},$$

which concludes the proof. □

Proof of Lemma 2.2.4.

First, consider the case when all eigenvalues of M_0 are different. We will choose all matrices $M_1, \ldots, M_\ell$ so that they have the same eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ as M_0 has. Let

$$\exp\left(\phi_1^m \sqrt{-1}\right), \ldots, \exp\left(\phi_n^m \sqrt{-1}\right)$$

denote the eigenvalues of M_m. Without loss of generality, we can assume that all ϕ_i^m belong to $(-\pi, \pi)$ for all $m \leq \ell$ and $i \leq n$.

Let θ be a positive number to be chosen later, and let the function $f_\theta(x)$ be defined as follows:

$$f_\theta(x) = \begin{cases} x - \theta & \text{if } x > \theta, \\ 0 & \text{if } |x| \leq \theta, \\ x + \theta & \text{if } x < -\theta. \end{cases}$$

Now let

$$\phi_i^{m+1} = f_\theta(\phi_i^m) \qquad \text{for all } i \leq n,\ m < \ell.$$

Note that f_θ is an odd function; therefore, we have $\phi_a^{m+1} = -\phi_b^{m+1}$ if $\phi_a^m = -\phi_b^m$ for some values of a and b. Hence, all corresponding M_m belong to $SO(n, \mathbb{R})$.

Since all matrices M_m have the same eigenvectors, we obtain that the eigenvalues of $M_{m+1}M_m^T$ are the following:

$$\exp\left(\left(\phi_1^{m+1} - \phi_1^m\right)\sqrt{-1}\right), \ldots, \exp\left(\left(\phi_n^{m+1} - \phi_n^m\right)\sqrt{-1}\right).$$

Observe that

$$|f_\theta(x) - x| \leq \theta.$$

Therefore, the real parts of the eigenvalues of $M_{m+1}M_m^T$ are at least $\cos\theta$.

Thus, by Lemma 2.2.1, we have

$$\begin{aligned} \|M_m - M_{m+1}\|_F &= \sqrt{2n - 2\operatorname{Tr} M_{m+1}M_m^T} \\ &\leq \sqrt{2n - 2n\cos\theta} \\ &< \sqrt{2n - 2n\left(1 - \frac{\theta^2}{2}\right)} \\ &= \theta\sqrt{n}. \end{aligned}$$

Let us choose $\theta = 1/\sqrt{n}$. We obtain $\|M_m - M_{m+1}\|_F < 1$. On the other hand, for every $i \leq n$ we get $\phi_i^\ell = 0$ if

$$\ell = \left\lceil \frac{\pi}{\theta} \right\rceil < \pi\sqrt{n} + 1.$$

Indeed, if $\phi_i^\ell > 0$ for some i, then

$$\phi_i^0 \geq \phi_i^\ell + \ell\theta > \pi\sqrt{n} \cdot \frac{1}{\sqrt{n}} = \pi,$$

which contradicts the assumption $\phi_i^0 < \pi$. Analogously, ϕ_i^ℓ cannot be negative. Therefore, all eigenvalues of M_ℓ are equal to 1, i.e. $M_\ell = I$.

In the case when eigenvalues of M_0 are not all different, choose any matrix M_1 within distance 1 from M_0 with distinct eigenvalues. Applying the above construction to M_1, we get the desired sequence of length at most $\pi\sqrt{n} + 2$. This completes the construction. □

Proof of Lemma 2.2.5.

Note that for any auxiliary matrix M_k there is a probability of failing to couple two matrices. Recall that we stop the coupling procedure if we fail at a certain stage.

For each matrix M_k, $1 \leq k \leq \ell$, we make $\tau = \left\lceil \frac{n^2}{2} \log \frac{4\ell}{\delta^2 \epsilon} \right\rceil$ steps. Then, by Lemma 2.2.3, we have

$$\mathbf{P}\left(\|A_\tau M_1 - B_\tau\|_F \geq \delta\right) < \frac{\epsilon}{\ell}.$$

and

$$\mathbf{P}\left(\|A_{k\tau} M_k - B_{k\tau}\|_F \geq \delta \;\middle|\; \left\|A_{(k-1)\tau} M_{k-1} - B_{(k-1)\tau}\right\|_F < \delta\right) < \frac{\epsilon}{\ell},$$

for all $2 \leq k \leq \ell$. Thus, after $\ell\tau$ steps we have:

$$\begin{aligned}
\mathbf{P}\left(\|A_{\ell\tau} - B_{lm\tau}\|_F \geq \delta\right) &< \sum_{k=1}^{\ell} \mathbf{P}\left(\|A_{k\tau} M_k - B_{k\tau}\|_F \geq \delta\right) \\
&< \sum_{k=1}^{\ell} \frac{\epsilon}{\ell} \\
&= \epsilon.
\end{aligned}$$

Note that $\ell < \pi\sqrt{n} + 2$. Hence, we can couple A_0 and B_0 with probability of success at least $1 - \epsilon$ in at most

$$t = 2n^{2.5} \log \frac{13\sqrt{n}}{\delta^2 \epsilon} \text{ steps,}$$

which accomplishes the proof of the lemma. $\square$

2.4 Remarks

Our result can be used to show convergence in $O\left(n^{2.5}\log\frac{n}{\varepsilon}\right)$ steps in the discrepancy metric D defined as follows:

$$D(P,Q)=\sup_{B}|P(B)-Q(B)|,$$

where the supremum is taken over all closed balls B with respect to the Frobenius metric.

It is unclear whether our results can be used to improve the super-exponential upper bound on the mixing time for the ℓ_1-distance in [13]. As far as we know, there is no general result establishing a connection in this case.

The main result in [25] (see also [49]) is the $O(n^3)$ upper bound on the convergence rate of the entropy. This is related to Kac's original question [29]. We hope that our results can be used to prove stronger bounds.

In our coupling construction for technical reasons (to avoid a small eigenvalue problem) we have to choose a $O(\sqrt{n})$ sequence of "target matrices" M_i. While this sequence cannot be easily shortened, we believe that our coupling process is in fact more efficient than our results suggest.

Recall the basic coupling process we constructed with same pair of coordinates and $\beta=\beta(\alpha)$ as in Lemma 2.2.2. We conjecture that in fact this process mixes in $O(n^2\log n)$ time, a result supported by experimental evidence. This would further improve the upper bound of the mixing time of Kac's walk to nearly match with the trivial lower bound.

As noted in [13], the study of Kac's random walk is strongly related to study of random walks on $SU(n,\mathbb{C})$ by random elementary 2-dimensional rotations. In fact, the technique in [13] and of this paper can be directly translated to this case. These walks are closely related and motivated by quantum computing [2, 46], more specifically by quantum random walks [1].

One can easily modify our coupling construction to obtain an $O(n^2\log n)$ upper bound for the corresponding random walk on the unit sphere $S^{n-1}\subset\mathbb{R}^n$. The improvement has to do with a $O(1)$ sequence of required "target matrices" in this case.

A lower bound $\Omega(n \log n)$ was conjectured in [3, 34]. Moreover, Maslin conjectures a sharp cutoff in this case [34]. It is also possible that our basic coupling mixes in $O(n \log n)$, but as of now such a result seems infeasible.

The analogue of Kac's random walk on $SL(n, \mathbb{F}_q)$ was studied in [38] and is shown to mix in $O(n^3 \log n)$ steps. Unfortunately, the stopping time approach in that paper does not seem to translate to the compact case. It is unclear whether the generalized coupling approach can be used to improve Pak's bounds.

It seems that the coupling for random walks on compact sets has rarely been used. Let us mention [41], where an exact coupling was used to bound the mixing time. We hope our version of the technique will lead to further examples.

Chapter 3

Coupon Collector's Process on Posets

3.1 Overview

In this chapter we will consider in detail the coupon collector's process. The discussion is organized as follows. In Section 3.2, we prove a few simple and auxiliary results. In Section 3.3 we will study the tail probabilities. Section 3.4 is devoted to the proofs of Theorems 1.2.4 and 1.2.6. Inequality (1.10) is proved in Section 3.6, and its generalization for higher moments is given in Section 3.9. We consider shallow posets, i.e. the posets in Theorem 1.2.8, and grids $\mathcal{R}_n$ in Section 3.7. Section 3.8 contains the proof of Theorem 3.8. Finally, in Section 3.10, we will consider combinatorial aspects of integrals over chain polytopes.

3.2 Auxiliary properties

3.2.1 Stochastic domination

Let X and Y be real random variables. We will say that X *stochastically dominates* Y and write $X \succeq Y$ (not to be confused with the partial order in posets) if for all $t \in \mathbb{R}$:

$$\mathbf{P}(X > t) \geq \mathbf{P}(Y > t) . \tag{3.1}$$

Proposition 3.2.1. *If, for sample probability vectors* $\mathbf{p} = (p_1, \dots, p_n)^T$ *and* $\mathbf{q} = (q_1, \dots, q_n)^T$*, we have*

$$p_i \leq q_i \quad \textit{for all } i,$$

then, for any poset $\mathcal{P}$ *on* n *elements,*

$$\tau(\mathbf{p}, \mathcal{P}) \succeq \tau(\mathbf{q}, \mathcal{P}).$$

Corollary 3.2.2. *The function* $\mathbb{E}\,\tau^k(\mathbf{p}, \cdot)$ *is decreasing with respect to each* p_i *for any positive integer* k*.*

Proof. In order to prove Proposition 3.2.1, note that we can turn the process with sample probabilities $\mathbf{q}$ into another with probabilities $\mathbf{p}$ as follows. Whenever we sample coupon k under $\mathbf{q}$, we accept it with probability $\frac{p_i}{q_i}$ and reject otherwise. Obviously, now coupon k will be sampled with probability p_k. Observe that under this reduction, if within the process governed by $\mathbf{p}$ we insert coupon k into the collection, we can insert this coupon within the process governed by $\mathbf{q}$. Therefore, we have

$$\mathbf{P}(\tau(\mathbf{q}, \mathcal{P}) < t) \leq \mathbf{P}(\tau(\mathbf{p}, \mathcal{P}) < t).$$

To prove Corollary 3.2.2, note that

$$\begin{aligned}
\mathbb{E}\,\tau_C^k(\mathbf{p}, \mathcal{P}) &= \int_0^\infty k t^{k-1} \mathbf{P}(\tau_C(\mathbf{p}, \mathcal{P}) > t)\, dt \\
&\geq \int_0^\infty k t^{k-1} \mathbf{P}(\tau_C(\mathbf{q}, \mathcal{P}) > t)\, dt \\
&= \mathbb{E}\,\tau_C^k(\mathbf{q}, \mathcal{P})
\end{aligned}$$

for the continuous version, and

$$\begin{aligned}
\mathbb{E}\,\tau_D^k(\mathbf{p}, \mathcal{P}) &= \sum_{m=0}^{\infty} \left[(m+1)^k - m^k\right] \mathbf{P}(\tau_D(\mathbf{p}, \mathcal{P}) > m) \\
&\geq \sum_{m=0}^{\infty} \left[(m+1)^k - m^k\right] \mathbf{P}(\tau_D(\mathbf{q}, \mathcal{P}) > m) \\
&= \mathbb{E}\,\tau_D(\mathbf{q}, \mathcal{P})
\end{aligned}$$

for the discrete process. It is not hard to see that if $p_i < q_i$ for some i, then the inequalities become strict. $\square$

We will also say that poset $\mathcal{Q} = (B, \prec_Q)$ is a *subposet* of poset $\mathcal{P} = (A, \prec_P)$ if $B \subseteq A$, and for any two elements $i, j \in B$ we have

$$i \prec_Q j \quad \text{if} \quad i \prec_P j.$$

Proposition 3.2.3. *If $\mathcal{Q} = (B, \prec_Q)$ is a subposet of $\mathcal{P} = (A, \prec_P)$, then*

$$\tau(\mathbf{p}, \mathcal{P}) \succeq \tau(\mathbf{q}, \mathcal{Q}),$$

provided the sample probabilities of elements in B given by the vector $\mathbf{p}$ coincide with the sample probabilities given by $\mathbf{q}$.

Corollary 3.2.4. *Under the assumptions of Proposition 3.2.3, we have*

$$\mathbb{E}\,\tau^k(\mathbf{p}, \mathcal{P}) \geq \mathbb{E}\,\tau^k(\mathbf{q}, \mathcal{Q})$$

for any positive integer k.

Proof. In order to prove Proposition 3.2.3, it remains to note that if at any moment we can insert coupon k into $\mathcal{P}$, we can insert it into $\mathcal{Q}$. The proof of Corollary 3.2.4 is similar to the proof of Corollary 3.2.2. $\square$

Example 3.2.5. Since any poset $\mathcal{P}$ on n elements is a subposet of $\mathcal{S}_n$ (appropriately ordered), and $\mathcal{I}_n$ is a subposet of $\mathcal{P}_n$, we easily obtain that

$$\tau(\mathbf{p}, \mathcal{I}_n) \preceq \tau(\mathbf{p}, \mathcal{P}) \preceq \tau(\mathbf{p}, \mathcal{S}_n).$$

3.2.2 Recursion

In this section, we will consider the poset $\mathcal{P} = (S, \prec)$ on n elements, whose sample probabilities are given by a vector $\mathbf{p}$. Whenever we deal with posets on subsets of S

with the same partial order $\prec$, we will assume that the sample probability of any element is the same as its sample probability in $\mathcal{P}$ given by $\mathbf{p}$. For the sake of brevity, we will omit $\mathbf{p}$ in notation.

Denote by $\mathcal{P} \setminus \{i\}$ the poset which is obtained from $\mathcal{P}$ by removing the element i. Let us also say that an element i is a *minimal* element of $\mathcal{P}$ if there is no element j in $\mathcal{P}$ such that $j \prec i$. Then we have the following recursion formula for the CCP on posets.

Theorem 3.2.6. *If elements* $1, \ldots, k$ *are all minimal elements of the poset* $\mathcal{P}$*, then*

$$\mathbb{E}\,\tau(\mathcal{P}) = \frac{1}{p_1 + \ldots + p_k} + \sum_{i=1}^{k} \frac{p_i}{p_1 + \ldots + p_k} \mathbb{E}\,\tau(\mathcal{P} \setminus \{i\}). \tag{3.2}$$

Proof. Let τ_i denote the stopping time of picking element i for the first time, and let

$$t = \min_{1 \le i \le k} \tau_i.$$

Then, for the discrete process we have

$$\begin{aligned} \mathbb{E}\,\tau(\mathcal{P}) &= \sum_{i=1}^{k} \Pr(\tau_i = t) \left[\mathbb{E}(\tau_i \mid \tau_i = t) + \mathbb{E}\,\tau(\mathcal{P} \setminus \{i\})\right] \\ &= \sum_{i=1}^{k} \Pr(\tau_i = t)\, \mathbb{E}(\tau_i \mid \tau_i = t) + \sum_{i=1}^{k} \Pr(\tau_i = t)\, \mathbb{E}\,\tau(\mathcal{P} \setminus \{i\}) \\ &= \mathbb{E}\,t + \sum_{i=1}^{k} \Pr(\tau_i = t)\, \mathbb{E}\,\tau(\mathcal{P} \setminus \{i\}). \end{aligned}$$

It is easy to see that

$$\Pr(\tau_i = t) = \frac{p_i}{p_1 + \ldots + p_k}.$$

Also the random variable t has the geometric distribution with parameter

$$p_1 + \ldots + p_k.$$

Thus, its expectation is

$$\mathbb{E}\, t = \frac{1}{p_1 + \ldots + p_k}.$$

The recursion for the continuous process can be established in a similar way. It also follows from Lemma 1.2.2. □

Corollary 3.2.7. *For a poset $\mathcal{P}$, the expected value $\mathbb{E}\,\tau(\mathbf{p}, \mathcal{P})$ is a function of n variables $p_1, \ldots, p_n$ such that*

$$\mathbb{E}\,\tau(\mathbf{p}, \mathcal{P}) = \frac{N(p_1, \ldots, p_n)}{D(p_1, \ldots, p_n)}, \tag{3.3}$$

where N and D are both homogeneous polynomials with positive coefficients and $\deg N = \deg D - 1$. *Moreover,*

$$D(p_1, \ldots, p_n) = \prod_{a \in \mathrm{ACh}(\mathcal{P})} \chi(a) \cdot \mathbf{p}. \tag{3.4}$$

Proof. Identity (3.2) implies that $\mathbb{E}(\mathbf{p}, \mathcal{P})$ is a rational function. Thus, there are polynomials N and D such that (3.3) holds. Also, from (3.2) we have that $\mathbb{E}\,\tau(\mathbf{p}, \mathcal{P})$ is homogeneous with degree of homogeneity equal to (-1). Thus,

$$\deg N = \deg D - 1.$$

It remains to note that relation (3.2) implies that

$$\mathbb{E}\,\tau(\mathbf{p}, \mathcal{P}) = \sum_{A} \prod_{a \in A} \frac{1}{\chi(a) \cdot \mathbf{p}},$$

where the sum is taken over certain sequences A of antichains, and in each sequence A every antichain appears only once. Therefore, we can choose D as in (3.4). □

Although we can give a short formula for D, we have no nice expression for N. In particular, for the classical coupon collector's process we obtain the following

Corollary 3.2.8. *Let elements of the poset $\mathcal{I}_n$ have equal sample probabilities p. Then*

$$\mathbb{E}\,\tau\,(\mathbf{p}, \mathcal{I}_n) = \frac{1}{p} \cdot H_n = \frac{1}{p}\left(1 + \frac{1}{2} + \ldots + \frac{1}{n}\right). \tag{3.5}$$

Proof. Let $E_n = \mathbb{E}\,(\mathbf{p}, \mathcal{I}_n)$. Since $\mathcal{I}_n \setminus \{i\} = \mathcal{I}_{n-1}$ for any element $i \in \mathcal{I}_n$, we get from (3.2) that

$$E_n = \frac{1}{np} + \sum_{i=1}^{n} \frac{p}{np} \cdot E_{n-1} = \frac{1}{np} + \frac{1}{p} \cdot E_{n-1}.$$

Now, using the fact that $E_1 = \frac{1}{p}$, we easily obtain (3.5) by induction. □

For $p = \frac{1}{n}$, this result was first proved in [18, IX.3].

3.3 Tail probability and the Markov chain approach

This section is devoted to the proof of Theorem 1.2.3. First, we will prove the theorem for the posets $\mathcal{S}_n$ and $\mathcal{I}_n$, and then, by using Example 3.2.5, we will extend it to any poset on n elements.

Lemma 3.3.1. *Theorem 1.2.3 holds for posets $\mathcal{S}_n$ and $\mathcal{I}_n$ for any positive integer n.*

Proof. Let us assume that all sample probabilities are different. Thus, without loss of generality we can assume that

$$0 < p_1 < \ldots < p_n < 1.$$

First, we consider the discrete process. Let τ_i denote the stopping time when we collect element i. Then for $\mathcal{I}_n$ we have

$$\mathbf{P}\left(\tau_D\left(\mathcal{I}_n\right) > m\right) = \mathbf{P}\left(\bigcup_{i=1}^{n}\left[\tau_i > m\right]\right),$$

and by the inclusion-exclusion principle

$$\mathbf{P}\left(\tau_D\left(\mathcal{I}_n\right) > m\right) = \sum_{J\in\mathbf{2}^n,\#J>0}(-1)^{\#J}\,\mathbf{P}\left(\bigcap_{i\in J}\left[\tau_i > m\right]\right),$$

where $\mathbf{2}^n$ stands for the family of all subposets of the set $\{1,\ldots,n\}$. It is not hard to see that

$$\mathbf{P}\left(\bigcap_{i\in J}\left[\tau_i > m\right]\right) = \left(1-\sum_{i\in J}p_i\right)^m.$$

Finally, we obtain

$$\mathbf{P}\left(\tau_D\left(\mathcal{I}_n\right) > m\right) = \sum_{J\in\mathbf{2}^n,\#J>0}(-1)^{\#J}\left(1-\sum_{i\in J}p_i\right)^m. \tag{3.6}$$

Thus, we have

$$\begin{aligned}
\lim_{m\to\infty}\frac{1}{m}\log\mathbf{P}\left(\tau_D\left(\mathcal{I}_n\right)>m\right) &= \lim_{m\to\infty}\frac{1}{m}\log\frac{\mathbf{P}\left(\tau_D\left(\mathcal{I}_n\right)>m\right)}{\left(1-p_1\right)^m}+\log\left(1-p_1\right)\\
&= \lim_{m\to\infty}\frac{1}{m}\log\sum_{J\in 2^n}(-1)^{\#J}\left(\frac{1-\sum_{i\in J}p_i}{1-p_1}\right)^m\\
&\quad+\log\left(1-p_1\right).
\end{aligned}$$

Now for any $J\neq\{1\}$ we get

$$\frac{1-\sum_{i\in J}p_i}{1-p_1}<1,$$

hence when m gets large, the sum in the last display tends to 1, and we proved (1.3) for $\mathcal{I}_n$.

Let us proceed with the proof of (1.3) for $\mathcal{S}_n$. By Lemmas 1.2.2 and 3.4.4, it suffices to consider the case when the elements are ordered as follows:

$$1\prec 2\prec\ldots\prec n.$$

In the proof we will consider a corresponding Markov chain.

Let state S_i for $0\leq i\leq n$ represent the event when we have collected i elements. Then, the transition probabilities have to be the following:

$$\begin{aligned}
p\left(S_i\to S_i\right) &= 1-p_{i+1} \quad \text{if } i<n,\\
p\left(S_n\to S_n\right) &= 1,\\
p\left(S_i\to S_{i+1}\right) &= 1-p_{i+1}.
\end{aligned}$$

The initial distribution $\boldsymbol{\pi}^0$ of the Markov process is given by the vector

$$(1,0,\ldots,0)^T,$$

and for any integer $m\geq 0$ we have

$$\boldsymbol{\pi}^m=M^m\boldsymbol{\pi}^0,$$

where the transition probability matrix M is defined as

$$M = \begin{pmatrix} 1-p_1 & 0 & \cdots & 0 & 0 \\ p_1 & 1-p_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1-p_n & 0 \\ 0 & 0 & \cdots & p_n & 1 \end{pmatrix}.$$

In addition,

$$\mathbf{P}\left(\tau_D\left(\mathcal{S}_n\right) \leq m\right) = \left(\boldsymbol{\pi}^m\right)_{n+1},$$

where $(\mathbf{v})_i$ denotes the i-th entry of vector $\mathbf{v}$.

It is not hard to see that the eigenvalues of matrix M are

$$1-p_1,\ 1-p_2,\ \ldots,\ 1-p_n,\ 1,$$

and each has the unique corresponding eigenvector $\mathbf{v}^i$ so that

$$M\mathbf{v}^i = (1-p_i)\,\mathbf{v}^i \quad \text{for all } 1 \leq i \leq n,$$

and

$$M\mathbf{v} = \mathbf{v}.$$

Observe that the vector $\mathbf{v}$ represents the stationary distribution $(0, 0, \ldots, 0, 1)^T$. Thus, decomposing $\boldsymbol{\pi}^0$ in the eigenbasis, we have

$$\boldsymbol{\pi}^m = \mathbf{v} + \sum_{i=1}^{n} \beta_i \mathbf{v}^i \left(1-p_i\right)^m$$

for some $\beta_1, \ldots, \beta_n$.

This implies

$$\begin{aligned}
\mathbf{P}\left(\tau_D\left(\mathcal{S}_n\right) > m\right) &= 1 - \left(\boldsymbol{\pi}^m\right)_{n+1} \\
&= 1 - (\mathbf{v})_{n+1} - \sum_{i=1}^{n} \beta_i \left(\mathbf{v}^i\right)_{n+1} (1-p_i)^m \\
&= -\sum_{i=1}^{n} \beta_i \left(\mathbf{v}^i\right)_{n+1} (1-p_i)^m .
\end{aligned}$$

In order to prove (1.3) for $\mathcal{S}_n$, it remains to show that

$$\beta_1 \cdot \left(\mathbf{v}^1\right)_{n+1} \neq 0.$$

It is not difficult to see that we can choose $\mathbf{v}^1$ as follows:

$$\mathbf{v}^1 = \begin{pmatrix} 1 \\ \frac{p_1}{p_2-p_1} \\ \frac{p_1}{p_2-p_1}\frac{p_2}{p_3-p_2} \\ \vdots \\ \frac{p_1}{p_2-p_1}\cdots\frac{p_{n-1}}{p_n-p_{n-1}}\frac{p_n}{1-p_n} \end{pmatrix}.$$

It is also transparent that the first entries of $\mathbf{v}^i$ for $i > 1$ are equal to 0. Since $\left(\boldsymbol{\pi}^0\right)_1 = 1$, for $\mathbf{v}^1$ chosen as above we have

$$\beta_1 = 1 \quad \text{and} \quad \left(\mathbf{v}^1\right)_{n+1} \neq 0.$$

This finishes the proof of (1.3) for $\mathcal{S}_n$.

In order to prove (1.4), note that in both cases the tail probability can be expressed as follows:

$$\mathbf{P}\left(\tau_D\left(\mathcal{P}\right) > m\right) = \sum_{a\in \mathrm{ACh}(\mathcal{P}),\#a>0} \gamma_a \left(1 - \sum_{i\in a} p_i\right)^m$$

for some γ_a indexed by antichains in the poset, and $\gamma_a \neq 0$ if $a = \{1\}$. From (1.2.2) we have

$$\begin{aligned}
\mathbf{P}\left(\tau_C\left(\mathcal{P}\right)>t\right) &= \sum_{m=0}^{\infty} e^{-t}\frac{t^m}{m!}\sum_{a\in \mathrm{ACh}(\mathcal{P}),\#a>0}\gamma_a\left(1-\sum_{i\in a}p_i\right)^m \\
&= \sum_{a\in \mathrm{ACh}(\mathcal{P}),\#a>0}\gamma_a \exp\left(-t\sum_{i\in a}p_i\right).
\end{aligned}$$

Since the minimal value of $\sum_{i\in a} p_i$ is achieved only on $a=\{1\}$, we have:

$$\frac{P\left(\tau_C\left(\mathcal{P}\right)>t\right)}{e^{-tp_1}} = \sum_{a\in \mathrm{ACh}(\mathcal{P}),\#a>0}\gamma_a \exp\left(-t\frac{\sum_{i\in a}p_i}{p_1}\right)\to 1$$

as $t\to\infty$, which yields (1.4) for both posets.

If more than one sample probabilities are equal to $p_{\min}$, then the lemma follows from Proposition 3.2.1. □

In order to prove Theorem 1.2.3, it is enough to note that any n element poset $\mathcal{P}$ is a subposet of $\mathcal{S}_n$, and $\mathcal{I}_n$ is a subposet of $\mathcal{P}$ (see Example 3.2.5). Thus, by Proposition 3.2.3 we have

$$\mathbf{P}\left(\tau_D\left(\mathcal{I}_n\right)>m\right)\le \mathbf{P}\left(\tau_D\left(\mathcal{P}\right)>m\right)\le \mathbf{P}\left(\tau_D\left(\mathcal{S}_n\right)>m\right),$$

and

$$\mathbf{P}\left(\tau_C\left(\mathcal{I}_n\right)>t\right)\le \mathbf{P}\left(\tau_C\left(\mathcal{P}\right)>t\right)\le \mathbf{P}\left(\tau_C\left(\mathcal{S}_n\right)>t\right),$$

and Theorem 1.2.3 follows from Lemma 3.3.1.

As an outcome of the proof of Lemma 3.3.1, we obtain the following fact.

Corollary 3.3.2. *The discrete coupon collector's process on poset $\mathcal{P}$ can be modeled as a Markov chain whose states correspond to the antichains in $\mathcal{P}$, and each state represents the set of minimal elements that have not been collected yet. For any antichain a, let $S_a(i)$ denote the state after adding element $i\in a$ to the collection.*

Then the transition probabilities of the Markov chain are given as follows:

$$p(S_a \to S_a) = 1 - \sum_{i \in a} p_i,$$

$$p(S_a \to S_a(i)) = p_i \quad \textit{for all } i \in a.$$

The transition matrix of the Markov chain has $\#\,\mathrm{ACh}(\mathcal{P})$ *eigenvalues which are*

$$1 - \sum_{i \in a} p_i \quad \textit{for all } a \in \mathrm{ACh}(\mathcal{P}),$$

and each has a full corresponding eigenspace.

There exist coefficients γ_a*, which are polynomials in* m *such that*

$$\mathbf{P}(\tau_D(\mathcal{P}) > m) = \sum_{a \in \mathrm{ACh}(\mathcal{P}), \# a > 0} \gamma_a \left(1 - \sum_{i \in a} p_i\right)^m,$$

$$\mathbf{P}(\tau_C(\mathcal{P}) > t) = \sum_{a \in \mathrm{ACh}(\mathcal{P}), \# a > 0} \gamma_a \exp\left(-t \sum_{i \in a} p_i\right)$$

for all $t > 0$ *and integer* $m \geq 0$ *(here the sums are taken over all non-empty antichains of* $\mathcal{P}$*). Moreover, the coefficient* γ_a *corresponding to the second largest eigenvalue* $(1 - p_{\min})$ *is positive.*

Remark. If

$$\sum_{i \in a_1} p_i \neq \sum_{i \in a_2} p_i,$$

for any distinct antichains a_1 and a_2, then the coefficients γ_a are independent of m and are some functions of $\mathbf{p}$. The polynomial dependence on m arises when the transition matrix of the corresponding Markov chain has eigenvalues with multiplicity greater than 1.

3.4 Continuous version

In this section, we will consider in detail the continuous CCP on an n element poset $\mathcal{P}$ with sample probabilities assigned by a vector $\mathbf{p} = (p_1, \ldots, p_n)^T$.

3.4.1 Comparison with the discrete version

First, we will prove Lemma 1.2.2, which establishes the connection between the discrete and the continuous version.

Proof. (Lemma 1.2.2) Let

$$p_0 = 1 - \sum_{i=1}^{n} p_i.$$

First, assume that $p_0 = 0$. Denote by $M(t)$ the number of coupons that have arrived (but not necessarily have been collected) by time t in the continuous process. Obviously, the quantity $M(t)$ assumes non-negative integer values. Then we have

$$\mathbf{P}(\tau_C > t) = \sum_{m=0}^{\infty} \mathbf{P}(M(t) = m)\, \mathbf{P}(\tau_C > t \mid M(t) = m).$$

Note that if we condition on the coupons that came prior to time t, the event $\tau_C > t$ will be the same as in the discrete process with the same first coupons. Therefore,

$$\mathbf{P}(\tau_C > t \mid M(t) = m) = \mathbf{P}(\tau_D > m).$$

Also, the arrivals of coupons occur as a Poisson process with rate $\sum_{i=1}^{n} p_i = 1$, hence

$$\mathbf{P}(M(t) = m) = \frac{t^m}{m!} e^{-t}.$$

Combining these facts, we get

$$\mathbf{P}(\tau_C > t) = \sum_{m=0}^{\infty} e^{-t} \frac{t^m}{m!} \mathbf{P}(\tau_D > m).$$

It order to deal with the case $p_0 > 0$, let us add a coupon with sample probability p_0, and ignore it each time it is drawn. Obviously, the relation (1.1) still holds; however, the random process on the original coupons is the same.

To conclude the proof of the lemma, observe that

$$\begin{aligned} \mathbb{E}\,\tau_C &= \int_0^\infty \mathbf{P}\left(\tau_C > t\right) dt \\ &= \sum_{m=0}^{\infty} \mathbf{P}\left(\tau_D > m\right) \left(\int_0^\infty e^{-t} \frac{t^m}{m!} dt\right). \end{aligned}$$

Since for any non-negative integer m

$$\int_0^\infty e^{-t} t^m dt = m!, \tag{3.7}$$

we have

$$\mathbb{E}\,\tau_C = \sum_{m=0}^{\infty} \mathbf{P}\left(\tau_D > m\right) = \mathbb{E}\,\tau_D,$$

as desired. □

Corollary 3.4.1. *For any positive integer k we have*

$$\mathbb{E}\,\tau_C^k = \sum_{m=1}^{k} \begin{bmatrix} k \\ m \end{bmatrix} \mathbb{E}\,\tau_D^k, \tag{3.8}$$

$$\mathbb{E}\,\tau_D^k = \sum_{m=1}^{k} (-1)^{k-m} \begin{Bmatrix} k \\ m \end{Bmatrix} \mathbb{E}\,\tau_C^k, \tag{3.9}$$

where $\left[{k \atop m}\right]$ and $\left\{{k \atop m}\right\}$ stand for the Stirling numbers of the first and the second kind, respectively.

Proof. We have

$$\begin{aligned} \mathbb{E}\,\tau_C^k &= \int_0^\infty k t^{k-1} \mathbf{P}\left(\tau_C > t\right) dt \\ &= \sum_{m=0}^{\infty} \mathbf{P}\left(\tau_D > m\right) \left(\int_0^\infty k t^{k-1} e^{-t} \frac{t^m}{m!} dt\right). \end{aligned}$$

Relation (3.7) yields

$$\begin{aligned}\mathbb{E}\,\tau_C^k &= \sum_{m=0}^{\infty} \mathbf{P}\left(\tau_D > m\right) \cdot k \frac{(m+k-1)!}{m!} \\ &= \sum_{m=0}^{\infty} \mathbf{P}\left(\tau_D > m\right) \cdot k \prod_{i=1}^{k-1} (m+i)\,.\end{aligned}$$

It was shown in [19, Chapter 6] that

$$k \prod_{i=1}^{k-1} (m+i) = \sum_{i=1}^{k} \begin{bmatrix} k \\ m \end{bmatrix} \left((m+1)^{i+1} - m^i\right).$$

Using this, we get

$$\begin{aligned}\mathbb{E}\,\tau_C &= \sum_{m=0}^{\infty} \mathbf{P}\left(\tau_D > m\right) \sum_{i=1}^{k} \begin{bmatrix} k \\ m \end{bmatrix} \left((m+1)^{i+1} - m^i\right) \\ &= \sum_{i=1}^{k} \begin{bmatrix} k \\ m \end{bmatrix} \sum_{m=0}^{\infty} \mathbf{P}\left(\tau_D > m\right) \left((m+1)^{i+1} - m^i\right) \\ &= \sum_{i=1}^{k} \begin{bmatrix} k \\ m \end{bmatrix} \mathbb{E}\,\tau_D.\end{aligned}$$

Reverting (3.8), we obtain (3.9) (see [19]). □

3.4.2 Parametrization

One of the advantages of the continuous process is that it admits a simple parametrization. Denote by τ_i the stopping time when we add element i to the collection, and let

$$x_i = \tau_i - \max_{j \prec i} \tau_j.$$

Note that the stopping times τ_i are almost always finite, thus, so are the random variables x_i.

Proposition 3.4.2. *For any $1 \leq i \leq n$, the random variable x_i has the exponential* *distribution with parameter* p_i*, namely*

$$\mathbf{P}(X_i > t) = e^{-p_i t}$$

for all $t \geq 0$.

Proof. In order to collect element i, we have to collect all the coupons preceding i, which takes time $y_i = \max_{j \prec i} \tau_j$. Then we have to wait until the coupon i comes for the first time, which takes x_i in addition. Note that y_i is a stopping time, and the Poisson process is a strong Markov process. Thus, the quantity

$$\mathbf{P}(x_i > t_1 \mid y_i = t_2)$$

depends on $t = t_1 - t_2$ only, and the random variable x_i has the exponential distribution with parameter p_i. ☐

Knowing the random variables x_i, we can easily reconstruct τ_i. Indeed, if element i is a minimal element in $\mathcal{P}$, then $\tau_i = x_i$. Therefore, for any element i we can recursively compute $\max_{j \prec i} \tau_j$ and set

$$\tau_i = x_i + \max_{j \prec i} \tau_j.$$

Proposition 3.4.3. *Let* $\mathrm{Ch}_i(\mathcal{P})$ *denote the set of all chains in* $\mathcal{P}$ *with the maximal element* i*. Then*

$$\tau_i = \max_{c \in \mathrm{Ch}_i(\mathcal{P})} \mathbf{x} \cdot \boldsymbol{\chi}(c), \tag{3.10}$$

and

$$\tau = \max_{c \in \mathrm{Ch}(\mathcal{P})} \mathbf{x} \cdot \boldsymbol{\chi}(c), \tag{3.11}$$

where $\mathbf{x} = (x_1, \ldots, x_n)^T$.

Proof. Let $\mathrm{Ch}'_i(\mathcal{P})$ stand for all chains in the poset $\mathcal{P}$, whose elements strictly precede element i. Then

$$\mathrm{Ch}'_i(\mathcal{P}) = \bigcup_{j \prec i} \mathrm{Ch}_j(\mathcal{P}),$$

and hence

$$\begin{aligned}\max_{c\in \mathrm{Ch}_i(\mathcal{P})} \mathbf{x}\cdot\boldsymbol{\chi}(c) &= x_i + \max_{c\in \mathrm{Ch}_i'(\mathcal{P})} \mathbf{x}\cdot\boldsymbol{\chi}(c) \\ &= x_i + \max_{j\prec i}\max_{c\in \mathrm{Ch}_j(\mathcal{P})} \mathbf{x}\cdot\boldsymbol{\chi}(c).\end{aligned}$$

Thus, by induction we obtain (3.10).

In order to get (3.11), it remains to note that

$$\mathrm{Ch}(\mathcal{P}) = \bigcup_{i\in\mathcal{P}} \mathrm{Ch}_i(\mathcal{P}), \qquad \text{and} \qquad \tau = \max_{i\in\mathcal{P}} \tau_i.$$

This completes the proof. □

Thus, we can use random variables x_i in order to analyze the continuous coupon collector's process. Combining Propositions 3.4.2 and 3.4.3, we obtain the following fact.

Lemma 3.4.4. *For any poset $\mathcal{P}$ on n elements with a sample probability vector $\mathbf{p}$, we have*

$$\mathbf{P}(\tau_C(\mathbf{p},\mathcal{P}) \le t) = \int_{t\mathcal{C}(\mathcal{P})} p_1 e^{-p_1x_1}\dots p_n e^{-p_nx_n} dx_1\dots dx_n, \tag{3.12}$$

and

$$\mathbb{E}\,\tau_C^k(\mathbf{p},\mathcal{P}) = \int_0^\infty\dots\int_0^\infty p_1e^{-p_1x_1}\dots p_ne^{-p_nx_n}\left[\max_{c\in\mathrm{Ch}(\mathcal{P})}\mathbf{x}\cdot\boldsymbol{\chi}(c)\right]^k dx_1\dots dx_n \tag{3.13}$$

for any $k > 0$.

Proof. Since we are going to deal with the continuous version only, for brevity we will omit the subscript C. From Proposition 3.4.2, we easily obtain for $\tau = \tau(p,\mathcal{P})$

$$\mathbf{P}(\tau\le t) = \int_{\tau\le t} p_1e^{-p_1x_1}\dots p_ne^{-p_nx_n}dx_1\dots dx_n,$$

and

$$\mathbb{E}\,\tau^k = \int_0^\infty\dots\int_0^\infty \tau^k\cdot p_1e^{-p_1x_1}\dots p_ne^{-p_nx_n}dx_1\dots dx_n.$$

Now relation (3.11) leads directly to (3.13).

Also Proposition 3.4.3 implies that condition $\tau < t$ is equivalent to

$$\mathbf{x} \cdot \chi(c) \leq t \qquad \text{for all chains } c \in \mathrm{Ch}(\mathcal{P}).$$

From the definition of the chain polytope (1.5, 1.6), we get

$$\tau \leq t \Longleftrightarrow \mathbf{x} \in \mathcal{C}(\mathcal{P}),$$

and thus, relation (3.12) is proved. □

Example 3.4.5. Let us consider the poset $\mathcal{I}_n$. Since all its elements are incomparable, the poset has n chains, each consisting of a single element. Therefore, the chain polytope of $\mathcal{P}$ is a hypercube given by the following inequalities

$$0 \leq x_i \leq 1 \quad \text{for all } 1 \leq i \leq n.$$

With this, the relation (3.12) becomes

$$\begin{aligned} \mathbf{P}(\tau_C(\mathbf{p}, \mathcal{P}) \leq t) &= \int_0^t \ldots \int_0^t p_1 \ldots p_n e^{-p_1 x_1} \ldots e^{-p_1 x_1} dx_1 \ldots dx_n \\ &= \left(1 - e^{-p_1 t}\right) \ldots \left(1 - e^{-p_n t}\right). \end{aligned}$$

Thus, the tail probability can be expressed as follows:

$$\begin{aligned} \mathbf{P}(\tau_C(\mathbf{p}, \mathcal{P}) > t) &= e^{-p_1 t} + \ldots + e^{-p_n t} - e^{-(p_1+p_2)t} - \ldots - e^{-(p_{n-1}+p_n)t} \\ &\quad + \ldots + (-1)^{n-1} e^{-(p_1+\ldots+p_n)t}, \end{aligned}$$

and we obtain the classical result for the expected value

$$\begin{aligned} \mathbb{E}\, \tau_C(\mathbf{p}, \mathcal{P}) &= \int_0^\infty \mathbf{P}(\tau_C(\mathbf{p}, \mathcal{P}) > t)\, dt \\ &= \frac{1}{p_1} + \ldots + \frac{1}{p_n} - \frac{1}{p_1 + p_2} - \ldots - \frac{1}{p_{n-1} + p_n} + \ldots + \frac{(-1)^{n-1}}{p_1 + \ldots + p_n}. \end{aligned}$$

A slightly different derivation of this result is given in the proof of Lemma 3.9.1.

3.4.3 Support functions

In order to simplify the calculations and clarify the intuition, we will use support functions when dealing with

$$\max_{c\in\mathcal{C}(\mathcal{P})} \mathbf{x}\cdot\boldsymbol{\chi}(c).$$

Recall that for any convex set $B\subset\mathbb{R}^n$ its *support function* $\varphi_B(\mathbf{x})$ is defined as

$$\varphi_B(\mathbf{x}) = \sup_{\mathbf{y}\in B} \mathbf{x}\cdot\mathbf{y}. \tag{3.14}$$

Note that if B is a polytope, we can replace sup by max.

Thus, if we consider the polytope $\mathcal{A}(\mathcal{P})$ in $\mathbb{R}^n$ which is the convex hull of

$$\boldsymbol{\chi}(c) \quad \text{for all chains } c\in\mathrm{Ch}(\mathcal{P}),$$

then relation (3.13) becomes

$$\mathbb{E}\,\tau_C^k(\mathbf{p},\mathcal{P}) = \int_0^\infty\!\!\dots\int_0^\infty p_1e^{-p_1x_1}\dots p_ne^{-p_nx_n}\varphi_{\mathcal{A}(\mathcal{P})}(\mathbf{x})^k\,dx_1\dots dx_n. \tag{3.15}$$

The polytope $\mathcal{A}(\mathcal{P})$ is usually called the *antichain polytope* of poset $\mathcal{P}$.

Despite the variety of facts about support functions, we will be mainly interested in the following properties.

Proposition 3.4.6. *Let B and C be bounded convex sets in $\mathbb{R}^n$. If $B\subseteq C$ then*

$$\varphi_B(\mathbf{x})\le\varphi_C(\mathbf{x})$$

for any vector $\mathbf{x}\in\mathbb{R}^n$.

Proposition 3.4.7. *Let B be a bounded convex set in $\mathbb{R}^n$. Then for any vectors $\mathbf{x},\mathbf{y}\in$*

$\mathbb{R}^n$ *and any non-negative* α *and* β *we have:*

$$\varphi_B(\alpha\mathbf{x}) = \alpha\varphi_B(\mathbf{x}), \tag{3.16}$$

$$\varphi_B(\alpha\mathbf{x}+\beta\mathbf{y}) \le \alpha\varphi_B(\mathbf{x}) + \beta\varphi_B(\mathbf{y}). \tag{3.17}$$

Proof. Proposition 3.4.6 and relation (3.16) follow directly from the definition of convex functions. Thus, it remains to prove (3.17).

Due to (3.4.6), we can assume that $\alpha = \beta = 1$. We have

$$\begin{aligned} \varphi_B(\mathbf{x}+\mathbf{y}) &= \sup_{\mathbf{z}\in B} \mathbf{z}\cdot(\mathbf{x}+\mathbf{y}) \\ &\le \sup_{\mathbf{z}\in B} \mathbf{z}\cdot\mathbf{x} + \sup_{\mathbf{z}\in B} \mathbf{z}\cdot\mathbf{y} \\ &= \varphi_B(\mathbf{x}) + \varphi_B(\mathbf{y}), \end{aligned}$$

as desired. □

For element i of poset $\mathcal{P}$ let $\mathcal{P}_i$ be the induced poset on elements which are comparable with i. Namely, the set of elements of $\mathcal{P}_i$ is

$$\{j \in \mathcal{P} \mid j \preceq i \text{ or } i \preceq j\}$$

and for any elements j_1 and j_2

$$j_1 \prec j_2 \text{ in } \mathcal{P}_i \iff j_1 \prec j_2 \text{ in } \mathcal{P}.$$

Lemma 3.4.8. *For any element* i *of an* n *element poset* $\mathcal{P}$*, we have*

$$\lim_{p_i\to 0}\left[\mathbb{E}\,\tau^k(\mathbf{p},\mathcal{P}) - \mathbb{E}\,\tau^k(\mathbf{p},\mathcal{P}_i)\right] = 0, \tag{3.18}$$

$$\lim_{p_i\to\infty}\mathbb{E}\,\tau^k(\mathbf{p},\mathcal{P}) = \mathbb{E}\,\tau^k(\mathbf{p},\mathcal{P}\setminus\{i\}), \tag{3.19}$$

where k *is positive integer.*

Remark. The limit in (3.19) does not make much probabilistic sense in the discrete version since the sum of all sample probabilities cannot exceed 1. However, considering $\mathbb{E}\,\tau_D^k$ as an abstract function of n variables, the limit exists and the relation (3.19) holds.

Example 3.4.9. Let us consider a poset $\mathcal{P}$ on elements a, b, c and d with the following partial order:

$$a \prec c, \quad b \prec c, \quad b \prec d.$$

Let the sample probabilities of these elements be p_a, p_b, p_c and p_d respectively. It is not difficult to show (for instance, through (3.2)) that

$$\mathbb{E}\,\tau\left(\mathcal{P}\right) = \frac{1}{p_a} + \frac{1}{p_b} + \frac{1}{p_c} + \frac{1}{p_d} - \frac{\left(p_a + p_b + p_d\right)\left(p_a + p_c + p_d\right)}{\left(p_a + p_b\right)\left(p_a + p_d\right)\left(p_c + p_d\right)}.$$

Taking the limit as $p_a \to \infty$, we get

$$\lim_{p_a \to \infty} \mathbb{E}\,\tau\left(\mathcal{P}\right) = \frac{1}{p_b} + \frac{1}{p_c} + \frac{1}{p_d} - \frac{1}{p_c + p_d},$$

which is equal to $\mathbb{E}\,\tau\left(\mathcal{P} \setminus \{a\}\right)$.

Also since element c is the only element comparable to a, we have

$$\mathbb{E}\,\tau\left(\mathcal{P}_a\right) = \frac{1}{p_a} + \frac{1}{p_c},$$

and thus,

$$\begin{aligned}
\lim_{p_a \to 0}\left(\tau\left(\mathcal{P}\right) - \tau\left(\mathcal{P}_a\right)\right) &= \lim_{p_a \to 0}\left(\frac{1}{p_b} + \frac{1}{p_d} - \frac{\left(p_a + p_b + p_d\right)\left(p_a + p_c + p_d\right)}{\left(p_a + p_b\right)\left(p_a + p_d\right)\left(p_c + p_d\right)}\right) \\
&= \frac{1}{p_b} + \frac{1}{p_d} - \frac{\left(p_b + p_d\right)\left(p_c + p_d\right)}{p_b p_d\left(p_c + p_d\right)} = 0,
\end{aligned}$$

as given by (3.18).

Before we proceed with the proof of Lemma 3.4.8, let us mention the probabilistic intuition behind it. If p_i is large, then in the continuous version the coupon i arrives much more often than the others. Hence, when it is possible to add the coupon i, i.e. when all preceding coupons are already in the collection, it takes a small amount of time in comparison with completing of the rest. Therefore, the process behaves as it would behave on the poset without element i. If instead $p_i \to 0$, then all moments of τ will go to infinity, and the bottleneck is the set of elements that are comparable with i.

Proof. Corollary 3.4.1 yields that it suffices to prove (3.18) and (3.19) for continuous version. Thus, to simplify notation we will omit the subscript C.

Without loss of generality we can assume that $i = n$. Let $\mathcal{A}_0$ and $\mathcal{A}_1$ be the convex hulls of vertices of $\mathcal{A}(\mathcal{P})$ whose first coordinate is 0 and 1 respectively. It is not hard to see that

$$\mathcal{A}_0 = \mathcal{A}(\mathcal{P} \setminus \{n\}),$$

$$\mathcal{A}_1 = \mathcal{A}(\mathcal{P}_n),$$

thus,

$$\varphi_{\mathcal{A}(\mathcal{P})}(\mathbf{x}) = \max\left(\varphi_{\mathcal{A}_0}(\mathbf{x}), \varphi_{\mathcal{A}_1}(\mathbf{x})\right). \tag{3.20}$$

Denote by the prime ($'$) the projection on the first $(n-1)$ coordinates.

Let

$$y_j = p_j x_j = \frac{x_j}{r_j}, \quad 1 \le j \le n,.$$

and denote for brevity

$$\varphi_0 = \varphi_{\mathcal{A}'_0}\left(\frac{y_1}{p_1}, \ldots, \frac{y_{n-1}}{p_{n-1}}\right) \quad \text{and} \quad \varphi_1 = \varphi_{\mathcal{A}'_1}\left(\frac{y_1}{p_1}, \ldots, \frac{y_{n-1}}{p_{n-1}}\right).$$

Now we obtain

$$\varphi_{\mathcal{A}(\mathcal{P})} = \max\left(\varphi_0, \varphi_1 + r_n y_n\right),$$

hence (3.13) becomes

$$\mathbb{E}\,\tau^k(\mathbf{p}, \mathcal{P}) = \int_0^\infty \cdots \int_0^\infty e^{-y_1 - \ldots - y_n} \max\left(\varphi_0, \varphi_1 + r_n y_n\right)^k dy_1 \ldots dy_n.$$

Let now $p_n \to \infty$, or, equivalently, $r_n \to 0$. In this case,

$$\lim_{p_n \to \infty} \mathbb{E}\, \tau^k (\mathbf{p}, \mathcal{P}) = \int_0^\infty \ldots \int_0^\infty e^{-y_1 - \ldots - y_n} \max (\varphi_0, \varphi_1)^k dy_1 \ldots dy_n.$$

Since all vertices of $\mathcal{A}_0$ are indicators of chains in $\mathcal{P} \setminus \{n\}$ and all vertices in $\mathcal{A}_2$ are chains in $\mathcal{P}$ that contain n, it is obvious that

$$\mathcal{A}_0' \supset \mathcal{A}_1',$$

Therefore, Proposition 3.4.6 gives

$$\varphi_0 \geq \varphi_1,$$

which implies

$$\lim_{p_n \to \infty} \mathbb{E}\, \tau^k (\mathbf{p}, \mathcal{P}) = \int_0^\infty \ldots \int_0^\infty e^{-y_1 - \ldots - y_n} \varphi_0^k \, dy_1 \ldots dy_n = \mathbb{E}\, \tau^k (\mathbf{p}, \mathcal{P} \setminus \{n\}).$$

Let now $p_n \to 0$, or $r_n \to \infty$. Since all chains in $\mathcal{P}_n$ contain element n, we have

$$\mathbb{E}\, \tau^k (\mathbf{p}, \mathcal{P}_n) = \int_0^\infty \ldots \int_0^\infty e^{-y_1 - \ldots - y_n} (\varphi_1 + r_n y_n)^k \, dy_1 \ldots dy_n,$$

and therefore,

$$\begin{aligned} \mathbb{E}\, \tau^k (\mathbf{p}, \mathcal{P}) - \mathbb{E}\, \tau^k (\mathbf{p}, \mathcal{P}_n) &= \int_{\varphi_0 > \varphi_1 + r_n y_n} e^{-y_1 - \ldots - y_n} \varphi_0^k \, dy_1 \ldots dy_n \\ &= \int_0^\infty \ldots \int_0^\infty e^{-y_1 - \ldots - y_{n-1}} \varphi_0^k \\ &\quad \times \left(1 - \exp\left(-\frac{\varphi_0 - \varphi_1}{r_n}\right)\right) dy_1 \ldots dy_{n-1}. \end{aligned}$$

Using inequality $1 - e^{-\xi} \leq \xi$ which holds for all real ξ, we obtain

$$1 - \exp\left(-\frac{\varphi_0 - \varphi_1}{r_n}\right) \leq \frac{\varphi_0 - \varphi_1}{r_n} = p_n (\varphi_0 - \varphi_1).$$

This leads to

$$\mathbb{E}\,\tau^k(\mathbf{p},\mathcal{P})-\mathbb{E}\,\tau^k(\mathbf{p},\mathcal{P}_n)\ \le\ p_n\int_0^\infty\!\!\cdots\int_0^\infty e^{-y_1-\ldots-y_{n-1}}\varphi_0^k(\varphi_0-\varphi_1)\,dy_1\ldots dy_{n-1}.$$

The integral in the last display is finite: it is bounded from above by

$$\int_0^\infty\!\!\cdots\int_0^\infty e^{-y_1-\ldots-y_{n-1}}\varphi_0^{k+1}dy_1\ldots dy_{n-1},$$

which is equal to $\mathbb{E}\,\tau^{k+1}(\mathbf{p},\mathcal{P}\setminus\{n\})$. Thus,

$$\lim_{p_n\to 0}\left[\mathbb{E}\,\tau^k(\mathbf{p},\mathcal{P})-\mathbb{E}\,\tau^k(\mathbf{p},\mathcal{P}_n)\right]\le 0.$$

On the other hand, poset $\mathcal{P}_n$ is a subposet of $\mathcal{P}$, therefore by Corollary 3.2.4

$$\mathbb{E}\,\tau^k(\mathbf{p},\mathcal{P})-\mathbb{E}\,\tau^k(\mathbf{p},\mathcal{P}_n)\ge 0.$$

This concludes the proof. □

Now we have all ingredients to prove Theorem 1.2.4.

Proof. (Theorem 1.2.4) Due to Corollary 3.4.1, it suffices to prove the theorem for the continuous version only.

Note that Lemma 3.4.4 shows that if two posets $\mathcal{P}$ and $\mathcal{Q}$ have the same chain polytopes, then

$$\mathbb{E}\,\tau_C^k(\mathbf{p},\mathcal{P})=\mathbb{E}\,\tau_C^k(\mathbf{p},\mathcal{Q}).$$

Hence we have to show the converse.

In [48], it was proved that the chain polytope of a poset $\mathcal{R}$ is the convex hull of indicators of all antichains of $\mathcal{R}$. Therefore, it suffices to show that if elements i and j are incomparable in $\mathcal{P}$ then they are also incomparable in $\mathcal{Q}$. Let $\mathcal{P}'$ and $\mathcal{Q}'$ be induced posets on the elements i and j. Taking the limit of $\mathbb{E}\,\tau_C^k$ as $p_M\to\infty$

for $m \neq i$ and $m \neq j$, by (3.19) we get

$$\mathbb{E}\,\tau_C^k\left(\mathcal{P}'\right) = \mathbb{E}\,\tau_C^k\left(\mathcal{Q}'\right) \tag{3.21}$$

for all sample probabilities p_i and p_j.

If elements i and j in $\mathcal{Q}'$ were comparable, we would have

$$\begin{aligned}
\mathbb{E}\,\tau_C^k\left(\mathcal{Q}'\right) &= \int_0^\infty \int_0^\infty p_i e^{-p_i x_i} p_j e^{-p_j x_j} \left(x_i + x_j\right)^k dx_i dx_j \\
&> \int_0^\infty \int_0^\infty p_i e^{-p_i x_i} p_j e^{-p_j x_j} \max\left(x_i, x_j\right)^k dx_i dx_j \\
&= \mathbb{E}\,\tau_C^k\left(\mathcal{P}'\right),
\end{aligned}$$

which contradicts (3.21). Thus, elements i and j are incomparable in $\mathcal{Q}'$, and the theorem follows. □

3.5 Convexity

Theorem 1.2.6 immediately follows from (3.15), because the expression for $\mathbb{E}\,\tau^k$ involves integration of the support function which is convex (Proposition 3.4.7). A more detailed proof is the following.

Proof. (Theorem 1.2.6). We will prove the convexity of $\mathbb{E}\,\tau$ for the continuous process, which will entail the convexity for the discrete version. First, let us show that $\mathbb{E}\,\tau^k$ is a convex function of $\mathbf{r}$ for any $k \geq 1$.

Let $\mathbf{a}$ and $\mathbf{b}$ be n-dimensional vectors. With slight abuse of notation, we denote by $\{\mathbf{a} \cdot \mathbf{b}\}$ and $\{\mathbf{a}/\mathbf{b}\}$ the vectors

$$(a_1 b_1, \ldots, a_n b_n) \quad \text{and} \quad \left(\frac{a_1}{b_1}, \ldots, \frac{a_n}{b_n}\right).$$

Then for variables $y_i = \frac{x_i}{p_i}$ we obtain from (3.13)

$$\begin{aligned} \mathbb{E}\,\tau^k &= \int_0^\infty \cdots \int_0^\infty e^{-y_1} \ldots e^{-y_n} \left[\max_{c \in \mathrm{Ch}(\mathcal{P})} \chi(c) \cdot \{\mathbf{y}/\mathbf{p}\}\right]^k dy_1 \ldots dy_n \\ &= \int_0^\infty \cdots \int_0^\infty e^{-y_1} \ldots e^{-y_n} \left[\max_{c \in \mathrm{Ch}(\mathcal{P})} \chi(c) \cdot \{\mathbf{y} \cdot \mathbf{r}\}\right]^k dy_1 \ldots dy_n. \end{aligned}$$

Now note that for vectors $\mathbf{r}$, $\mathbf{r}'$ and $0 \leq \lambda \leq 1$ we have

$$\begin{aligned} \max_{c \in \mathrm{Ch}(\mathcal{P})} \chi(c) \cdot \{\mathbf{y}\,(\lambda \mathbf{r} + (1-\lambda)\,\mathbf{r}')\} \;\leq\; & \lambda \max_{c \in \mathrm{Ch}(\mathcal{P})} \chi(c) \cdot \{\mathbf{y} \cdot \mathbf{r}\} \\ & + (1-\lambda) \max_{c \in \mathrm{Ch}(\mathcal{P})} \chi(c) \cdot \{\mathbf{y} \cdot \mathbf{r}'\}. \end{aligned}$$

Thus, the function

$$f(\mathbf{r}) = \max_{c \in \mathrm{Ch}(\mathcal{P})} \chi(c) \cdot \{\mathbf{y} \cdot \mathbf{r}\}$$

is convex in $\mathbf{r}$ for any vector $\mathbf{y}$. So is function $f(\mathbf{r})^k$, since $k \geq 1$, which implies the convexity of $\mathbb{E}\,\tau^k$.

In order to show the convexity with respect to $\mathbf{p}$, it remains to prove that function

$$g(\mathbf{p}) = \max_{c \in \mathrm{Ch}(\mathcal{P})} \boldsymbol{\chi}(c) \cdot \{\mathbf{y}/\mathbf{p}\}$$

is convex for any $\mathbf{y}$. In the similar way, we obtain

$$\begin{aligned} g(\lambda \mathbf{p} + (1-\lambda)\mathbf{p}') &= \max_{c \in \mathrm{Ch}(\mathcal{P})} \boldsymbol{\chi}(c) \cdot \left\{ \frac{\mathbf{y}}{\lambda \mathbf{p} + (1-\lambda)\mathbf{p}'} \right\} \\ &\leq \max_{c \in \mathrm{Ch}(\mathcal{P})} \boldsymbol{\chi}(c) \cdot \left(\lambda \left\{ \frac{\mathbf{y}}{\mathbf{p}} \right\} + (1-\lambda) \left\{ \frac{\mathbf{y}}{\mathbf{p}'} \right\} \right) \\ &\leq \lambda \max_{c \in \mathrm{Ch}(\mathcal{P})} \boldsymbol{\chi}(c) \cdot \left\{ \frac{\mathbf{y}}{\mathbf{p}} \right\} + (1-\lambda) \max_{c \in \mathrm{Ch}(\mathcal{P})} \boldsymbol{\chi}(c) \cdot \left\{ \frac{\mathbf{y}}{\mathbf{p}'} \right\} \\ &= \lambda g(\mathbf{p}) + (1-\lambda) g(\mathbf{p}') . \end{aligned}$$

This accomplishes the proof of Theorem 1.2.6. □

Remark 3.5.1. We believe that higher moments for the discrete CCP are also convex. However, since we do not have any convenient tool to analyze it, this question remains open.

In [45] it was proved that

$$\mathbf{P}(\tau_D(\mathbf{p}, \mathcal{I}_n) > m)$$

is minimized on the sheet $p_1 + p_2 + \ldots + p_n = 1$ by $\mathbf{p} = \frac{1}{n} \cdot \mathbf{1}$. Although in [7] it was proved that $\mathbf{P}(\tau_D(\mathbf{p}, \mathcal{I}_n) > m)$ is a Schur-convex function of $\mathbf{p}$, it is not convex. Moreover, we can easily show that for any poset $\mathcal{P}$ containing more than one element neither of the following quantities

$$\mathbf{P}(\tau_D(\mathbf{p}, \mathcal{P}) > m), \quad \mathbf{P}(\tau_C(\mathbf{p}, \mathcal{P}) > t)$$

is a convex function of $\mathbf{p}$. The main idea is to take the limit when all but two sample probabilities tend to infinity. Doing this, we can reduce the problem to posets $\mathcal{S}_2$ or $\mathcal{I}_2$, and for them we can verify the statement directly.

3.6 Estimates for posets

In this section we will prove Theorem 1.2.7. The lower bound follows easily from Corollary 3.2.4, and the upper bound is an outcome of convexity and estimates of $\mathbb{E}\,\tau$ on the boundary of an appropriate convex set.

Proof. (Theorem 1.2.7) Due to Lemma 1.2.2, we will consider the continuous process only.

In order to obtain the lower bound, let us consider the chain c' for which

$$L = \chi(c') \cdot \mathbf{r}.$$

Obviously, the chain c' is a subposet of $\mathcal{P}$, and therefore,

$$\mathbb{E}\,\tau(\mathcal{P}) \geq \mathbb{E}\,\tau(c') = L.$$

In order to prove the upper bound, note that condition

$$\max_{c \in \mathrm{Ch}(\mathcal{P})} \chi(c) \cdot \mathbf{r} = L.$$

is equivalent to the following conditions on $\mathbf{r} = (r_1, \ldots, r_n)$:

$$r_i \geq 0 \quad \text{for all} \quad 1 \leq i \leq n,$$

$$\chi(c) \cdot \mathbf{r} \leq L \quad \text{for any chain} \quad c \in \mathrm{Ch}(\mathcal{P}).$$

By the definition of the chain polytope, we get

$$\mathbf{r} \in L \cdot \mathcal{C}(\mathcal{P}).$$

In [48] it was shown that the vertices of the polytope $\mathcal{C}(\mathcal{P})$ are indicators of antichains of $\mathcal{P}$. Therefore, for any $\mathbf{r}$ there exist non-negative numbers α_a indexed by antichains that

$$\mathbf{r} = L \sum_{a \in \mathrm{ACh}(\mathcal{P})} \alpha_a \boldsymbol{\chi}(a) \quad \text{and} \quad \sum_{a \in \mathrm{ACh}(\mathcal{P})} \alpha_a = 1.$$

Now with slight abuse of notation, by convexity of $\mathbb{E}\,\tau$ (Theorem 1.2.6) we have:

$$\begin{aligned} \mathbb{E}\,\tau(\mathbf{r}) &= \mathbb{E}\,\tau\left(\sum_{a \in \mathrm{ACh}(\mathcal{P})} \alpha_a \cdot L \cdot \boldsymbol{\chi}(a)\right) \\ &\leq \sum_{a \in \mathrm{ACh}(\mathcal{P})} \alpha_a \,\mathbb{E}\,\tau(L \cdot \boldsymbol{\chi}(a)). \end{aligned}$$

However, the quantity $\mathbb{E}\,\tau(L \cdot \boldsymbol{\chi}(a))$ is easy to compute: this is the expected time of completing the collection of incomparable elements with equal sample probabilities $\frac{1}{L}$. The number of these elements is the size of antichain a which is at most M, therefore from Corollary 3.2.8 we obtain

$$\mathbb{E}\,\tau(L \cdot \boldsymbol{\chi}(a)) \leq L \cdot H_M,$$

and finally

$$\begin{aligned} \mathbb{E}\,\tau(\mathbf{r}) &\leq \sum_{a \in \mathrm{ACh}(\mathcal{P})} \alpha_a \,\mathbb{E}\,\tau(L \cdot \boldsymbol{\chi}(a)) \\ &\leq \sum_{a \in \mathrm{ACh}(\mathcal{P})} \alpha_a L \cdot H_M \\ &= L \cdot H_M. \end{aligned}$$

This completes the proof of the theorem. □

3.7 Chain bound

In this section, we will consider posets on elements with equal sample probabilities p, i.e. the probability vector is $p \cdot \mathbf{1}$.

Lemma 3.7.1. *Let $\mathcal{S}_n$ denote the completely ordered poset (the chain) on n elements. Then*

$$\mathbf{P}\left(\tau_D\left(p \cdot \mathbf{1}, \mathcal{S}_n\right) > m\right) \leq \exp\left(-mp + n + n \log \frac{mp}{n}\right), \tag{3.22}$$

$$\mathbf{P}\left(\tau_C\left(p \cdot \mathbf{1}, \mathcal{S}_n\right) > t\right) \leq \exp\left(-tp + n + n \log \frac{tp}{n}\right) \tag{3.23}$$

for any integer $m > 0$ and any real $t > 0$.

Proof. Let us first consider the discrete process. As usual, denote by τ_i be the stopping time when we collect element i, and let $\tau_0 = 0$. If now $x_i = \tau_i - \tau_{i-1}$, then

$$\tau_D = x_1 + \ldots + x_n,$$

and $x_1, \ldots, x_n$ are independent random variables that have geometric distribution with parameter p.

For fixed m, consider m i.i.d. random variables $y_1, \ldots, y_m$ such that

$$\mathbf{P}\left(y_i = 1\right) = p \quad \text{and} \quad \mathbf{P}\left(y_i = 0\right) = 1 - p.$$

Observe that for the discrete process we have

$$\mathbf{P}\left(\tau_i > k\right) = \mathbf{P}\left(y_1 + \ldots + y_k < i\right),$$

where $k \leq m$ and $i \leq n$. In other words, each random variable y_i indicates whether we add a new coupon (if any) to the collection at step i. Therefore,

$$\mathbf{P}\left(\tau_D\left(p \cdot \mathbf{1}, \mathcal{S}_n\right) > m\right) = \mathbf{P}\left(y_1 + \ldots + y_m < n\right).$$

Applying the Chernoff bound, we get for any $\delta > 0$

$$\mathbf{P}\left(y_1 + \ldots + y_m < (1-\delta)\, mp\right) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{mp},$$

and choosing δ such that

$$(1-\delta)\, mp = n,$$

we obtain

$$\begin{aligned}\mathbf{P}\left(\tau_D\left(p \cdot \mathbf{1}, \mathcal{S}_n\right) > m\right) &\leq \exp\left(-mp\delta - mp\,(1-\delta)\log(1-\delta)\right) \\ &= \exp\left(-mp + n - n\log\frac{n}{mp}\right).\end{aligned}$$

For the continuous version, note that if p' is fixed positive number and $p \to 0$, the random variable $\frac{1}{p'} \cdot px_1$ converges in distribution to an exponential random variable with parameter p'. Therefore, if $p \to 0$ then in distribution

$$\frac{1}{p'} \cdot p\tau_D\left(p \cdot \mathbf{1}, \mathcal{S}_n\right) \to \tau_C\left(p' \cdot \mathbf{1}, \mathcal{S}_n\right).$$

This implies

$$\mathbf{P}\left(\tau_C\left(p' \cdot \mathbf{1}, \mathcal{S}_n\right) > t\right) \leq \exp\left(-tp' + n + \log\frac{tp'}{n}\right).$$

This concludes the proof of the lemma. □

Remark 3.7.2. In the proof, we use the reduction of the random variable τ to a sequence of binary i.i.d. random variables, and then apply the Markov inequality to the moment-generating function of the sequence (this is how the Chernoff inequality is established). We could also omit the reduction and apply the same procedure to τ directly. However, in that case it was difficult to choose the value of parameter λ in $\mathbb{E}\, e^{\lambda\tau}$ to get a useful upper bound.

With Lemma 3.7.1 in hands, we can proceed with the proof of Theorem 1.2.8.

Proof. (Theorem 1.2.8) First, we will prove (1.11). Since expectations of τ_C and τ_D coincide, we can consider the continuous process only.

If $k = 1$ then all elements in the poset are incomparable, thus $\mathbb{E}\,\tau\left(p\cdot\mathbf{1},\mathcal{P}\right) = \frac{1}{p}H_n$, and both inequalities hold. Hence, we can assume that $k > 1$.

Consider a maximal element i in $\mathcal{P}$ and the corresponding subposet $\mathcal{P}_i$. Let ℓ be a linear extension of $\mathcal{P}_i$. Obviously, the poset $\mathcal{P}_i$ is a subposet of ℓ. Therefore, from Proposition 3.2.3 we have

$$\mathbf{P}\left(\tau\left(p\cdot\mathbf{1},\mathcal{P}_i\right) > t\right) \leq \mathbf{P}\left(\tau\left(p\cdot\mathbf{1},\ell\right) > t\right)$$

for any $t > 0$. However, the poset ℓ is a chain of length at most k, hence Lemma 3.7.1 yields

$$\mathbf{P}\left(\tau\left(p\cdot\mathbf{1},\mathcal{P}_i\right) > t\right) \leq \exp\left(-tp + k + k\log\frac{tp}{k}\right).$$

If $\mathcal{M}\left(\mathcal{P}\right)$ is a set of all maximal elements of $\mathcal{P}$, then

$$\tau\left(p\cdot\mathbf{1},\mathcal{P}\right) = \max_{i\in\mathcal{M}(\mathcal{P})}\tau_i = \max_{i\in\mathcal{M}(\mathcal{P})}\tau\left(p\cdot\mathbf{1},\mathcal{P}_i\right),$$

and since $\mathcal{M}\left(\mathcal{P}\right)$ contains at most n elements, the union bound gives

$$\mathbf{P}\left(\tau\left(p\cdot\mathbf{1},\mathcal{P}\right) > t\right) \leq n\exp\left(-tp + k + k\log\frac{tp}{k}\right).$$

Let now

$$f\left(t\right) = n\exp\left(-tp + k + k\log\frac{tp}{k}\right),$$

and let $t_0 > 0$ be the smallest real such that for all $t > t_0$ we have

$$f\left(t\right) < 1.$$

Note that

$$\begin{aligned} f\left(\frac{1}{p}\left(\log n + k\right)\right) &= n\exp\left(-\left(\log n + k\right) + k + k\log\frac{\log n + k}{k}\right) \\ &= \exp\left(k\log\left(1 + \frac{\log n}{k}\right)\right) \\ &\geq 2^k. \end{aligned}$$

Thus, we have

$$t_0 > \frac{1}{p}\left(\log n + k\right). \tag{3.24}$$

On the other hand, we can show that

$$t_0 < \frac{1}{p}\left(\log n + k \log\log n + 2k\right). \tag{3.25}$$

Indeed, if $t' = \frac{1}{p}\left(\log n + k\log\log n + 2k\right)$, then

$$f\left(t'\right) = \exp\left(-k\log\log n - k + k\log\frac{\log n + k\log\log n + 2k}{k}\right),$$

and inequality $f\left(t'\right) < 1$ is equivalent to

$$e\log n > \frac{\log n}{k} + \log\log n + 2.$$

However, for $n \geq 3$ we have

$$\left(e - \frac{1}{2}\right)\log n > \log\log n + 2,$$

and the previous inequality follows since $k \geq 2$. It remains to note that the function $f\left(t\right)$ is decreasing for $t > t'$, thus for all $t > t'$ we get

$$f\left(t\right) < f\left(t'\right) < 1.$$

Now we obtain

$$\begin{aligned}
\mathbb{E}\,\tau\left(p\cdot\mathbf{1},\mathcal{P}\right) &= \int_0^\infty \mathbf{P}\left(\tau\left(p\cdot\mathbf{1},\mathcal{P}\right) > t\right)dt \\
&= \int_0^{t_0} \mathbf{P}\left(\tau\left(p\cdot\mathbf{1},\mathcal{P}\right) > t\right)dt + \int_{t_0}^\infty \mathbf{P}\left(\tau\left(p\cdot\mathbf{1},\mathcal{P}\right) > t\right)dt \\
&\leq t_0 + \int_0^\infty f\left(t_0 + \xi\right)d\xi.
\end{aligned}$$

Observe that

$$\begin{aligned} f(t_0+\xi) &= n\exp\left(-t_0p-\xi p+k+k\log\frac{(t_0+\xi)\,p}{k}\right) \\ &= f(t_0)\exp\left(-\xi p+k\log\frac{t_0+\xi}{t_0}\right). \end{aligned}$$

Inequalities $f(t_0)\leq 1$ and

$$\log\frac{t_0+\xi}{t_0}=\log\left(1+\frac{\xi}{t_0}\right)\leq\frac{\xi}{t_0}$$

give

$$\mathbb{E}\,\tau\,(p\cdot\mathbf{1},\mathcal{P})\leq t_0+\int_0^\infty\exp\left(-\xi p+\frac{k}{t_0}\xi\right)d\xi.$$

Using (3.24) and (3.25), we get

$$\begin{aligned} t_0+\int_0^\infty\exp\left(-\xi p+\frac{k}{t_0}\xi\right)d\xi &< \frac{1}{p}t'+\int_0^\infty\exp\left(-\xi p+\frac{k}{\log n+k}p\xi\right)d\xi \\ &< \frac{1}{p}t'+\int_0^\infty\exp\left(-\frac{1}{2}p\xi\right)d\xi \\ &= \frac{1}{p}\left(t'+2\right). \end{aligned}$$

Finally,

$$\mathbb{E}\,\tau\,(p\cdot\mathbf{1},\mathcal{P})<\frac{1}{p}\left(\log n+k\log\log n+2k+2\right),$$

which implies the upper bound in (1.11). The lower bound follows from Corollary 3.2.4 and the fact that the poset on n incomparable elements is a subposet of $\mathcal{P}$.

In the similar way we can estimate the second moment of the continuous process:

$$\begin{aligned} \mathbb{E}\,\tau_C^2\,(p\cdot\mathbf{1},\mathcal{P}) &= \int_0^\infty 2t\mathbf{P}\left(\tau_C\,(p\cdot\mathbf{1},\mathcal{P})>t\right)dt \\ &= \int_0^{t_0} 2t\mathbf{P}\left(\tau_C\,(p\cdot\mathbf{1},\mathcal{P})>t\right)dt+\int_{t_0}^\infty 2t\mathbf{P}\left(\tau_C\,(p\cdot\mathbf{1},\mathcal{P})>t\right)dt \\ &\leq t_0^2+\int_0^\infty 2\,(t_0+\xi)\,f\,(t_0+\xi)\,d\xi \end{aligned}$$

We already established that

$$f\left(t_0+\xi\right) \leq \exp\left(-\frac{1}{2}p\xi\right),$$

and using this, we get

$$\begin{aligned}
\mathbb{E}\,\tau_C^2\left(p\cdot\mathbf{1},\mathcal{P}\right) &< t_0^2+\int_0^\infty 2\left(t_0+\xi\right)\exp\left(-\frac{1}{2}p\xi\right)d\xi \\
&= t_0^2+\frac{8+4pt_0}{p^2}.
\end{aligned}$$

Therefore, recalling that $k=o\left(\frac{\log n}{\log\log n}\right)$, we have for $n\to\infty$

$$\begin{aligned}
\frac{\operatorname{Var}\tau_C\left(p\cdot\mathbf{1},\mathcal{P}\right)}{\left(\mathbb{E}\,\tau_C\left(p\cdot\mathbf{1},\mathcal{P}\right)\right)^2} &< \frac{p^2t_0^2-\log^2 n+8+4pt_0}{\log^2 n} \\
&< \frac{\left(k\log\log n+2k\right)\left(2\log n+k\log\log n+2k\right)}{\log^2 n} \\
&\quad +4\frac{\log n+k\log\log n+2k+2}{\log^2 n} \\
&= k\frac{\log\log n}{\log n}\left(2+k\frac{\log\log n}{\log n}+o\left(1\right)\right)+o\left(1\right) \\
&= o\left(1\right).
\end{aligned}$$

It remains to note that for the discrete process Corollary 3.4.1 implies the following:

$$\begin{aligned}
\operatorname{Var}\tau_D\left(p\cdot\mathbf{1},\mathcal{P}\right) &= \mathbb{E}\,\tau_D^2\left(p\cdot\mathbf{1},\mathcal{P}\right)-\mathbb{E}\,\tau_D^2\left(p\cdot\mathbf{1},\mathcal{P}\right) \\
&= \left(\mathbb{E}\,\tau_C^2\left(p\cdot\mathbf{1},\mathcal{P}\right)-\mathbb{E}\,\tau_C\left(p\cdot\mathbf{1},\mathcal{P}\right)\right)-\mathbb{E}\,\tau_C\left(p\cdot\mathbf{1},\mathcal{P}\right) \\
&< \mathbb{E}\,\tau_C^2\left(p\cdot\mathbf{1},\mathcal{P}\right)-\mathbb{E}\,\tau_C\left(p\cdot\mathbf{1},\mathcal{P}\right) \\
&= \operatorname{Var}\tau_C\left(p\cdot\mathbf{1},\mathcal{P}\right),
\end{aligned}$$

hence

$$\sqrt{\operatorname{Var}\tau_D\left(p\cdot\mathbf{1},\mathcal{P}\right)}=o\left(\mathbb{E}\,\tau_D\left(p\cdot\mathbf{1},\mathcal{P}\right)\right),$$

as desired. □

Applying Lemma 3.7.1 for all chains in the poset, we can prove Theorem 1.2.10.

Proof. (Theorem 1.2.10) Again due to (1.2) we can consider the continuous process only. Using the notation of Section 3.4, denote

$$\tau(c) = \boldsymbol{\chi}(c) \cdot \mathbf{x}$$

for any chain c in $\mathcal{P}$. Then we have

$$\tau(p \cdot \mathbf{1}, \mathcal{P}) = \max_{c \in \mathrm{MCh}(\mathcal{P})} \tau(c),$$

and hence by the union bound we obtain

$$\mathbf{P}(\tau(p \cdot \mathbf{1}, \mathcal{P}) > t) = \mathbf{P}\left(\max_{c \in \mathrm{MCh}(\mathcal{P})} \tau(c) > t\right) \leq \sum_{c \in \mathrm{MCh}(\mathcal{P})} \mathbf{P}(\tau(c) > t)$$

for any $t > 0$. Since any chain c has length at most ℓ, we have from Lemma 3.7.1

$$\mathbf{P}(\tau(c) > t) \leq \exp\left(-pt + \ell + \ell \log \frac{pt}{\ell}\right),$$

which leads to

$$\mathbf{P}(\tau(p \cdot \mathbf{1}, \mathcal{P}) > t) \leq \#\,\mathrm{MCh}(\mathcal{P}) \cdot \exp\left(-pt + \ell + \ell \log \frac{pt}{\ell}\right).$$

For the function

$$f(t) = \#\,\mathrm{MCh}(\mathcal{P}) \cdot \exp\left(-pt + \ell + \ell \log \frac{pt}{\ell}\right),$$

the equation $f(t) = 1$ is equivalent to

$$\frac{pt}{\ell} - \log \frac{pt}{\ell} = 1 + \frac{\log \#\,\mathrm{MCh}(\mathcal{P})}{\ell}.$$

Thus, for all $t > \frac{\ell}{p}\xi^*$ we have $f(t) < 1$. Finally, we obtain

$$\begin{aligned}
\mathbb{E}\,\tau(p\cdot\mathbf{1},\mathcal{P}) &= \int_0^\infty \mathbf{P}\left(\tau(p\cdot\mathbf{1},\mathcal{P}) > t\right) dt \\
&= \int_0^{\ell\xi^*/p} \mathbf{P}\left(\tau(p\cdot\mathbf{1},\mathcal{P}) > t\right) dt + \int_{\ell\xi^*/p}^\infty \mathbf{P}\left(\tau(p\cdot\mathbf{1},\mathcal{P}) > t\right) dt \\
&\leq \frac{\ell}{p}\xi^* + \int_0^\infty f\left(\frac{\ell}{p}\xi^* + t\right) dt \\
&= \frac{\ell}{p}\xi^* + \int_0^\infty \#\,\mathrm{MCh}\,(\mathcal{P}) \exp\left(-\ell\xi^* - pt + \ell + \ell\log\frac{\ell\xi^* + pt}{\ell}\right) dt \\
&= \frac{\ell}{p}\xi^* + \int_0^\infty f\left(\frac{\ell}{p}\xi^*\right) \exp\left(-pt + \ell\log\left(1 + \frac{pt}{\ell\xi^*}\right)\right) dt \\
&\leq \frac{\ell}{p}\xi^* + \int_0^\infty \exp\left(-pt + \frac{pt}{\xi^*}\right) dt \\
&= \frac{1}{p}\left(\ell\xi^* + \frac{\xi^*}{\xi^* - 1}\right).
\end{aligned}$$

The lower bound easily follows from the observation that the longest maximal chain is a subposet of $\mathcal{P}$. □

Corollary 3.7.3. *For any integer $k > 0$ and any positive $a_1, \ldots, a_k$ the limit*

$$\lim_{n\to\infty} \frac{1}{n}\,\mathbb{E}\,\tau_C\left(p\cdot\mathbf{1}, \mathcal{R}\left(a_1 n, \ldots, a_k n\right)\right) \tag{3.26}$$

exists for every $p > 0$.

Proof. Denote for convenience

$$\mathcal{R}_n = \mathcal{R}\left(a_1 n, \ldots, a_k n\right).$$

First, we will show that the sequence $\mathbb{E}\,\tau\left(p\cdot\mathbf{1},\mathcal{R}_n\right)$ is super-additive, namely that

$$\mathbb{E}\,\tau\left(p\cdot\mathbf{1},\mathcal{R}_{n+m}\right) \geq \mathbb{E}\,\tau\left(p\cdot\mathbf{1},\mathcal{R}_n\right) + \mathbb{E}\,\tau\left(p\cdot\mathbf{1},\mathcal{R}_m\right) \tag{3.27}$$

for any n and m.

Indeed, let us consider the poset $\mathcal{R}_{n+m}$. The quantity in the right-hand side can be bounded from above by

$$\mathbb{E}\left(\max_{C} \mathbf{x} \cdot \chi(c)\right),$$

where the maximum is taken over all maximal chains which contain the element

$$(\lfloor a_1 n \rfloor, \ldots, \lfloor a_k n \rfloor).$$

However, in the left-hand side the maximum is taken over all maximal chains in poset $\mathcal{R}_{n+m}$. Therefore, inequality (3.27) holds, and the sequence $\frac{1}{n}\mathbb{E}\,\tau(p\cdot\mathbf{1},\mathcal{R}_n)$ is increasing.

Note that all chains in $\mathcal{R}_n$ have the same length which is

$$L_n = n(a_1 + \ldots + a_k - 1).$$

Also the number of maximal chains in $\mathcal{R}_n$ is

$$k^{n(a_1+\ldots+a_k-1)}.$$

Denote by ξ_n^* the maximal root of

$$\xi_n^* - \log \xi_n^* = 1 + \frac{\log \#\,\mathrm{MCh}(\mathcal{R}_n)}{L_n}.$$

Then the sequence ξ_n^* converges to the solution of

$$\xi - \log \xi = 1 + k,$$

Hence by for sufficiently large n there exists a constant $A > 0$ independent of n such that

$$L_n \xi_n^* + \frac{\xi_n^*}{\xi_n^* - 1} < A L_n.$$

Theorem 1.2.10 implies

$$\mathbb{E}\,\tau(p\cdot\mathbf{1},\mathcal{R}_n) \le \frac{1}{p}\cdot A L_n.$$

Thus, the sequence $\frac{1}{n}\,\mathbb{E}\,\tau\,(p\cdot\mathbf{1},\mathcal{R}_n)$ is bounded from above and the limit in (3.26) exists. $\square$

The value of the limit can be computed exactly when $k=2$, see [27] for detailed analysis. In higher dimensions existence of the limit was discovered in many sources (i.e. [12, 32]), however, its value remains unknown.

3.8 Linear extensions of trees

Let $\#\mathcal{P}$ denote the number of elements in the poset $\mathcal{P}$ and let $\mathrm{Lin}\,(\mathcal{P})$ be the number of linear extensions of $\mathcal{P}$. Let $\mathcal{U}\,(i)$ stand for the set of elements of $\mathcal{P}$ that precede or equal i. Then the *hook-length formula* (see, for instance, [47]) for forests gives

$$\mathrm{Lin}\,(\mathcal{P}) = \frac{(\#\mathcal{P})!}{\prod_{i\in\mathcal{P}} \#\mathcal{U}\,(i)}. \tag{3.28}$$

This is the key ingredient for the proof of Theorem 1.2.11.

Proof. (Theorem 1.2.11) Let us consider the sample probability vector $\mathbf{p}$ for which

$$p_i = \lambda\,\#\mathcal{U}\,(i)\,, \quad i \in \mathcal{P}$$

where

$$\lambda = \left(\sum_{i\in\mathcal{P}} \#\mathcal{U}\,(i)\right)^{-1}.$$

Suppose that poset $\mathcal{P}$ has k minimal elements indexed by $1, 2, \ldots, k$. Then any linear extension has to begin with one of these elements, and the number of linear extensions which start with $1 \le i \le k$ is

$$\mathrm{Lin}\,(\mathcal{P} \setminus \{i\})\,.$$

From (3.28) we have

$$\mathrm{Lin}\,(\mathcal{P} \setminus \{i\}) = \frac{(\#\mathcal{P} - 1)!}{\prod_{j\in\mathcal{P}, j\neq i} \#\mathcal{U}\,(i)},$$

which gives

$$\frac{\mathrm{Lin}\,(\mathcal{P} \setminus \{i\})}{\mathrm{Lin}\,(\mathcal{P} \setminus \{j\})} = \frac{\#\mathcal{U}\,(i)}{\#\mathcal{U}\,(j)} = \frac{p_i}{p_j}$$

for all $1 \le i \le k$ and $1 \le j \le k$. Thus, we obtain the correct distribution of the first element in the linear extension.

However, after getting the first element we have to produce the first element of a linear extension in the remaining forest. By the same argument as above, the vector $\mathbf{p}$

will induce the correct distribution. Finally, the distribution on linear extensions will be uniform. □

Theorem 1.2.11 might not hold for posets that are not forests. The hook-length formula for the Young diagrams implies that we will need to change sample probabilities after collecting a few first elements in order to produce the uniform distribution on the linear extensions.

3.9 Higher moments of the continuous process and complete homogeneous symmetric polynomials

In some cases, we can compute the moments of the continuous process explicitly. Recall that the *complete homogeneous symmetric polynomial* $h_k(z_1, \ldots, z_n)$ is defined as a sum of all different monomials

$$z_1^{a_1} z_2^{a_2} \ldots z_n^{a_n}$$

whose degree $a_1 + \ldots + a_n$ is equal to k, i.e.

$$h_k(z_1, \ldots, z_n) = \sum_{a_1+\ldots+a_n=k} z_1^{a_1} \ldots z_n^{a_n}. \tag{3.29}$$

Lemma 3.9.1. *For posets $\mathcal{S}_n$ and $\mathcal{I}_n$ we have*

$$\mathbb{E}\, \tau_C^k(\mathbf{p}, \mathcal{S}_n) = k! \cdot h_k\left(\frac{1}{p_1}, \ldots, \frac{1}{p_n}\right) = k! \cdot h_k(r_1, \ldots, r_n). \tag{3.30}$$

$$\mathbb{E}\, \tau_C^k(\mathbf{p}, \mathcal{I}_n) = k! \sum_{J \in \mathbf{2}^n, \#J>0} \frac{(-1)^{\#J+1}}{\left(\sum_{i \in J} p_i\right)^k}. \tag{3.31}$$

In particular,

$$\mathbb{E}\, \tau_C^k(p \cdot \mathbf{1}, \mathcal{S}_n) = \frac{k!}{p^k} \cdot h_k(1, 1, \ldots, 1) = \frac{k!}{p^k} \cdot \binom{n+k-1}{k}, \tag{3.32}$$

$$\mathbb{E}\, \tau_C^k(p \cdot \mathbf{1}, \mathcal{I}_n) = \frac{k!}{p^k} \cdot h_k\left(1, \frac{1}{2}, \ldots, \frac{1}{n}\right). \tag{3.33}$$

Since the moments of the continuous CCP are convex (Theorem 1.2.6), we obtain the following fact.

Corollary 3.9.2. *For any non-negative integers n and k the function $h_k(x_1, \ldots, x_n)$ is convex on the domain*

$$\{(x_1, x_2, \ldots, x_n) \mid x_1 \geq 0, x_2 \geq 0, \ldots, x_n \geq 0\}.$$

Proof. For a chain $\mathcal{S}_n$ we obtain:

$$\begin{aligned}
\mathbb{E}\,\tau_C^k(\mathbf{p},\mathcal{S}_n) &= \int_0^\infty \cdots \int_0^\infty p_1 e^{-p_1 x_1} \dots p_n e^{-p_n x_n} (x_1+\ldots+x_n)^k \, dx_1 \ldots dx_n \\
&= \int_0^\infty \cdots \int_0^\infty \Bigg(p_1 e^{-p_1 x_1} \dots p_n e^{-p_n x_n} \\
&\quad \times \sum_{a_1+\ldots+a_n=k} \frac{k!}{a_1!\ldots a_n!} x_1^{a_1} \ldots x_n^{a_n} \Bigg) dx_1 \ldots dx_n \\
&= \sum_{a_1+\ldots+a_n=k} \frac{k!}{a_1!\ldots a_n!} \cdot \frac{a_1!}{p_1^{a_1}} \cdots \frac{a_n!}{p_n^{a_n}} \\
&= k! \cdot h_k\left(\frac{1}{p_1}, \ldots, \frac{1}{p_n}\right).
\end{aligned}$$

Here all a_i are assumed to be non-negative integers, and the summation is taken over distinct combinations $(a_1, \ldots, a_n)$.

In order to get (3.31), note that

$$\begin{aligned}
\mathbb{E}\,\tau_C^k(\mathbf{p},\mathcal{I}_n) &= \int_0^\infty \cdots \int_0^\infty p_1 e^{-p_1 x_1} \dots p_n e^{-p_n x_n} \left(\max_{1\le i\le n} x_i\right)^k dx_1 \ldots dx_n \\
&= \sum_{i=1}^n \int_0^\infty p_i e^{-p_i x_i} x_i^k \Bigg(\int_0^{x_i} \cdots \int_0^{x_i} p_1 e^{-p_1 x_1} \dots p_{i-1} e^{-p_{i-1} x_{i-1}} \\
&\quad \times p_{i+1} e^{-p_{i+1} x_{i+1}} \dots p_n e^{-p_n x_n} dx_1 \ldots dx_{i-1} dx_{i+1} \ldots dx_n \Bigg) dx_i \\
&= \sum_{i=1}^n \int_0^\infty p_i e^{-p_i x_i} x_i^k \prod_{j\ne i} \left(1 - e^{-p_j x_i}\right) dx_i.
\end{aligned}$$

Expanding the product and integrating with respect to x_i, we have

$$\begin{aligned}
\mathbb{E}\,\tau_C^k(\mathbf{p},\mathcal{I}_n) &= \sum_{i=1}^n \int_0^\infty p_i x_i^k \sum_{J\in \mathbf{2}^n, i\in J} (-1)^{\#J+1} \exp\left(-x_i \sum_{j\in J} p_j\right) dx_i \\
&= \sum_{i=1}^n p_i \sum_{J\in \mathbf{2}^n, i\in J} (-1)^{\#J+1} \frac{k!}{\left(\sum_{j\in J} p_j\right)^{k+1}} \\
&= k! \sum_{J\in \mathbf{2}^n, \#J>0} \frac{(-1)^{\#J+1}}{\left(\sum_{j\in J} p_j\right)^{k+1}} \cdot \sum_{j\in J} p_j,
\end{aligned}$$

and (3.31) follows.

Plugging in $p_i = p$, we get

$$\begin{aligned}
\mathbb{E}\,\tau_C^k\left(p\cdot\mathbf{1},\mathcal{I}_n\right) &= k!\sum_{J\in\mathbf{2}^n,\#J>0}\frac{(-1)^{\#J+1}}{(p\cdot\#J)^k}\\
&= k!\sum_{i=1}^{n}\binom{n}{i}\frac{(-1)^{i+1}}{i^k}\\
&= k!\cdot H_k\left(n\right).
\end{aligned}$$

Therefore, it remains to show that

$$H_k\left(n\right) = h_k\left(1,\frac{1}{2},\ldots,\frac{1}{n}\right).$$

For $k=1$ we have

$$\begin{aligned}
H_1\left(n\right) &= \sum_{i=0}^{n-1}\binom{n}{i+1}\frac{(-1)^i}{i+1}\\
&= \sum_{i=0}^{n-1}\binom{n}{i+1}(-1)^i\left(\int_0^1\xi^i d\xi\right)\\
&= \int_0^1\frac{-1}{\xi}\sum_{i=1}^{n}\binom{n}{i}(-\xi)^i\,d\xi\\
&= \int_0^1\frac{1-(1-\xi)^n}{\xi}d\xi.
\end{aligned}$$

Finally, substitution $\zeta = 1-\xi$ leads to

$$\begin{aligned}
H_1\left(n\right) &= \int_0^1\frac{1-\zeta^n}{1-\zeta}d\zeta\\
&= \int_0^1\left(1+\zeta+\zeta^2+\ldots+\zeta^{n-1}\right)d\zeta\\
&= 1+\frac{1}{2}+\ldots+\frac{1}{n}\\
&= h_1\left(1,\frac{1}{2},\ldots,\frac{1}{n}\right).
\end{aligned}$$

Let $k > 1$. In the similar way we obtain

$$\begin{aligned} H_k(n) &= \sum_{i=0}^{n-1} \binom{n}{i+1} (-1)^i \left(\int_0^1 \xi_1^i d\xi_1\right) \dots \left(\int_0^1 \xi_k^i d\xi_k\right) \\ &= \int_0^1 \dots \int_0^1 \frac{-1}{\xi_1 \dots \xi_n} \sum_{i=1}^{n} \binom{n}{i} (-\xi_1 \dots \xi_n)^i d\xi_1 \dots d\xi_k \\ &= \int_0^1 \dots \int_0^1 \frac{1-(1-\xi_1\dots\xi_k)^n}{\xi_1\dots\xi_k} d\xi_1 \dots d\xi_k. \end{aligned}$$

If $n = 1$, then

$$H_k(1) = 1 = h_k(1).$$

If $n > 1$, we have

$$\begin{aligned} H_k(n) &= \int_0^1 \dots \int_0^1 \sum_{i=0}^{n-1} (1-\xi_1\dots\xi_k)^i d\xi_1 \dots d\xi_k \\ &= H_k(n-1) + \int_0^1 \dots \int_0^1 (1-\xi_1\dots\xi_k)^{n-1} d\xi_1 \dots d\xi_k. \end{aligned}$$

Integration with respect to ξ_k gives

$$\begin{aligned} H_k(n) &= H_k(n-1) + \int_0^1 \dots \int_0^1 \frac{1-(1-\xi_1\dots\xi_{k-1})^n}{n\xi_1\dots\xi_{k-1}} d\xi_1 \dots d\xi_{k-1} \\ &= H_k(n-1) + \frac{1}{n} H_{k-1}(n). \end{aligned}$$

However, from the definition of complete symmetric functions we obtain

$$h_k(z_1, \dots, z_n) = h_k(z_1, \dots, z_{n-1}) + z_n h_{k-1}(z_1, \dots, z_n). \tag{3.34}$$

Therefore, quantities $H_k(n)$ and $h_k\left(1, \dots, \frac{1}{n}\right)$ coincide when either $k = 1$ or $n = 1$, and satisfy the same recurrence. Thus, they are equal, as desired.

The relation (3.32) follows directly from (3.30) and (3.34). □

Theorem 3.9.3. *For finite poset $\mathcal{P}$ and integer $k \geq 1$ we have*

$$\max_{(i_1,\dots,i_m)\in \mathrm{Ch}(\mathcal{P})} h_k(p_{i_1}, \dots, p_{i_m}) \leq \mathbb{E}\, \tau_C^k(\mathbf{p}, \mathcal{P}) \leq L^k \cdot k! H_k(M), \tag{3.35}$$

where M is the maximal size of an antichain in the poset, and

$$H_k(M) = \sum_{i=1}^{M} \binom{M}{i} \frac{(-1)^{i+1}}{i^k} = h_k\left(1, \frac{1}{2}, \ldots, \frac{1}{M}\right). \tag{3.36}$$

Proof. The theorem can be proved in the same way as we established Theorem 1.2.7. Indeed, let

$$c = (i_1, i_2, \ldots, i_m)$$

be a chain in the poset $\mathcal{P}$. Corollary 3.2.4 gives

$$\mathbb{E}\,\tau_C^k(c) \leq \mathbb{E}\,\tau_C^k(\mathcal{P}).$$

Now using Lemma 3.9.1 and taking the maximum over all chains, we obtain the lower bound.

With notations in the proof of Theorem 1.2.7, we get by Lemma 3.9.1 that

$$\mathbb{E}\,\tau_C^k(L \cdot \boldsymbol{\chi}(a)) \leq L^k \cdot k! H_k(M),$$

and convexity gives the upper bound. □

3.10 Integration over the chain polytope

3.10.1 Motivation

Let us consider the continuous CCP on the poset $\mathcal{P}$ that consists of three elements a_1, a_2 and a_3 so that

$$a_1 \prec a_2 \prec a_3.$$

Assuming the corresponding rates to be p_1, p_2 and p_3 respectively, let us compute $\mathbf{P}(\tau_C < t)$. Since the only maximal chain of $\mathcal{P}$ is (a_1, a_2, a_3), then the chain polytope of $\mathcal{P}$ is given by the following inequalities:

$$x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0,$$
$$x_1 + x_2 + x_3 \leq 1,$$

where variable x_i corresponds to element a_i $(1 \leq i \leq 3)$. Therefore, formula (3.12) gives

$$\mathbf{P}(\tau < t) = \int_0^t dx_1 \int_0^{t-x_1} dx_2 \int_0^{t-x_1-x_2} dx_3\ p_1 p_2 p_3 \exp\left(-p_1 x_1 - p_2 x_2 - p_3 x_3\right).$$

Evaluating the integral, we get

$$\mathbf{P}(\tau < t) = 1 - \frac{e^{-p_1 t} p_2 p_3}{(p_1 - p_2)(p_1 - p_3)} - \frac{e^{-p_2 t} p_1 p_3}{(p_2 - p_1)(p_2 - p_3)} - \frac{e^{-p_3 t} p_1 p_2}{(p_3 - p_1)(p_3 - p_2)}. \quad (3.37)$$

The Markov chain approach used in Section 3.3 also gives the same answer. So does the Brion's formula ([8]) which is a standard tool to integrate exponents of a linear function over a polytope.

Although relation (3.37) is not complicated, it not trivial why there is no singularity if $p_1 \to p_2$. A possible way to cope with these problems is to compute a power series of $\mathbf{P}(\tau < t)$ with respect to t. We leave this technical procedure to the reader and give the final answer

$$\mathbf{P}(\tau < t) = p_1 p_2 p_3 \left(\frac{t^3}{3!} - \frac{t^4}{4!} \cdot h_1(p_1, p_2, p_3) + \frac{t^5}{5!} \cdot h_2(p_1, p_2, p_3) - \dots \right), \tag{3.38}$$

where h_k are the complete homogenous symmetric functions introduced in Section 3.9. Now the problem with singularity is resolved, since each coefficient is symmetric in variables p_1, p_2, p_3, and it is not difficult to prove that the series converges to a C^∞ function. See formula (3.44) for the generalization of (3.38).

3.10.2 Proofs

As we mentioned in the introduction, representation (3.38) is a consequence of Theorem 1.2.12. In order to prove it, we need an auxiliary fact.

Lemma 3.10.1. *For any n element poset $\mathcal{P}$ and n-dimensional vector*

$$\mathbf{m} = (m_1, m_2, \dots, m_n)$$

with positive integer entries we have

$$\int_{C(\mathcal{P})} x_1^{m_1-1} \dots x_n^{m_n-1} \, dx_1 \dots dx_n = \frac{(m_1-1)! \dots (m_n-1)!}{|\mathbf{m}|!} e(\mathbf{m}). \tag{3.39}$$

We postpone the proof of the lemma for a moment, and show how Theorem 1.2.12 follows.

Proof. (Theorem 1.2.12) Relation (3.39) implies

$$\int_{tC(\mathcal{P})} x_1^{m_1-1} \dots x_n^{m_n-1} \, dx_1 \dots dx_n = \frac{t^{|\mathbf{m}|}}{|\mathbf{m}|!} (m_1-1)! \dots (m_n-1)! \, e(\mathbf{m}),$$

or

$$\int_{tC(\mathcal{P})} \frac{a_1 \ldots a_n}{(|\mathbf{m}| - n)!} \frac{(|\mathbf{m}| - n)!}{(m_1 - 1)! \ldots (m_n - 1)!} (a_1 x_1)^{m_1 - 1} \ldots (a_n x_n)^{m_n - 1} \, dx_1 \ldots dx_n = \frac{t^{|\mathbf{m}|}}{|\mathbf{m}|!} e(\mathbf{m}) \, a_1^{m_1} \ldots a_n^{m_n}. \quad (3.40)$$

Note that

$$\sum_{|\mathbf{m}|=k} \frac{(|\mathbf{m}| - n)!}{(m_1 - 1)! \ldots (m_n - 1)!} (a_1 x_1)^{m_1 - 1} \ldots (a_n x_n)^{m_n - 1} = (a_1 x_1 + \ldots + a_n x_n)^{k-n},$$

where the sum is taken over all n-dimensional vectors $\mathbf{m}$ with positive integer entries that add up to k. Taking this sum of both sides in (3.40), we obtain

$$a_1 \ldots a_n \int_{tC(\mathcal{P})} \frac{(a_1 x_1 + \ldots + a_n x_n)^{k-n}}{(k-n)!} dx_1 \ldots dx_n = \sum_{|\mathbf{m}|=k} \frac{t^{|\mathbf{m}|}}{|\mathbf{m}|!} e(\mathbf{m}) \, a_1^{m_1} \ldots a_n^{m_n}.$$

It is not difficult to see that k is integer and $k \geq n$. Thus, taking the summation of the integrand over all values of k, we have

$$\sum_{k=n}^{\infty} \frac{(a_1 x_1 + \ldots + a_n x_n)^{k-n}}{(k-n)!} = \exp(a_1 x_1 + \ldots + a_n x_n),$$

and the theorem follows. □

Before we proceed with the proof of Lemma 3.10.1, let us consider the case

$$m_1 = m_2 = \ldots = m_n = 1.$$

Then formula (3.39) becomes

$$\int_{C(\mathcal{P})} dx_1 dx_2 \ldots dx_n = \frac{1}{n!} e(1, 1, \ldots, 1) = \frac{1}{n!} e(\mathbf{1}). \quad (3.41)$$

From the definition of $e(\cdot)$, we obtain that $e(\mathbf{1})$ is the number of linear extensions of the poset $\mathcal{P}$. Indeed, any map f corresponding to vector $\mathbf{1}$ is a bijection between elements of $\mathcal{S}_n$ and $\mathcal{P}$. Thus, map f^{-1} is an order-preserving bijection from $\mathcal{P}$ to $\mathcal{S}_n$, therefore, f^{-1} is a linear extension. Thus, in this case we obtain the result given in [48] concerning the volume of the chain polytope.

Proof. (Lemma 3.10.1) For the poset $\mathcal{P}$ on elements

$$s_1, s_2, \ldots, s_n$$

and a vector $\mathbf{m}$ let us consider a poset $\mathcal{P}^{\mathbf{m}}$ constructed as follows. The elements of $\mathcal{P}^{\mathbf{m}}$ are

$$s_1^1, s_1^2, \ldots, s_1^{m_1}, s_2^1, s_2^2, \ldots, s_2^{m_2}, \ldots, s_n^1, s_n^2, \ldots, s_n^{m_n},$$

and

$$s_i^a \prec s_j^b \text{ in } \mathcal{P}^{\mathbf{m}} \quad \Longleftrightarrow \quad s_i \prec s_j \text{ in } \mathcal{P}$$

for any $1 \leq i, j \leq n$ and $1 \leq a \leq m_i$, $1 \leq b \leq m_j$. In other words, instead of a single element s_i we use m_i incomparable copies of it with induced partial order. Note that $\mathcal{P}^{\mathbf{m}}$ has $|\mathbf{m}|$ elements.

Now consider a linear extension of $\mathcal{P}^{\mathbf{m}}$, i.e. a bijection $\ell \colon \mathcal{P}^{\mathbf{m}} \to \mathcal{S}_{|\mathbf{m}|}$. Let a map $\pi \colon \mathcal{P}^{\mathbf{m}} \to \mathcal{P}$ be defined as follows:

$$\pi\left(s_i^a\right) = s_i \quad \text{for all } 1 \leq i \leq n,\ 1 \leq a \leq m_i.$$

Then $f = \pi\ell^{-1}$ is a map from $\mathcal{S}_{|\mathbf{m}|}$ to $\mathcal{P}$ which maps exactly m_i elements of $\mathcal{S}_{|\mathbf{m}|}$ to element s_i. Also, if $x, y \in \mathcal{S}_{|\mathbf{m}|}$ and $x \prec y$, then $\ell^{-1}(x) \not\succ \ell^{-1}(y)$, otherwise ℓ is not a linear extension. Thus,

$$f(x) \not\succ f(y),$$

which means that f does not break the partial order. This shows that any linear extension ℓ gives a map $f \colon \mathcal{S}_{|\mathbf{m}|} \to \mathcal{P}$ corresponding to vector $\mathbf{m}$.

Note that π does not distinguish different copies of the same element in $\mathcal{P}$. Hence for any map f there are exactly $m_1!m_2!\ldots m_n!$ linear extensions of $\mathcal{P}^{\mathbf{m}}$ that result to f. Therefore, the number of linear extensions of $\mathcal{P}^{\mathbf{m}}$ is

$$m_1!\, m_2! \dots m_n! \cdot e\left(m_1, m_2, \dots, m_n\right) = m_1!\, m_2! \dots m_n! \cdot e\left(\mathbf{m}\right).$$

Then relation (3.41) yields for the volume of the chain polytope

$$\int_{\mathcal{C}(\mathcal{P}^{\mathbf{m}})} dx_1^1 \dots dx_1^{m_1} \dots dx_n^1 \dots dx_n^{m_n} = \frac{m_1!\, m_2! \dots m_n!}{|\mathbf{m}|!}\, e\left(\mathbf{m}\right). \tag{3.42}$$

Here upper indices indicate the number of corresponding copies of elements in $\mathcal{P}$.

Let for brevity

$$\mathbf{x}^{\mathbf{m}} = \left(x_1^1, \dots, x_1^{m_1}, \dots x_n^1, \dots, x_n^{m_n}\right)^T,$$

and let

$$\mathbf{x} = \left(x_1, \dots, x_n\right)^T, \quad \text{where } x_i = \max_{1 \le a \le m_i} x_i^a \quad \text{for all } 1 \le i \le n.$$

Note that any chain in $\mathcal{P}^{\mathbf{m}}$ cannot contain more than one copy of element s_i. This shows that

$$s_1' \prec s_2' \prec \dots \prec s_k'$$

is a chain in $\mathcal{P}^{\mathbf{m}}$ if and only if

$$\pi\left(s_1'\right) \prec \pi\left(s_2'\right) \prec \dots \prec \pi\left(s_n'\right)$$

is a chain in $\mathcal{P}$. Hence,

$$\mathbf{x}^{\mathbf{m}} \in \mathcal{C}\left(\mathcal{P}^{\mathbf{m}}\right) \iff \mathbf{x} \in \mathcal{C}\left(\mathcal{P}\right). \tag{3.43}$$

Now consider i.i.d. random variables

$$X_1^1, \dots, X_1^{m_1}, X_2, \dots, X_2^{m_2}, \dots, X_n^1, \dots, X_n^{m_n}$$

that have the uniform distribution on $[0,1]$, and let

$$X_i = \max_{1 \le a \le m_i} X_i^a \quad \text{for all } 1 \le i \le n.$$

Then it is easy to see that

$$\int_{\mathcal{C}(\mathcal{P}^{\mathbf{m}})} dx_1^1 \dots dx_1^{m_1} \dots dx_n^1 \dots dx_n^{m_n} = \mathbf{P}\left(\mathbf{X}^{\mathbf{m}} \in \mathcal{C}\left(\mathcal{P}^{\mathbf{m}}\right)\right),$$

where $\mathbf{X}^{\mathbf{m}} = \left(X_1^1, \dots, X_1^{m_1}, \dots X_n^1, \dots, X_n^{m_n}\right)^T$. Equivalence (3.43) implies

$$\mathbf{P}\left(\mathbf{X}^{\mathbf{m}} \in \mathcal{C}\left(\mathcal{P}^{\mathbf{m}}\right)\right) = \mathbf{P}\left(\mathbf{X} \in \mathcal{C}\left(\mathcal{P}\right)\right),$$

where $\mathbf{X} = \left(X_1, \dots, X_n\right)^T$. It remains to note that the density of X_i is

$$m_i x_i^{m_i - 1} dx_i,$$

therefore

$$\mathbf{P}\left(\mathbf{X} \in \mathcal{C}\left(\mathcal{P}\right)\right) = \int_{\mathcal{C}(\mathcal{P})} m_1 x_1^{m_1-1} dx_1 \dots m_n x_n^{m_n-1} dx_n.$$

Plugging this into (3.42), we get

$$\int_{\mathcal{C}(\mathcal{P})} m_1 x_1^{m_1-1} dx_1 \dots m_n x_n^{m_n-1} dx_n = \frac{m_1!\, m_2! \dots m_n!}{|\mathbf{m}|!} e\left(\mathbf{m}\right),$$

and the lemma follows. □

3.10.3 Examples

Totally ordered posets

Let

$$\mathcal{P} = \mathcal{S}_n = \left\{s_1 \prec s_2 \prec \dots \prec s_n\right\}.$$

Then for any vector $\mathbf{m}$ we obtain

$$e(\mathbf{m}) = 1.$$

Indeed, the smallest m_1 elements of $\mathcal{S}_{|\mathbf{m}|}$ should be mapped into s_1, otherwise we would break the partial order. Also the next m_2 elements should be mapped into s_2, and so on.

Therefore, Theorem 1.2.12 implies

$$\mathbf{P}\left(\tau_C\left(\mathbf{p}, \mathcal{S}_n\right) < t\right) = \sum_{\mathbf{m}} \frac{t^{|\mathbf{m}|}}{|\mathbf{m}|!} p_1^{m_1} \dots p_n^{m_n}.$$

Recalling the definition of complete homogeneous symmetric functions, we get

$$\mathbf{P}\left(\tau_C\left(\mathbf{p}, \mathcal{S}_n\right) < t\right) = p_1 \dots p_n \left(\frac{t^n}{n!} - \frac{t^{n+1}}{(n+1)!} h_1(\mathbf{p}) + \frac{t^{n+2}}{(n+2)!} h_2(\mathbf{p}) - \dots\right), \quad (3.44)$$

where $h_k(\mathbf{p})$ stands for $h_k(p_1, \dots, p_n)$. Obviously, formula (3.38) is a special case of (3.44).

Sum of i.i.d. exponential variables

Let now $\mathbf{p} = p \cdot \mathbf{1}$. In the proof of Lemma 3.9.1, we have shown that

$$h_k(\mathbf{1}) = \binom{n+k-1}{k},$$

which leads to

$$\begin{aligned} \mathbf{P}\left(\tau\left(\mathbf{1}, \mathcal{S}_n\right) < t\right) &= \sum_{k=n}^{\infty} \frac{(pt)^k}{k!} (-1)^{k-n} h_{k-n}(\mathbf{1}) \\ &= \sum_{k=n}^{\infty} \frac{(pt)^k}{k!} (-1)^{k-n} \binom{k-1}{n-1}. \end{aligned}$$

Expanding the binomial coefficients, we obtain

$$\begin{aligned}\mathbf{P}\left(\tau\left(\mathbf{1},\mathcal{S}_n\right)<t\right) &= \sum_{k=n}^{\infty}(-1)^{k-n}\frac{(pt)^k}{k!}\frac{(k-1)!}{(k-n)!\,(n-1)!}\\ &= \frac{1}{(n-1)!}\sum_{k=n}^{\infty}\frac{(-1)^{k-n}}{(k-n)!}\frac{(pt)^k}{k}.\end{aligned}$$

Finally, using the fact that

$$\int_0^{pt}\xi^{k-1}d\xi=\frac{(pt)^k}{k},$$

we get

$$\begin{aligned}\mathbf{P}\left(\tau\left(\mathbf{1},\mathcal{S}_n\right)<t\right) &= \frac{1}{(n-1)!}\int_0^{pt}\xi^{n-1}\sum_{k=n}^{\infty}\frac{(-1)^{k-n}}{(k-n)!}\xi^{k-n}d\xi\\ &= \frac{1}{(n-1)!}\int_0^{pt}\xi^{n-1}e^{-\xi}d\xi.\end{aligned}$$

However, the random variable $\tau\left(\mathbf{1},\mathcal{S}_n\right)$ can be represented as $X_1+X_2+\ldots+X_n$, where X_i are i.i.d. exponential random variables with parameter p. Thus, we obtain the following classical fact known in the probability theory.

Proposition 3.10.2. *For i.i.d. exponential random variables $X_1, X_2, \ldots, X_n$ with parameter p we have*

$$\mathbf{P}\left(X_1+X_2+\ldots+X_n<t\right)=\frac{1}{(n-1)!}\int_0^{pt}\xi^{n-1}e^{-\xi}d\xi. \tag{3.45}$$

Sum of i.i.d. geometric random variables

For the discrete coupon collector's process on $\mathcal{S}_n$ with sample probabilities given by vector $p\cdot\mathbf{1}$ the stopping time τ_D can be represented as a sum of n i.i.d. random variables $G_1,\ldots,G_n$ that have the geometric distribution with parameter p. Lemma 1.2.2 and Proposition 3.10.2 lead to the following result.

Proposition 3.10.3. *For i.i.d. geometric random variables $G_1, G_2, \ldots, G_n$ with pa-*

rameter p we have

$$\mathbf{P}(G_1+\ldots+G_n\le m)=\begin{cases} n\binom{m}{n}\int_0^p \xi^{n-1}(1-p)^{m-n}\,d\xi, & \text{if } m\ge n,\\ 0, & \text{if } m<n.\end{cases} \tag{3.46}$$

Proof. Lemma 1.2.2 gives that $\frac{1}{m!}\mathbf{P}(\tau_D(p\cdot\mathbf{1},\mathcal{S}_n)\le m)$ is a coefficient against t^m in the power series of

$$e^t\,\mathbf{P}(\tau_C(p\cdot\mathbf{1},\mathcal{S}_n)\le t).$$

Formula (3.44) shows that the smallest power of t in the power series is n, therefore,

$$\mathbf{P}(\tau_D(p\cdot\mathbf{1},\mathcal{S}_n)\le m)=0\quad\text{if } m<n.$$

Let us assume $m\ge n$. Expanding $e^t\,\mathbf{P}(\tau_C(p\cdot\mathbf{1},\mathcal{S}_n)\le t)$, we obtain that

$$\frac{1}{m!}\mathbf{P}(\tau_D(p\cdot\mathbf{1},\mathcal{S}_n)\le m)=\frac{1}{(n-1)!}\sum_{k=n}^{m}\frac{(-1)^{k-n}p^k}{k\,(k-n)!\,(m-k)!}$$

By changing the index of summation, we get

$$\begin{aligned}\mathbf{P}(\tau_D\le m) &= \frac{m!}{(n-1)!}\sum_{k=0}^{m-n}\frac{(-1)^k p^{k+n}}{(k+n)\,k!\,(m-k-n)!}\\ &= \frac{m!}{(n-1)!\,(m-n)!}\sum_{k=0}^{m-n}(-1)^k\frac{p^{k+n}}{k+n}\binom{m-n}{k}.\end{aligned}$$

Substituting $\int_0^p \xi^{k+n-1}d\xi$ for $\frac{p^{k+n}}{k+n}$, we have

$$\begin{aligned}\mathbf{P}(\tau_D\le m) &= \frac{m!}{(n-1)!\,(m-n)!}\int_0^p \xi^{n-1}\sum_{k=0}^{m-n}(-1)^k\binom{m-n}{k}\xi^k d\xi\\ &= n\binom{m}{n}\int_0^p \xi^{n-1}(1-\xi)^{m-n}\,d\xi.\end{aligned}$$

This accomplishes the proof for $m\ge n$. □

Sequences of elements

Let us consider the discrete coupon collector's process on a poset $\mathcal{P}$ on n elements whose sample probabilities are equal to $p = \frac{1}{n}$. Since all sample probabilities add up to 1, at each step we draw a coupon. Let

$$c_1, c_2, \ldots, c_\ell$$

be the sequence of coupons drawn at first ℓ trials. It is not hard to see that we complete the collection if among these ℓ elements we can find a subsequence of distinct elements

$$c_{i_1}, c_{i_2}, \ldots, c_{i_n}$$

such that

$$i_1 < i_2 < \ldots < i_n,$$

and for any $1 \leq a < b \leq n$ either $c_{i_a} \prec c_{i_b}$ or elements c_{i_a} and c_{i_b} are incomparable. In other words, we complete the collection if the sequence of the coupon drawn at first ℓ trials contains at least one linear extension of the poset $\mathcal{P}$ as a subsequence. Denote the number of such sequences by $W_\ell(\mathcal{P})$. The choice of the sample probabilities implies

$$\mathbf{P}\left(\tau_D\left(p \cdot \mathbf{1}, \mathcal{P}\right) \leq \ell\right) = \frac{W_\ell(\mathcal{P})}{n^\ell}.$$

Having this, we can use Theorem 1.2.12 and Lemma 1.2.2 to evaluate W_ℓ.

Proposition 3.10.4. *For an n element poset we have*

$$W_\ell(\mathcal{P}) = \begin{cases} \sum_{k=n}^{\ell} (-1)^{k-n} \binom{\ell}{k} E_k n^{\ell-k}, & \text{if } \ell \geq n, \\ 0, & \text{if } \ell < n. \end{cases} \tag{3.47}$$

Here

$$E_k = \sum_{|\mathbf{m}|=k} e(\mathbf{m}),$$

where $\mathbf{m}$ *are* n*-dimensional vectors with positive integer entries.*

Example 3.10.5. Let us consider a poset $\mathcal{P}$ on three elements a, b and c ordered as follows:

$$a \prec c, \quad b \prec c.$$

The poset $\mathcal{P}$ has two linear extensions: abc and bac, therefore, $e(1,1,1) = 2$. It is also easy to see that

$$\begin{aligned} e(2,1,1) = 3 &\ : \ aabc, \quad abac, \quad baac, \\ e(1,2,1) = 3 &\ : \ abbc, \quad babc, \quad bbac, \\ e(1,1,2) = 2 &\ : \ abcc, \quad bacc. \end{aligned}$$

Here $aabc$ indicates that the corresponding map from $\mathcal{S}_4$ maps the first two elements to element a, the third element — to b, and the fourth one is mapped to c.

Thus, we have

$$\begin{aligned} E_3 &= e(1,1,1) = 2, \\ E_4 &= e(2,1,1) + e(1,2,1) + e(1,1,2) = 8. \end{aligned}$$

Formula (3.47) gives

$$W_4(\mathcal{P}) = \binom{4}{3} E_3 \cdot 3 - \binom{4}{4} E_4 = 12E_3 - E_4 = 16,$$

and the corresponding sequences are (linear extensions are underlined)

$$\begin{aligned} &\underline{abc}a, \quad \underline{abc}b, \quad \underline{abc}c, \quad a\underline{bac}, \\ &\underline{ab}b\underline{c}, \quad a\underline{abc}, \quad \underline{a}c\underline{bc}, \quad b\underline{abc}, \\ &c\underline{abc}, \quad \underline{bac}a, \quad \underline{bac}b, \quad \underline{bac}c, \\ &\underline{ba}a\underline{c}, \quad b\underline{bac}, \quad \underline{b}c\underline{ac}, \quad c\underline{bac}. \end{aligned}$$

Proof. It is obvious that $W_\ell(\mathcal{P}) = 0$ if $\ell < n$, since any linear extension consists of n elements. Thus, we can assume $\ell \geq n$.

Plugging $a_i = -\frac{1}{n}$ into (1.16), we obtain from (3.12) that

$$\mathbf{P}(\tau_C \le t) = \sum_{\mathbf{m}} \frac{t^{|\mathbf{m}|}}{|\mathbf{m}|!} \frac{(-1)^{|\mathbf{m}|-n}}{n^{|\mathbf{m}|}} e(\mathbf{m}).$$

Combining the terms with the same value of $|\mathbf{m}|$, we have

$$\mathbf{P}(\tau_C \le t) = \sum_{k=n}^{\infty} \frac{t^k}{k!} \frac{(-1)^{k-n}}{n^k} E_k.$$

Taking the coefficient against t^ℓ in the power series of $e^t \mathbf{P}(\tau_C \le t)$, by Lemma 1.2.2 we get that

$$\frac{1}{\ell!} \mathbf{P}(\tau_D \le \ell) = \sum_{k=n}^{\ell} \frac{(-1)^{k-n}}{n^k} \frac{E_k}{k!\,(\ell-k)!}.$$

This leads to

$$W_\ell(\mathcal{P}) = n^\ell \mathbf{P}(\tau_D \le \ell) = \sum_{k=n}^{\ell} \frac{(-1)^{k-n} n^\ell}{n^k} \frac{\ell!}{k!\,(\ell-k)!} E_k,$$

and the proposition follows. □

Remark 3.10.6. Formula (3.47) looks like a typical Inclusion-Exclusion summation. However, we do not know how to apply the Inclusion-Exclusion principle in order to compute $W_\ell(\mathcal{P})$ directly.

Bibliography

[1] D. Aharonov, A. Ambainis, J. Kempe and U. Vazirani, Quantum walks on graphs, *Proc. 33rd STOC* (2001), 50–59.

[2] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error, *Proc. 29th STOC* (1997), 176–188.

[3] N. Ailon and B. Chazelle, Approximate Nearest Neighbors and the Fast Johnson-Lindenstrauss Transform, *Proc. 38th STOC* (2006), 557–563.

[4] D. Aldous, Random walks on finite groups and rapidly mixing Markov chains, in *Lecture Notes in Math.* **986** (1983), 243–297.

[5] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, monograph in preparation, available at `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

[6] D. Aldous, L. Lovász and P. Winkler, Mixing times for uniformly ergodic Markov chains, *Stochastic Process. Appl.* **71** (1997), No. 2, 165–185.

[7] S. Boneh and M. Hofri, The coupon-collector problem revisited — a survey of engineering problems and computational methods, *Stochastic Models* **13** (1997), 39–66.

[8] M. Brion, Points entiers dans les polyedres convexes, *Annales scientifiques de l'Ecole Normale Superieure* **21** (1988), 653–663.

[9] R. Bubley and M. E. Dyer, Path Coupling: A Technique for Proving Rapid Mixing in Markov Chains, *Proc. 38th FOCS* (1997), 223–231.

[10] E. A. Carlen, M. C. Carvalho and M. Loss, Determination of the spectral gap for Kac's master equation and related stochastic evolution, *Acta Math.* **191** (2003), No. 1, 1–54.

[11] R. Caron, M. Hlynka and J. McDonald, On the best-case performance of probabilistic methods for detecting necessary constraints, *Technical report WMSR-88-02, Dept. of Math. and Stats, Univ. of Windsor* (1988).

[12] J. T. Cox, A. Gandolfi, P. Griffin and H. Kesten, Greedy lattice animals I: Upper bounds. Ann. Appl. Probab. **3** (1993), 1151–1169.

[13] P. Diaconis and L. Saloff-Coste, Bounds for Kac's master equation, *Comm. Math. Phys.* **209** (2000), 729–755.

[14] W. Doeblin, Exposé de la theorie des chaînes simples constantes de Markov à un nombre fini d'etats (in French), *Rev. Math. de l'Union Interbalkanique* **2** (1933), 77–105.

[15] R. M. Dudley, *Real Analysis and Probability*, Cambridge University Press, 2002.

[16] M. E. Dyer, L. A. Goldberg, C. S. Greenhill, M. Jerrum and M. Mitzenmacher, An extension of path coupling and its application to the Glauber dynamics for graph colourings, *SIAM J. Comput.* **30** (2001), 1962–1975.

[17] P. Erdös and A. Rényi, On a classical problem of probability theory, *Magyar Tud. Akad. Mat. Kutató Int. Közl* **6** (1961), 215–220.

[18] W. Feller, *An introduction to probability theory and applications*, Vol. 1, Wiley, 1968.

[19] R. L. Graham, D. E. Knuth and O. Patashnik, *Concrete Mathematics,* 2nd edition, Addison-Wesley, 1998.

[20] W. Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika* **57** (1970), 97–109.

[21] T. P. Hayes and E. Vigoda, Coupling with the stationary distribution and improved sampling for colorings and independent sets, *Ann. Appl. Probab.* **16** (2006), 1297–1318.

[22] L. Holst, On birthday, collector's, occupancy and other classical urn problems, *Intern. Stat. Rev.* **54** (1986), 15–27.

[23] L. Holst, Extreme value distribution for random coupon collector and birthday problems, *Extremes* **4 (2)** (2001), 129–145.

[24] A. Hurwitz, Über die erzengung der invarianten durch integration, *Nach. Gesell. Wissen* (1897), 71–90. Reprinted in A. Hurwitz, *Mathematische Werke* **2** (1963), 546–564.

[25] E. Janvresse, Bounds on Semigroups of Random Rotations on $SO(n)$, *Theory Probab. Appl.* **47** (2003), 526–532.

[26] M. Jerrum and A. Sinclair, Approximating the Permanent, *SIAM J. Comput.* **18** (1989), 1149–1178.

[27] K. Johansson, Shape fluctuations and random matrices, *Communications in Math. Phys.* **209** (2000), 437–476.

[28] N. Johnson and S. Kotz, *Urn models and their applications. An approach to modern discrete probability theory,* Wiley, 1977.

[29] M. Kac, Probability and Related Topics in Physical Science, Interscience, New York, 1959.

[30] T. Lindvall, *Lectures on the Coupling Method,* New York, Wiley, 1992.

[31] L. Lovász and P. Winkler, Mixing times, in *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.* **41** (1998), 85–133.

[32] J. B. Martin, Linear growth for greedy lattice animals, *Stochastic Process. Appl.* **98** (2002), 43–66.

[33] S. Martinez, Some bounds on the coupon collector problem, *Rand. Struct. and Alg.* **25** (2004), 208–226.

[34] D. Maslin, The eigenvalues of Kac's master equation, *Math. Z.* **243** (2003), 291–331.

[35] H. McKean, Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas, *Arch. Rational Mech. Anal.* **2** (1966), 343–367.

[36] R. Montenegro and P. Tetali, Mathematical Aspects of Mixing Times in Markov Chains, *Foundations and Trends in Theoretical Computer Science*, to appear.

[37] R. Motwani and P. Raghavan, *Randomized Algorithms,* Cambrige University Press, 1995.

[38] I. Pak, Using stopping times to bound mixing times, *Proc. 10th SODA* (1999), 953–954.

[39] J. W. Pitman, On coupling of Markov chains, *Probab. Theory Rel. Fields* **35** (1976), 315–322.

[40] U. Porod, The cut-off phenomenon for random reflections, *Ann. Probab.* **24** (1996), 74–96.

[41] D. Randall and P. Winkler, Mixing points on an interval, *Proc. 7th ALENEX* (2005), 218–221.

[42] J. S. Rosenthal, Random Rotations: Characters and Random Walks on $SO(N)$, *Ann. Probab.* **22** (1994), 398–423.

[43] L. Saloff-Coste, Lectures on finite Markov chains, in *Lecture Notes in Math.* **1665**, Springer, (1997), 301–413.

[44] L. Saloff-Coste, Random walks on finite groups, in *Enc. Math. Sci.* **110**, Springer (2004), 263–346.

[45] S. Shioda, Some upper and lower bounds on the coupon collector problem, *J. of Comp. and Appl. Math.* **200** (2007), 154–167.

[46] P. W. Shor, Quantum computing, *Proc. ICM Berlin*, Doc. Math. **1** (1998), 467–486.

[47] R. Stanley, *Enumerative Combinatorics*, Cambridge University Press, 1999.

[48] R. Stanley, Two poset polytopes, *Discrete Comput. Geom.* **1** (1986), 9–23.

[49] C. Villani, Cercignani's conjecture is sometimes true and always almost true, *Comm. Math. Phys.* **234** (2003), 455–490.

[50] E. T. Whittaker and G. N. Watson, *A Course in Modern Analysis*, 4th edition, Cambridge University Press, Cambridge, England, 1990.