Phase transition for parking blocks, Brownian excursion and coalescence

Philippe Chassaing, Guy Louchard

Phase transition for parking blocks, Brownian excursion and coalescence[1]

P. Chassaing [2] & G. Louchard [3]

Abstract. In this paper, we consider hashing with linear probing for a hashing table with m places, n items ($n < m$), and $\ell = m - n$ empty places. For a non computer science-minded reader, we shall use the metaphore of n cars parking on m places: each car c_i chooses a place p_i at random, and if p_i is occupied, c_i tries successively $p_i + 1$, $p_i + 2$, until it finds an empty place. Pittel [42] proves that when ℓ/m goes to some positive limit $\beta < 1$, the size $B_1^{m,\ell}$ of the largest block of consecutive cars satisfies $2(\beta - 1 - \log \beta)B_1^{m,\ell} = 2 \log m - 3 \log \log m + \Xi_m$, where Ξ_m converges weakly to an extreme-value distribution. In this paper we examine at which level for n a phase transition occurs between $B_1^{m,\ell} = o(m)$ and $m - B_1^{m,\ell} = o(m)$. The intermediate case reveals an interesting behaviour of sizes of blocks, related to the standard additive coalescent in the same way as the sizes of connected components of the random graph are related to the multiplicative coalescent.

Key words. Hashing with linear probing, parking, Brownian excursion, empirical processes, coalescence.

A.M.S. Classification. 60C05, 60J65, 60F05, 68P10, 68R05.

[1]Random Structures Algorithms 21 (2002), no. 1, 76-119.
[2]Institut Elie Cartan, INRIA, CNRS and Université Henri Poincaré,
BP 239, 54 506 Vandoeuvre Cedex, France.
chassain@iecn.u-nancy.fr
[3]Université Libre de Bruxelles, Département d'Informatique,
Campus Plaine, CP 212, Bvd du Triomphe, 1050 Bruxelles, Belgium.
louchard@ulb.ac.be

1 Main results

1.1 Emergence of a giant block

We consider hashing with linear probing for a hashing table with a set of m places, $\{1, 2, ..., m\}$, n items $\{c_1, c_2, ..., c_n\}$, and $\ell = m - n$ empty places ($\ell > 0$). Hashing with linear probing is a fundamental object in analysis of algorithms: its study goes back to the 1960's [29, 31] and is still active [2, 23, 30, 42]. For a non computer science-minded reader, we shall use, all along the paper, the metaphore of n cars parking on m places, leaving ℓ places empty: each car c_i chooses a place p_i at random, and if p_i is occupied, c_i tries successively $p_i + 1$, $p_i + 2$, until it finds an empty place. We use the convention that place $m + 1$ is also place 1.

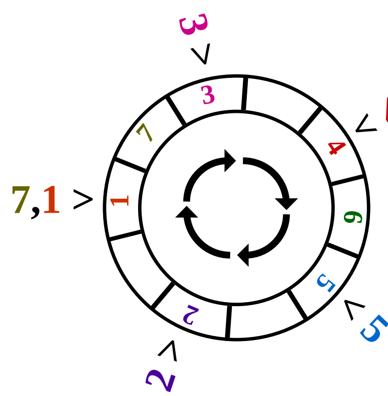

Figure 1: An example where $(m, n) = (10, 7)$ and $B^{10,3} = (3, 3, 1, 0, 0, \dots)$

Under the name of parking function, hashing with linear probing has been and is still studied by combinatorists [25, 26, 45, 49, 50, 51]. Section 4 of [23] contains nice developments on the connections between parking functions and many other combinatorial objects. In this paper, we use mainly a - maybe less exploited - connection between parking functions and *empirical processes* of mathematical statistics (see also [15, 39]) .

Let $B_k^{m,\ell}$ denote the size of the k^{th} largest block of consecutive cars, and let $B^{m,\ell} = (B_k^{m,\ell})_{k\geq 1}$ be the decreasing sequence of sizes of blocks, ended by an infinite sequence of 0's. Pittel [42] proves that when ℓ/m goes to some positive limit β, $B_1^{m,\ell}$ satisfies

$$B_1^{m,\ell} = \frac{2\log m - 3\log\log m + \Xi_m}{2(\beta - 1 - \log\beta)},$$

where Ξ_m converges weakly to an extreme-value distribution. This paper is concerned with what we would call the "emergence of a giant block", by reference to the emergence of a giant component [4, 9, 14, 22, 28]. We have:

Theorem 1.1 *For m and n going jointly to $+\infty$*

(i) if $\sqrt{m} = o(\ell)$, $B_1^{m,\ell}/m \xrightarrow{P} 0$;

(ii) if $\ell = o(\sqrt{m})$, $B_1^{m,\ell}/m \xrightarrow{P} 1$.

Thus a phase transition occurs for $\ell = \Theta(\sqrt{m})$. The main result of this paper is the description of this phase transition with the help of Brownian motion theory, following [4]. More precisely, as in [4], the asymptotic behaviour of blocks' sizes is described by widths of excursions of stochastic processes related to the Brownian motion. It turns out, by nature of the problem, and also owing to previous works of Aldous & Pitman [8], that the description given here (specially by Theorem 1.3) is more precise than in [4].

1.2 Phase transition and Brownian motion

Recall some notations and definitions from Brownian motion theory. An *excursion* (from 0) of the function f is the restriction of f to an interval $[a, b]$ such that

$$f(a) = f(b) = 0 \text{ and } |f(x)| > 0 \;\; \forall x \in]a, b[;$$

$b-a$ is the *width* or *length* of the excursion, a is the *starting point* (or the beginning) of the excursion, b the end of the excursion. Let us adopt the notation of [54, Lecture 4] for the Brownian scaling of a function f over some interval $[a, b]$:

$$f^{[a,b]} = \left(\frac{1}{\sqrt{b-a}} f(a + t(b-a)), \;\; 0 \le t \le 1\right).$$

If f is the standard linear Brownian motion, and g (resp. d) is the last zero of f before 1 (resp. the first zero of f after 1), then $e = \left|f^{[g,d]}\right|$ is called the *normalized Brownian excursion.* When it is convenient, we regard the normalized Brownian excursion $e(t)$ as defined on the whole real line, being periodic with period 1. We define, for $\lambda \ge 0$, the operator Ψ_λ on the set of bounded functions on the line by

$$\begin{aligned} \Psi_\lambda f(t) &= f(t) - \lambda t - \inf_{-\infty < s \le t} (f(s) - \lambda s) \\ &= \sup_{s \le t} \left(f(t) - f(s) - \lambda(t-s)\right). \end{aligned} \tag{1.1}$$

If f has period 1, then so has $\Psi_\lambda f$. Evidently, $\Psi_\lambda f$ is nonnegative, and we have

$$\begin{aligned} \Psi_\lambda e(x) &= e(x) - \lambda x - \inf_{0 \le y \le x} (e(y) - \lambda y), \\ \Psi_\lambda e(0) &= \Psi_\lambda e(1) = 0. \end{aligned}$$

Let $B(\lambda) = (B_k(\lambda))_{k \ge 1}$ be the sequence of widths of excursions of $\Psi_\lambda e$, sorted in decreasing order. The sequence $B(\lambda)$ is a random element of the simplex

$$\{x_1 \ge x_2 \ge ... \ge x_n \ge \;...\; \ge 0, \sum_{i \ge 1} x_i = 1\}.$$

We have:

Theorem 1.2 *If* $\lim \frac{\ell}{\sqrt{m}} = \lambda \ge 0$,

$$\frac{B^{m,\ell}}{m} \xrightarrow{law} B(\lambda).$$

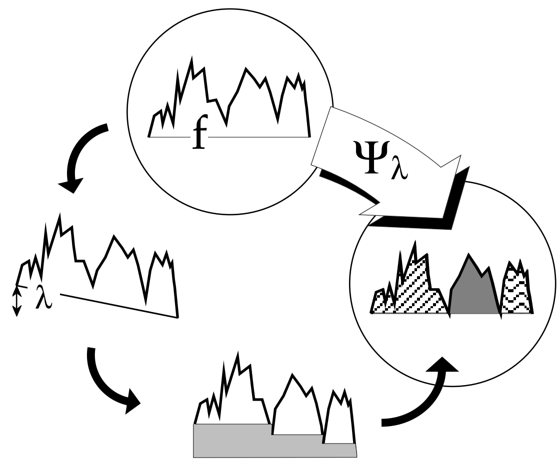

Figure 2: $\Psi_\lambda f$ and its excursions

For instance, to complete Theorem 1.1, note that

$$B_1^{m,\ell}/m \xrightarrow{law} B_1(\lambda).$$

Before we discuss the law of $B(\lambda)$, in the next Subsection, let us pursue the description of the asymptotics of the phase transition for parking blocks: up to now, we only considered the parking process frozen at a given time $n = m - \ell$, that is, just after the arrival of car c_n. The next Theorem describes the evolution of blocks' sizes, as cars arrive, during the phase transition: asymptotically, the joint law of sequences of blocks' sizes $B^{m,\lceil \lambda_i \sqrt{m} \rceil}$ after successive arrivals of cars $c_{n(i)}$, $n(i) = m - \lceil \lambda_i \sqrt{m} \rceil$, $\lambda_1 > \lambda_2 > \cdots > \lambda_k$, once these sequences are normalized, converges to the joint law of $(B(\lambda_i))_{i=1,2,\ldots,k}$. More formally, set

$$B^{(m)}(\lambda) \quad = \quad \frac{B^{m,\lceil \lambda \sqrt{m} \rceil}}{m}.$$

We have:

Theorem 1.3 *The finite-dimensional distributions of* $\left(B^{(m)}(\lambda)\right)_{\lambda \geq 0}$ *converge weakly to the finite-dimensional distributions of* $(B(\lambda))_{\lambda \geq 0}$.

Though, in the random graph model, the asymptotic distribution of sizes of clusters (connected components) has a description similar to that given at Theorem 1.2, the analog of Theorem 1.3 is false, as observed by Aldous [4]: in coalescence models based on excursions of stochastic processes, clusters (excursions) can only merge with their neighbors, while this is not true for connected components of the random graph. In Section 4, at the price of heavier notations, we give the analog of Theorem 1.3 for the asymptotic behaviour of sizes *and positions* of blocks.

1.3 Size-biased permutations

As a consequence of [34, Theorem 4], we have

Theorem 1.4 *The distribution function* $\Pr(B_1(\lambda) \leq x)$ *has the following expression:*

$$1 + \lambda^3 e^{\lambda^2/2} \sum_{k\geq 1} \frac{(-1)^k}{k!} \int_{D(\lambda,x,k)} \frac{\lambda^{2k} \exp\{-\lambda^4/2(\lambda^2 - x_1 - \ldots - x_k)\} dx_1 \ldots dx_k}{(2\pi)^{k/2}(x_1 \ldots x_k(\lambda^2 - x_1 - \ldots - x_k))^{3/2}},$$

in which

$$D(\lambda, x, k) = \left\{(x_i)_{1\leq i\leq k} \; : \; x_i \geq \lambda^2 x, 1 \leq i \leq k, \; \text{and} \; \sum x_i \leq \lambda^2\right\}.$$

Theorem 4 of [34] gives the limit law of the largest tree in a random forest: it turns out that forests and parking schemes are in one-to-one correspondence (see Subsection 5.1). Flajolet & Salvy [24] have a direct approach, to the computation of the density of $B_1(\lambda)$, by methods based on Cauchy coefficient integrals to which the saddle point method is applied: the density they obtain is a variant of the Dickman function [52, Ch. III, Sec. 5.3].

In view of Theorem 1.4, the joint law of $(B_1(\lambda), B_2(\lambda), \ldots, B_k(\lambda))$ seems out of reach, but we are more lucky with the joint law of the first terms of a sequence $R(\lambda)$ obtained by permutation of the terms of $B(\lambda)$. Roughly speaking, in the *size-biased permutation* $R(\lambda)$ of a random probability distribution such as $B(\lambda)$, the largest terms of the sequence $B(\lambda)$ appear with a high probability at the beginning of the sequence $R(\lambda)$: we have

$$\Pr\left(R_1(\lambda) = B_k(\lambda) \mid B(\lambda)\right) = B_k(\lambda), \tag{1.2}$$

the k^{th} term of $R(\lambda)$ being also drawn randomly with a probability proportional to its size, but among the terms that did not appear before. A more formal definition of size-biased permutations, by construction through a rejection method, is given in [38]: consider a sequence of independent, positive, integer-valued random variables $(I_k)_{k\geq 1}$, distributed according to $B(\lambda)$:

$$\Pr\left(I_k = j \mid B(\lambda)\right) = B_j(\lambda).$$

With probability 1 the terms of $B(\lambda)$ are positive, as $\Psi_\lambda e$ has infinitely many excursions, so each positive integer appears at least once in the sequence $(I_k)_{k\geq 1}$. Erase each repetition after the first occurence of a given integer in the sequence: there remains a random permutation $(\sigma(k))_{k\geq 1}$ of the positive integers. Set:

$$R_k(\lambda) = B_{\sigma(k)}(\lambda). \tag{1.3}$$

We have:

Theorem 1.5 *The law of the size-biased permutation* $R(\lambda)$ *of* $B(\lambda)$ *satisfies*

$$\left(R_1(\lambda) + R_2(\lambda) + \ \ldots \ + R_k(\lambda)\right)_{k\geq 1} \stackrel{law}{=} \left(\frac{N_1^2 + N_2^2 + \ \ldots \ + N_k^2}{\lambda^2 + N_1^2 + N_2^2 + \ \ldots \ + N_k^2}\right)_{k\geq 1},$$

in which the N_k *are standard Gaussian and independent.*

Actually, Theorem 1.5 gives an implicit description of the law of $B(\lambda)$, for instance it proves that almost surely each $R_k(\lambda)$ is positive, and thus a.s. $0 < B_k(\lambda) < 1$. Size-biased permutations of random discrete probabilities have been studied, among others, by Aldous [1] and Pitman [37, 38]. The most celebrated example is the size-biased permutation of the sequence of limit sizes of cycles of a random permutation. While the limit distribution of the sizes of the largest, second largest ... cycle have a complicated expression [19, 47], the successive terms $R_1, R_2, \ \dots$ of their size-biased permutation satisfies

$$(R_1 + R_2 + \ \dots \ + R_k)_{k\geq 1} \stackrel{law}{=} (1 - U_1 U_2 \ \dots \ U_k)_{k\geq 1},$$

in which the U_k are uniform on $[0, 1]$ and independent. Actually, it is common that the distribution of the size-biased permutation of a sequence has a simpler distribution than the original sequence, when the sequence is related to a Poisson point process, a famous example being the Poisson-Dirichlet distribution [10, 11, 35, 36, 41]. The distribution of $R(\lambda)$, as described in Theorem 1.5, already appeared as the law of the Δ-*valued fragmentation process* derived from the continuum random tree, introduced by Aldous & Pitman in their study of the standard additive coalescent [8, Corollary 5]: this is commented in the next Subsection.

As in the case of sizes of cycles, the unnatural size-biased permutation of $B(\lambda)$ is the limit of a natural permutation of $B^{(m)}(\lambda)$: define $R^{m,\ell} = \left(R_k^{m,\ell}\right)_{k\geq 1}$ as the sequence of sizes of blocks when the blocks are sorted by increasing date of birth (in increasing order of first arrival of a car). If $\ell \geq m$, or if there are less than k blocks, set $R_k^{m,\ell} = 0$. For instance, on Figure 1, $B^{m,\ell} = (3, 3, 1, 0, \ \dots \ )$ and $R^{m,\ell} = (3, 1, 3, 0, \ \dots \ )$. Concerning $R^{m,\ell}$, we have an analog of Theorem 1.2 :

Theorem 1.6 *If* $\lim m^{-1/2}\ell = \lambda \geq 0$,

$$\frac{R^{m,\ell}}{m} \xrightarrow{law} R(\lambda).$$

Set

$$R^{(m)}(\lambda) \quad = \quad \frac{R^{m,\lceil\lambda\sqrt{m}\rceil}}{m}.$$

For an analog of Theorem 1.3 to hold true, giving the convergence of finite dimensional distributions of $R^{(m)}$, we should define $(R(\lambda))_{\lambda\geq 0}$ as a process. This is not straightforward, as there are many possible definitions of the size-biased permutation σ_λ that throws $B(\lambda)$ on $R(\lambda)$: σ_λ has to be defined as a process too. For sake of brevity, we shall only state a result for the first component of $R(\lambda)$. Consider a random number ρ_1, uniform on $[0, 1]$ and independent of e, and define $R_1(\lambda)$ as the width of the excursion of $\Psi_\lambda e$ that contains ρ_1 (see Figure 3). Then $R_1(\lambda)$ is defined, and satisfies (1.2), simultaneously for each value of λ. We have:

Theorem 1.7 *The finite-dimensional distributions of* $\left(R_1^{(m)}(\lambda)\right)_{\lambda\geq 0}$ *converge weakly to the finite-dimensional distributions of* $(R_1(\lambda))_{\lambda\geq 0}$.

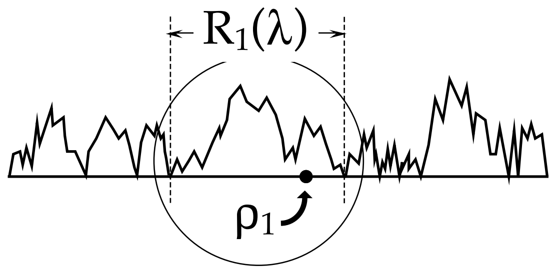

Figure 3: Random generation of $R_1(\lambda)$

It turns out that the limit process has a rather simple description: set

$$\Sigma(\lambda) = -1 + \frac{1}{R_1(\lambda)}.$$

We have

Theorem 1.8 $\Sigma = (\Sigma(\lambda))_{\lambda \geq 0}$ *is the stable subordinator with exponent* $1/2$.

It is well known that the family of first hitting times of levels λ by the Brownian motion is also the stable subordinator with exponent $1/2$. The following description of the stable subordinator with exponent $1/2$, by its finite dimensional distributions, will be useful for the proof: for any k and any k-tuple of positive numbers $(\lambda_i)_{1\leq i\leq k}$,

$$(\Sigma(\lambda_1 + \lambda_2 + ... + \lambda_i))_{1\leq i\leq k} \stackrel{law}{=} \left(\frac{\lambda_1^2}{N_1^2} + \frac{\lambda_2^2}{N_2^2} + ... + \frac{\lambda_i^2}{N_i^2} \right)_{1\leq i\leq k}, \tag{1.4}$$

in which the N_k are standard Gaussian and independent.

The well known fact that $\Sigma(\lambda)$ is a pure jump process makes sense in the parking scheme context, since the block of car c_1 is known to increase by $O(m)$ while only $O(\sqrt{m})$ cars arrived: it can only be explained by coalescence with other blocks of size $O(m)$, that is, by *instantaneous jumps*. Incidentally, let $L(\lambda)$ denote the length of the excursion of $\Psi_\lambda e$ beginning at 0, and set

$$\tilde{\Sigma}(\lambda) = -1 + \frac{1}{L(\lambda)}.$$

Bertoin [12] nicely proves that (Σ, R_1) and $(\tilde{\Sigma}, L)$ have the same law. For the moment, we do not see any combinatorial explanation of this identity between R_1 and L.

1.4 Coalescence

We give here a brief account of coalescence, which is masterfully surveyed in [5, 6]. We essentially quote the two previously cited references. Models of coalescence (aggregation, coagulation, gelation ...) have been studied in many scientific disciplines, essentially physical chemistry, but also astronomy, bubble swarms, mathematical genetics, and recently random graph theory [6, Section 1.4]. In a basic model, clusters with different masses move through space, and when two clusters (say, with masses x and y) are sufficiently close, there is some chance that they merge into a single cluster of mass $x+y$ [6, Section 1.1]. The probability that they merge is quantified, in some sense, by a *rate kernel* $K(x,y)$. As far as parking is concerned, the growth of clusters (parking blocks) is due partly to cars' arrivals, partly to aggregation with other blocks, but we saw that during the phase transition the coalescence factor is preponderant.

A complete model for coalescence, detailing mass, position, and velocity of each cluster, is too complicated for analysis, so recent works focused on the evolution of masses of clusters through time: the *general stochastic coalescent* [21] is the continuous-time Markov process whose state space is the infinite-dimensional simplex

$$\Delta = \left\{ (x_i)_{i\geq 1} \ : \ x_i \geq 0, \ \sum x_i = 1 \right\},$$

(the x_i's are the sizes of clusters) and that evolves according to the rule

$$\text{each pair } (x_i, x_j) \text{ of clusters merges at rate } K(x_i, x_j).$$

It means that, if at time t the state of the system is $(x_i)_{i\geq 1}$, the next pair (I,J) of clusters that will merge and the time $t+T$ when they merge are jointly distributed as follows: assume we are given a set of independent random variables $(T_{i,j})_{1\leq i<j}$ with distribution described by

$$\Pr\left(T_{i,j} > t\right) = \exp\left(-K(x_i, x_j)t\right),$$

and set

$$\inf_{1\leq i<j} T_{i,j} = T_{I,J} = T.$$

It turns out that the way connected components merge in the random graph process is somehow related to the *multiplicative coalescent* ($K(x,y) = xy$) [4]. One rather expects the parking to be related to the *additive coalescent* ($K(x,y) = x+y$): given that a parking scheme with m places, n cars and $\ell = m - n$ empty places has two blocks with size x and y, the probability that these two blocks merge at the next arrival is

$$\frac{x+y+2}{(\ell-1)m}, \tag{1.5}$$

as the number of empty places after block x but before block y is random uniform on $1, 2, \ldots, \ell-1$, and, given that this number is 1 (resp. $\ell-1$, $\notin \{1, \ell-1\}$) the conditional probability that the two blocks merge at the next arrival is $\frac{x+1}{m}$ $\left(\text{resp. } \frac{y+1}{m}, 0\right)$. Aldous & Pitman [8] give a construction of the additive coalescent through a fragmentation process $Y = (Y(\lambda))_{\lambda\geq 0}$: the Δ-valued random variable $Y(\lambda)$ is the ranked sequence of masses of tree components of continuum forests obtained by cutting the "edges" of the *Brownian continuum random tree* by a Poisson process of cuts with rate λ by unit length. As more or less expected, according to Theorem 1.5 and to [8, Corollary 5], the distributions of $Y(\lambda)$ and $B(\lambda)$ are the same.

Furthermore, let $\rho_1^\star$ be a leaf, of the Brownian continuum random tree, picked uniformly at random according to the mass measure, and let $Y_1^\star(\lambda)$ denote the mass of the tree component of the random forest that contains $\rho_1^\star$ when the cutting intensity is λ. Then, according to Theorem 1.8 and to [8, Theorem 6], the distributions of the stochastic processes $Y_1^\star$ and R_1 are the same. These facts suggest that

Theorem 1.9 *The processes B and Y have the same distribution.*

Theorem 1.9 is actually the main result of a recent paper by Bertoin [12]. In Section 7, we give an alternative proof of Theorem 1.9, that relies on Theorem 1.10, a path decomposition result for $\Psi_\lambda e$.

Note that Theorem 1.8 is not a mere consequence of Theorem 1.9 and [8, Theorem 6], as the very similar selection mechanisms leading to $Y_1^\star$ (resp. R_1) depend not only on the stochastic processes Y (resp. B), but on underlying richer structures, a family of Poisson point processes of cuts of a Brownian continuum random tree on one hand, and the family of stochastic processes $\Psi_\lambda e$ on the other hand. Even if one of the constructions of the Brownian continuum random tree uses the normalized Brownian excursion [3, Corollary 22], we do not know for the moment any extension of Theorem 1.9 to these richer structures, that would yield a direct proof of the identity between the distributions of the stochastic processes $Y_1^\star$ and R_1. However, in the concluding remarks, we give a rather convincing combinatorial explanation of the connection between the two richer structures.

1.5 Decomposition of sample paths of $\Psi_\lambda e$

In previous subsections, objects from Brownian motion theory allowed to describe phase transition for parking schemes. In this subsection, we translate the parking schemes combinatorial identity:

$$m^n = \sum_{k=1}^{n} C_{n-1}^{k-1} m(k+1)^{k-1}(m-k-1)^{n-k-1}(m-n-1)$$

to obtain Theorem 1.10, a property of decomposition of sample paths of $\Psi_\lambda e$ used in Section 7 to give simple proofs of Theorems 1.5, 1.8 and 1.9. Let ρ_1 be a random variable uniformly distributed on $[0,1]$ and independent of e. Almost surely, $\Psi_\lambda e(\rho_1)$ is positive. Let $g(\lambda)$ (resp. $d(\lambda)$) denote the last zero of $\Psi_\lambda e$ in the interval $[0,\rho_1)$ (resp. the first zero in the interval $(\rho_1,1]$), so that $R_1(\lambda) = d(\lambda) - g(\lambda)$. To avoid the extensive use of notation $\{x\}$ for the fractional part of the real number x, we shall extend $\Psi_\lambda e$, as well as other functions defined on $[0,1]$, such as q or r defined below, to periodic functions on the line. We set

$$\begin{aligned} q &= (\Psi_\lambda e)^{[g(\lambda),d(\lambda)]}, \\ r &= (\Psi_\lambda e)^{[d(\lambda),g(\lambda)+1]}. \end{aligned}$$

Let τ_x denote the shift operator for functions on the line, defined by

$$(\tau_x f)(y) = f(x+y).$$

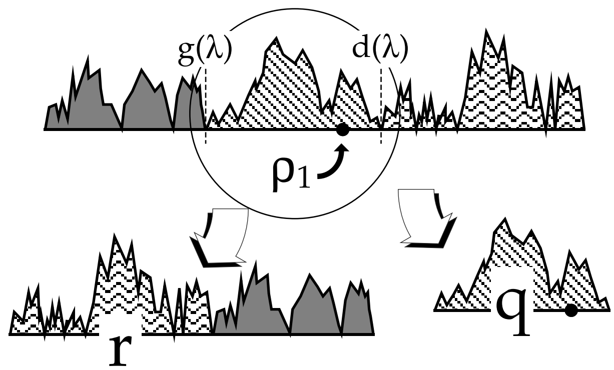

Figure 4: Decomposition of $\Psi_\lambda e$.

Theorem 1.10 *We have:*

(i) $R_1(\lambda)$ *has the same distribution as* $\frac{N^2}{\lambda^2+N^2}$*, in which* N *is standard Gaussian ;*

(ii) q *is a normalized Brownian excursion, independent of* $R_1(\lambda)$ *;*

(iii) Let w *be uniformly distributed on* $[0,1]$ *and independent of* e*. Given* (q,ρ_1) *and* $R_1(\lambda)=x$*,* $\tau_w r$ *has the same distribution as* $\tau_w \Psi_{\frac{\lambda}{\sqrt{1-x}}} e$*.*

Actually, not only the conditional distribution of $\tau_w r$, but also the conditional distribution of r has a simple description in terms of the Brownian motion, and also as a nonuniform random shift of $\Psi_{\frac{\lambda}{\sqrt{1-x}}} e$ [16]. However, the weaker form (*iii*) fills our needs for the proofs of Theorems 1.5, 1.8 and 1.9.

The paper is organized as follows. Section 2 analyses the block containing a given car or a given site, leading to the proof of Theorem 1.1. At Section 3, we give the proof of the main result, Theorem 1.2, with the help of a close coupling between *empirical processes* of mathematical statistics and the *profile* obtained by assuming that each car lays a $1/\sqrt{m}$-thick layer of sediment on the way between its first try and its final place (see Figure 5). We extend these arguments at Section 4 to obtain the asymptotic of the joint law, at different times, of widths and *positions* of blocks. Distributional results, Theorems 1.4 and 1.6, are proven at Section 5 by combinatorial arguments. We prove Theorems 1.7 and 1.10 at Section 6, with the help of Theorem 3.1, about weak convergence of profiles. Finally, in Section 7, Theorems 1.5, 1.8 and 1.9 are shown to be consequences of Theorem 1.10. Section 8 concludes the paper with an attempt of combinatorial explanation for the connections between our paper and [8].

2 On the block containing a given car, or a given site

In this Section, we prove Theorem 1.1, with the help of a weaker form of Theorem 1.6, concerning the size $R_1^{m,\ell}$ of the block containing car c_1: we have

Theorem 2.1 *If $m^{-1/2}\ell \longrightarrow \lambda > 0$,*

$$\frac{R_1^{m,\ell}}{m} \stackrel{law}{\longrightarrow} \frac{N^2}{\lambda^2 + N^2},$$

in which N is standard Gaussian.

Proof : Let $f(\lambda, x)$ denote the density of $\frac{N^2}{\lambda^2+N^2}$, and let $\varphi(m, n, k)$ denote the probability that, when parking $n = m - \ell$ cars on m places, the block containing car c_1 has k elements. We have

$$\begin{aligned} f(\lambda, x) &= \frac{\lambda}{\sqrt{2\pi}}\, x^{-1/2}(1-x)^{-3/2} \exp\left(-\frac{\lambda^2 x}{2(1-x)}\right) 1_{]0,1[}(x), \\ \varphi(m, n, k) &= C_{n-1}^{k-1} \frac{(k+1)^{k-1}}{m^n}\, m(m-k-1)^{n-k-1}(\ell-1). \end{aligned} \tag{2.6}$$

From the change of variable $x = y/(\lambda^2 + y)$, leading to

$$\int_0^1 f(\lambda, x)dx = \int_0^{+\infty} \frac{e^{-y/2}dy}{\sqrt{2\pi y}},$$

we deduce that $f(\lambda, x)$ is a density of probability and that if some random variable X has the density $f(\lambda, x)$, then $\lambda^2 X/(1-X)$ has a $\gamma_{1/2,1/2}$ law, the law of the square of a standard Gaussian random variable. That is, $N^2/(\lambda^2+N^2)$ has density $f(\lambda, x)$.

To explain (2.6), first we remark that the number of parking schemes for n cars on m places is m^n. If we specify that the last place has to be empty, we get what is called a *confined* parking scheme: there are $(m - n)m^{n-1}$ confined parking schemes [23, 30], as each orbit drawn by the group of rotations has m elements, among which $m-n$ are confined. A block with k cars can be seen as a confined parking scheme of k cars on $k + 1$ places, so there are $(k + 1)^{k-1}$ ways to build such a block. Turning to (2.6), one has to choose the set of $k - 1$ cars that belong to the same block as c_1, giving the factor C_{n-1}^{k-1}, the place where this block begins, giving the factor m, the way these k cars are allocated on these k places, giving the factor $(k + 1)^{k-1}$, and finally one has to park the $n - k$ remaining cars on the $m - k - 2$ remaining places, leaving one empty place at the beginning and at the end of the block containing car c_1. This can be done in $(m - k - 1)^{n-k-1}(\ell - 1)$ ways, the number of *confined* parking schemes of $n - k$ cars on $m - k - 1$ places. Note that these computations would hold for any given car instead of c_1.

For $0 < a < b < 1$ and $m^{-1/2}\ell \longrightarrow \lambda$,

$$\lim_m \Pr(am \le R_1^{m,\ell} \le bm) = \int_a^b f(\lambda, x)dx,$$

is a straightforward consequence of

Lemma 2.2 *For any $0 < \varepsilon < 1/2$ there exists a constant $C(\varepsilon)$ such that, whenever, simultaneously, $\varepsilon \le \frac{k}{m} \le 1 - \varepsilon$ and $\varepsilon \le \frac{\ell}{\sqrt{m}} \le \frac{1}{\varepsilon}$, we have:*

$$\left|\varphi(m, n, k) - \frac{1}{m} f\left(\frac{\ell}{\sqrt{m}}, \frac{k}{m}\right)\right| \le C(\varepsilon)m^{-3/2}.$$

Lemma 2.2 is proven at the end of this Section. $\diamondsuit$

Proof of Theorem 1.1(ii). We assume $\ell = o(\sqrt{m})$. Provided that $\ell \le \lambda\sqrt{m}$,

$$B_1^{m,\ell} \ge mR_1^{(m)}(\lambda).$$

Thus, for any $\lambda > 0$ and for m large enough:

$$\Pr(B_1^{m,\ell} < mx) \le \Pr\left(R_1^{(m)}(\lambda) < x\right).$$

Due to Theorem 2.1, we obtain that for any $\lambda > 0$

$$\limsup_m \Pr(B_1^{m,\ell} < mx) \le \Pr\left(\frac{N^2}{\lambda^2 + N^2} < x\right).$$

Clearly, for $x < 1$,

$$\inf_{\lambda > 0} \Pr\left(\frac{N^2}{\lambda^2 + N^2} < x\right) = 0.$$

Proof of Theorem 1.1(i). Let $L^{(m)}(\lambda)$ be the length, normalized by m, of the block of cars containing *place* 1 when car $c_{\lfloor m - \lambda\sqrt{m}\rfloor}$ has parked. We have, for $k > 0$,

$$\Pr\left(L^{(m)}(\lambda) = \frac{k}{m}\right) = \frac{\lfloor m - \lambda\sqrt{m}\rfloor}{m} \Pr\left(R_1^{(m)}(\lambda) = \frac{k}{m}\right),$$

and place 1 is empty with probability:

$$\Pr\left(L^{(m)}(\lambda) = 0\right) = \frac{\lceil \lambda\sqrt{m}\rceil}{m}.$$

We also have

$$\Pr\left(L^{(m)}(\lambda) = \frac{k}{m} \;\middle|\; B_1^{(m)}(\lambda) = \frac{k}{m}\right) \ge \frac{k}{m},$$

and thus

$$\begin{aligned} E\left[L^{(m)}(\lambda)\right] &\ge \sum_k \frac{k}{m} \Pr\left(L^{(m)}(\lambda) = B_1^{(m)}(\lambda) = \frac{k}{m}\right) \\ &\ge E\left[\left(B_1^{(m)}(\lambda)\right)^2\right]. \end{aligned}$$

Owing to $\sqrt{m} = o(\ell)$, we obtain that for any $\lambda > 0$,

$$B_1^{m,\ell}(\omega) \le mB_1^{(m)}(\lambda, \omega),$$

when m is large enough, not depending on ω, so that:

$$\begin{aligned} \limsup_m E\left[\left(\frac{B_1^{m,\ell}}{m}\right)^2\right] &\le \inf_{\lambda > 0} \lim_m E\left[L^{(m)}(\lambda)\right] \\ &= \inf_{\lambda > 0} E\left[\frac{N^2}{\lambda^2 + N^2}\right], \end{aligned}$$

yielding (i). $\diamondsuit$

Proof of Lemma 2.2. Setting, for brevity, $x = k/m$ and $\lambda = m^{-1/2}\ell$, we can write

$$\varphi(m, m - \lambda\sqrt{m}, k) = \Phi_{m,1}\Phi_{m,2}\Phi_{m,3},$$

in which:

$$\begin{aligned}
\Phi_{m,1} &= C^{xm-1}_{m-\lambda\sqrt{m}-1} \\
\Phi_{m,2} &= (xm+1)^{xm-1} m^{-m+\lambda\sqrt{m}-1} \\
\Phi_{m,3} &= (m - xm - 1)^{m-\lambda\sqrt{m}-xm-1}(\lambda\sqrt{m} - 1).
\end{aligned}$$

We obtain

$$\begin{aligned}
\Phi_{m,1} &= \frac{x^{-xm+1/2} e^{\kappa(m)-\beta(m)} \left(1 + O(1/k) + O\left(\sqrt{m}/(m-k)\right)\right)}{\sqrt{2\pi m} \quad (1-x)^{(1-x)m-\lambda\sqrt{m}+1/2}} \\
\kappa(m) &= 1 + (m - \lambda\sqrt{m} - 1/2)\log(1 - \lambda/\sqrt{m} - 1/m) \\
&= -\lambda\sqrt{m} + \lambda^2/2 + O(m^{-1/2}) \\
\beta(m) &= (m - \lambda\sqrt{m} - xm + 1/2)\log\left(1 - \frac{\lambda}{(1-x)\sqrt{m}}\right) \\
&= -\lambda\sqrt{m} + \frac{\lambda^2}{2(1-x)} + O\left((m-k)^{-1/2}\right). \\
\Phi_{m,2} &= x^{xm-1} m^{xm-m+\lambda\sqrt{m}}(1 + O(1/k))e \\
\Phi_{m,3} &= \frac{\lambda m^{(1-x)m-\lambda\sqrt{m}}(1-x)^{(1-x)m-\lambda\sqrt{m}}}{e(1-x)\sqrt{m}} \left(1 + O\left(\sqrt{m}\right) + O\left(\frac{\sqrt{m}}{m-k}\right)\right)
\end{aligned}$$

and finally:

$$\begin{aligned}
\varphi(m, m - \lambda\sqrt{m}, k) &= \frac{\lambda(1-x)^{-3/2}}{m\sqrt{2\pi x}} \exp\left(-\frac{\lambda^2 x}{2(1-x)}\right)(1 + \eta(m,k)) \\
|\eta(m,k)| &\leq \frac{K_1}{k} + \frac{K_2\sqrt{m}}{m-k} + \frac{K_3}{\sqrt{m}} + \frac{K_4}{\sqrt{m-k}}. \qquad \diamondsuit
\end{aligned}$$

3 Profiles of parking schemes

Let H_k denote the number of cars that tried to park on place k, *successfully or not*, and let h_m denote the *profile* of the parking scheme, defined by:

$$h_m(t) = \frac{H_{\lfloor mt \rfloor}}{\sqrt{m}}.$$

As $H_k = 0$ if and only if place k is empty, the width of an excursion of h_m turns out to be the length of some block of cars, normalized by $1/m$. Set:

$$h_\lambda(t) = \Psi_\lambda e(\{-v + t\}),$$

in which v denotes a uniform random variable independent of e.

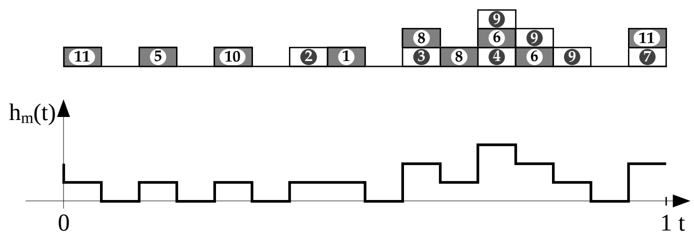

Figure 5: A parking scheme and its profile $(m = 16, n = 11)$

In this Section we give the proof of Theorem 1.2, that has roughly speaking three steps: as a first result, we establish in Subsection 3.1 a close coupling between H_k and the empirical processes of mathematical statistics. In Subsection 3.2, using Theorems of Donsker and Vervaat, we prove the following Theorem, which is the key to this paper.

Theorem 3.1 *If* $\lim_m \frac{\ell}{\sqrt{m}} = \lambda$,

$$h_m \stackrel{weakly}{\longrightarrow} h_\lambda.$$

Theorem 1.2 states the convergence of widths of excursions of h_m to widths of excursions of h_λ. Its proof, given in Subsection 3.3, requires some care, as the sequence of widths of excursions is not a continuous functional of h_m: the proof relies on an extension of the invariance principle that we learned from [4, 7]. Further consequences of Theorem 3.1 are Theorem 1.10 and also some results about stochastic processes developped in [16]. Theorem 1.3 is the consequence of Theorem 4.1, an extension of Theorem 3.1. The case $\lambda = 0$, $\ell = 1$ of Theorem 3.1 was developped in [15, Section 4] for the study of the width of labeled trees.

3.1 Connection between parking and empirical processes

Propositions 3.3 and 3.4, at the end of this subsection, are the key points for the convergence of blocks' sizes. Given a sequence $(U_k)_{k\geq 1}$ of independent uniform random variables, we assume that first try of car c_k is place $\lceil mU_k \rceil$, all parking schemes being thus equiprobable. If Y_k denotes the number of cars whose first try was place k, then we have:

$$H_{k+1} = Y_{k+1} + (H_k - 1)_+, \tag{3.7}$$

since either place k is occupied by car c_i and, among the H_k cars that tried place k, only car c_i won't visit place $k+1$, so that $H_{k+1} = Y_{k+1} + (H_k - 1)$, or place k is empty and $H_{k+1} = Y_{k+1}$. We understand this equation, when $k = m$, as $H_1 = Y_1 + (H_m - 1)_+$. This induction alone does not give the H_k's, since we do not have any starting value. The gap is filled by Proposition 3.2, that gives the connection between hashing (or parking) and empirical processes.

Given a sample $(U_1, U_2, ..., U_n)$ of uniform random variables, the *empirical distribution* F_n and the *empirical process* α_n (see [18, 43, 48] for background) are respectively defined by

$$\begin{aligned} F_n(t) &= \frac{\#\{1 \le i \le n \mid U_i \le t\}}{n} \\ &= t + \frac{\alpha_n(t)}{\sqrt{n}}. \end{aligned}$$

The process α_n gives a measure of the accuracy of the approximation of the true distribution function t by the empirical distribution function $F_n(t)$, and was, as such, extensively studied in mathematical statistics. Let V be defined by

$$\begin{aligned} a &= \min\{\alpha_n(k/m) \mid 1 \le k \le m\} \\ V &= \min\{j \mid 1 \le j \le m \text{ and } \alpha_n(j/m) = a\}. \end{aligned}$$

Proposition 3.2 *Place V is empty.*

Proof of Proposition 3.2. Set:

$$\begin{aligned} A_k &= \sqrt{n}\,\alpha_n(k/m) \\ &= \#\{1 \le i \le n \mid U_i \le k/m\} - k\,\frac{n}{m}. \end{aligned}$$

Since we have:

$$Y_k = \#\{j \mid 1 \le j \le n \text{ and } \lceil mU_j \rceil = k\},$$

it follows that:

$$A_{k+1} = A_k + Y_{k+1} - n/m. \tag{3.8}$$

As $A_0 = A_m$, we can extend A_k, and Y_k as well, to periodic sequences, so that (3.8) holds true for any integer. Thus

$$\begin{aligned} Y_{V-k+1} + Y_{V-k+2} + \cdots + Y_V &= kn/m - A_{V-k} + A_V \\ &\le \lfloor kn/m \rfloor \le k-1, \end{aligned} \tag{3.9}$$

the first inequality by definition of V. We remark that if the set of places $\{j-k+1, j-k+2, \cdots, j\}$ is full while no more than $k-1$ cars had their first try in it, then necessarily place $j-k$ is occupied. Thus if place V is not empty, by letting $k=1$ in (3.9) we obtain that $V-1$ is not empty either, and by induction on k, still using (3.9), no place is empty. $\diamondsuit$

For $0 \le k \le m$, set:

$$\begin{aligned} C_k &= \sqrt{n}\,\alpha_n(k/m) - k\,\frac{\ell}{m} \\ &= Y_1 + Y_2 + \cdots + Y_k - k, \end{aligned}$$

and extend it to any integer, through $C_{k+m} = C_k - \ell$. With the convention that H_k is periodic as well, (3.7) holds true for any integer, and we can use it to compute H_k, starting from $H_V = 0$:

Proposition 3.3 *For any $k \in \{1, 2, \ \dots\ , m-1\}$,*

$$H_{V+k} = C_{V+k} - C_V + \max_{1 \le i \le k}\{(C_{V-1} - C_{V+i-1})_+\}.$$

Since blocks of cars are just blocks of consecutive indices k such that $H_k > 0$ (*excursions* of H_k), our study relies essentially on this expression, that connects blocks of cars with empirical processes. A similar line of proof is used in [4, Subsection 1.3] for the study of connected components of random graphs.

Proof of Proposition 3.3. Set $\gamma_0 = \tilde{\gamma}_0 = 0$, and, for $k \geq 1$:

$$\begin{aligned} \tilde{\gamma}_k &= \#\{\text{empty places in the set } \{V, V+1, \ \dots \ , V+k-1\}\} \\ &= \#\{j \mid V \leq j \leq V+k-1 \text{ and } H_j = 0\} \\ &= 1 + \#\{j \mid V+1 \leq j \leq V+k-1 \text{ and } H_j = 0\} \\ \gamma_k &= \max_{1 \leq i \leq k} \{C_{V-1} - C_{V+i-1}\}. \end{aligned}$$

Relation (3.7) yields at once:

$$\begin{aligned} H_{V+k} &= H_V + Y_{V+1} + Y_{V+2} + \ \dots \ + Y_{V+k} - k + \tilde{\gamma}_k \\ &= C_{V+k} - C_V + \tilde{\gamma}_k, \end{aligned}$$

so the proof of Proposition 3.3 reduces to that of $\tilde{\gamma}_k = \gamma_k$.

We already have $\gamma_1 = \tilde{\gamma}_1 = 1$. For $k \geq 1$, note that either $\gamma_{k+1} = \gamma_k + 1$ or $\gamma_{k+1} = \gamma_k$. First consider the case $\gamma_{k+1} = \gamma_k$: there exists j such that $0 \leq j \leq k-1$ and $C_{V+j} \leq C_{V+k}$. This can be rewritten:

$$Y_{V+j+1} + \ \dots \ + Y_{V+k} \geq k - j,$$

meaning that more than $k-j-1$ cars want to park on only $k-j$ places. Thus the last place, $V+k$, is necessarily occupied, i.e. $\tilde{\gamma}_{k+1} = \tilde{\gamma}_k$.

Assume now that $\gamma_{k+1} = \gamma_k + 1$: for any j such that $0 \leq j \leq k-1$, we have $C_{V+j} \geq C_{V+k} + 1$, or equivalently:

$$Y_{V+j+1} + \ \dots \ + Y_{V+k} \leq k - j - 1.$$

Using this inequality in the same way as we used relation (3.9) previously, we conclude that if $\tilde{\gamma}_{k+1} \neq \tilde{\gamma}_k + 1$ or, equivalently, if $V+k$ is not empty, then the sets of places $\{V+j, \cdots, V+k\}$ have to be full, for any j such that $0 \leq j \leq k-1$, including thus V. ◇

We just proved that

Proposition 3.4 *Place $V+k$ is empty if and only if $\gamma_{k+1} = \gamma_k + 1$, or if and only if $-C_j$ has a record at $j = V+k$.*

Sequence γ_k will be easier to handle than $\tilde{\gamma}_k$, when dealing with uniform convergence in the next subsection.

3.2 Proof of Theorem 3.1

Recall that Donsker (1952), following an idea of Doob, proved that:

Theorem 3.5 *Let $b = (b(t))_{0 \leq t \leq 1}$ be a Brownian bridge. We have:*

$$\alpha_n \overset{weakly}{\longrightarrow} b.$$

We shall also need:

Theorem 3.6 *(Vervaat, 1979 [53]) Let v be the almost surely unique point such that $b(v) = \min_{0\le t\le 1} b(t)$. Then v is uniform and $e = (e(t))_{0\le t\le 1}$, defined by $e(t) = b(\{v+t\}) - b(v)$, is a normalized Brownian excursion, independent of v.*

Owing to the Skorohod representation theorem [46, II.86.1], we assume the joint existence, on some probabilistic triplet (Ω, A, P), of a sequence of copies of empirical processes, also denoted α_m, and of a Brownian bridge b, such that, for almost any $\omega \in \Omega$, $t \to \alpha_m(t,\omega)$ converges uniformly on $[0,1]$ to $t \to b(t,\omega)$. We also assume, in the definition of h_λ, that e and v are generated from b, using Vervaat's Theorem, so that $h_\lambda = \Psi_\lambda b$.

The idea of the proof is to build a sequence of copies of h_m that converges almost surely uniformly to a copy of h_λ: first α_n defines sequences

$$\begin{aligned} A_k^{m,\ell} &= \sqrt{n}\,\alpha_n(k/m), \\ C_k^{m,\ell} &= \sqrt{n}\,\alpha_n(k/m) - \ell\,\frac{k}{m}. \end{aligned}$$

Then, from $C_k^{m,\ell}$, we can define, through Proposition 3.3, $H_k^{m,\ell}$ that is distributed as H_k, though no underlying parking scheme has been defined. Actually we can also define

$$\left((A_k^{m,\ell}, C_k^{m,\ell}, H_k^{m,\ell}, \gamma_k^{m,\ell}, Y_k^{m,\ell})_{-\infty\le k\le+\infty}, V(m,\ell)\right),$$

with the same distribution as $((A_k, C_k, H_k, \gamma_k, Y_k)_{-\infty\le k\le+\infty}, V)$ in the previous Subsection. Thus $\tilde{h}_m$ defined by

$$\tilde{h}_m(t) = \frac{1}{\sqrt{m}} H_{\lfloor mt\rfloor}^{m,\ell}$$

is distributed as h_m, and we shall drop the tilda in what follows. We also set:

$$z_m(t) = \frac{1}{\sqrt{m}}\left(C_{V(m,\ell)+\lfloor mt\rfloor}^{m,\ell} - C_{V(m,\ell)}^{m,\ell}\right).$$

We have

Lemma 3.7 *If $\lim_m \frac{\ell}{\sqrt{m}} = \lambda$, then for almost any ω,*

$$\alpha_n(\lfloor mt\rfloor/m) \overset{uniformly}{\longrightarrow} b(t); \tag{3.10}$$

$$\lim_m \frac{V(m,\ell)}{m} = v; \tag{3.11}$$

$$z_m(t) \overset{uniformly}{\longrightarrow} z(t) = e(t) - \lambda t; \tag{3.12}$$

$$h_m(\{t + (V(m,\ell)/m)\}) \overset{uniformly}{\longrightarrow} \Psi_\lambda e(t). \tag{3.13}$$

Theorem 3.1 is a reformulation of (3.13), since we have:

$$\Psi_\lambda e(\{t-(V(m,\ell)/m)\}) \overset{uniformly}{\longrightarrow} \Psi_\lambda e(\{t-v\}) = h_\lambda(t).$$

Proof of (3.10). Set

$$M_m = \max_{0<k\le m} Y_k^{m,\ell}.$$

We have:

$$\begin{aligned}|\alpha_n(\lfloor mt\rfloor/m)-\alpha_n(t)| &\le \frac{\sqrt{n}}{m}+\frac{M_m}{\sqrt{n}}\\ &\le \frac{1+M_m}{\sqrt{n}},\end{aligned}$$

and, as $Y_k^{m,\ell}$ follows the binomial distribution with parameters $(n,1/m)$,

$$\begin{aligned}\Pr(M_m \ge C\log m) &\le m\Pr(Y_1^{m,\ell}\ge C\log m)\\ &\le mE[\exp(KY_1^{m,\ell})]\exp(-KC\log m)\\ &\le Am^{1-KC}. \end{aligned} \tag{3.14}$$

Thus Borel-Cantelli Lemma entails that, for a suitable C, with probability 1 the supremum norm of $\alpha_n(\lfloor mt\rfloor/m)-\alpha_n(t)$ vanishes as quickly as $\frac{C\log m}{\sqrt{n}}$. ♢

Proof of (3.11). For this proof and the next one, we consider an ω such that simultaneously $\alpha_n(t,\omega)$ and $\alpha_n(\lfloor mt\rfloor/m,\omega)$ converges uniformly (for $t\in[0,1]$) to $b(t,\omega)$, and such that $t\to b(t,\omega)$ reaches its minimum only once (we know that the set of such ω's has measure 1). We set:

$$\begin{aligned}\varepsilon_{m,1} &= \sup_{0\le t\le 1}|\alpha_n(\lfloor mt\rfloor/m)-\alpha_n(t)|,\\ \varepsilon_{m,2} &= \sup_{0\le t\le 1}|b(t)-\alpha_n(t)|,\\ \varepsilon_{m,3} &= \sup_{0\le t\le 1}\left|b(v+t)-b\left(\frac{V(m,\ell)}{m}+t\right)\right|.\end{aligned}$$

From the continuity property of b, the first minimum, $V(m,\ell)/m$, of $\alpha_n(\lfloor mt\rfloor/m)$ converges to the only minimum of b (i.e. v): we have

$$\begin{aligned}b(V(m,\ell)/m) &\le \alpha_n(V(m,\ell)/m)+\varepsilon_{m,2}\\ &\le \alpha_n(\lfloor mv\rfloor/m)+\varepsilon_{m,2}\\ &\le b(v)+\varepsilon_{m,1}+2\varepsilon_{m,2}.\end{aligned}$$

Now the minimum of $b(t)$ over the set $[v-\varepsilon,v+\varepsilon]^c\cap[0,1]$ is $b(v)+\eta$ for some positive η, and thus, if $\varepsilon_{m,1}+2\varepsilon_{m,2}<\eta$, then necessarily $|v-V(m,\ell)/m|<\varepsilon$. ♢

Proof of (3.12). One checks easily that:

$$|z_m(t)-z(t)| \le 2(\varepsilon_{m,1}+\varepsilon_{m,2}+\varepsilon_{m,3})+\frac{\lambda}{m}+\left|\lambda-\frac{\ell}{\sqrt{n}}\right|. \qquad ♢$$

Proof of (3.13). According to Proposition 3.3, we have:

$$h_m\left(\{t+(V(m,\ell)/m)\}\right) = z_m(t)+\max_{0\le s\le t}(-z_m(s))+\varepsilon_{m,4}(t),$$

where

$$\begin{aligned}\varepsilon_{m,4}(t) &= \max_{0\le s\le t}\left(\frac{Y^{m,\ell}_{V(m,\ell)+\lfloor ms\rfloor}}{\sqrt{n}}-z_m(s)\right)-\max_{0\le s\le t}(-z_m(s)),\\ |\varepsilon_{m,4}(t)| &\le \frac{M_m}{\sqrt{n}}.\end{aligned}$$

Thus (3.13) follows from the uniform convergence of z_m to z. $\diamondsuit$

3.3 An extension of the invariance principle

This section is the last step of the proof of Theorem 1.2. The widths of excursions of $h_m(t)$ above zero are the sizes of the blocks of cars of the corresponding parking scheme, normalized by m. Unfortunately, uniform convergence of h_m to h does not entails convergence of sizes of excursions. However the excursions of z_m above its current minimum are exactly the excursions of h_m above 0, up to the random shift $V(m,\ell)/m$, and, according to [4, Section 2.3], the uniform convergence of z_m to z entails convergence of sizes of excursions of z_m above its current minimum to sizes of excursions of z above its current minimum, provided that z does never reach its current minimum *two times*. It is known that this last condition holds true for almost each sample path z, so that we have almost sure convergence of sizes of excursions of z_m, or equivalently of sizes of blocks. Similarly, excursions of z above its current minimum are also excursions of $\Psi_\lambda e$ above 0, yielding Theorem 1.2.

Let us give some details and notations. We shall apply to z_m and z the following weakened form of [4, Lemma 7, p. 824]:

Lemma 3.8 *Suppose ζ: $[0,+\infty[\longrightarrow R$ is continuous. Let E be the set of nonempty intervals $I=(l,r)$ such that:*

$$\zeta(r)=\zeta(l)=\min_{s\le l}\zeta(s), \qquad \zeta(s)>\zeta(l) \text{ for } l<s<r.$$

Suppose that, for intervals $I_1,\ I_2\in E$ with $l_1<l_2$ we have

$$\zeta(l_1)>\zeta(l_2).$$

Suppose also that the complement of $\cup_{I\in E}(l,r)$ has Lebesgue measure 0. Let $\Theta=\{(l,r-l):\ (l,r)\in E\}$. Now let $\zeta_m\longrightarrow\zeta$ uniformly on $[0,1]$. Suppose $(t_{m,i},\ i\ge 1)$ satisfy the following:

(i) $0=t_{m,1}<t_{m,2}<\ldots<t_{m,k+1}=1$;

(ii) $\zeta_m(t_{m,i})=\min_{u\le t_{m,i}}\zeta_m(u)$;

(iii) $\lim_m \max_i(\zeta_m(t_{m,i})-\zeta_m(t_{m,i+1}))=0$

Write $\Theta^{(m)} = \{(t_{m,i}, t_{m,i+1} - t_{m,i});\ 1 \le i \le k\}$*. Then* $\Theta^{(m)} \longrightarrow \Theta$ *for the vague topology of measures on* $[0,1] \times (0,1]$*.*

Set of points, such as $\Theta^{(m)}$ or Θ, can also be seen as *point processes* (i.e. measures that are infinite sums of Dirac masses): we identify the set A and the measure

$$\sum_{x \in A} \delta_x.$$

For point processes on $[0,1] \times (0,1]$, the following criterium of convergence holds:

Proposition 3.9 $\Theta^{(m)} \longrightarrow \Theta$ *for the vague topology if and only if, for any* $y > 0$ *such that* $\Theta([0,1] \times \{y\}) = 0$*,*

(i) for m *large enough,* $\Theta^{(m)}([0,1] \times [y,1]) = \Theta([0,1] \times [y,1])$*;*

(ii) for any $x \in [0,1] \times [y,1]$ *such that* $\Theta(\{x\}) > 0$ *there is a sequence of points* x_m*,* $\Theta^{(m)}(x_m) > 0$*, such that* $x_m \longrightarrow x$*.*

As an easy consequence, partly due to the fact that second components add up to 1:

Corollary 3.10 *If* $\Theta^{(m)} \longrightarrow \Theta$ *for the vague topology, then the sequence of second components of points of* $\Theta^{(m)}$*, sorted in decreasing order, converge componentwise and in* ℓ_1 *to the corresponding sequence for* Θ*.*

One can find the proofs of Lemmata and Propositions of this subsection, and also of the stochastic calculus points in the next proof, in [17, pp. 30-34].

Let us choose $(\zeta_m, \zeta) = (z_m, z)$, defined at Subsection 3.2. Let the $t_{m,i}$'s of Lemma 3.8 be the successive positive records of $-z_m(t)$ so, due to Lemma 3.4, the $V(m,\ell) + mt_{m,i}$'s are the ℓ empty places of the corresponding parking scheme, counted starting at $V(m,\ell)$. The sequence of second components of $\Theta^{(m)}$ (resp. of Θ) is nothing else but $\frac{1}{m}B^{m,\ell}$ (resp. $B(\lambda)$). Thus Theorem 1.2 follows from Lemma 3.8 and Corollary 3.10, applied to $\zeta(t) = z(t)$, $\zeta_m(t) = z_m(t)$. Let us check the assumptions of Lemma 3.8. First, not depending on i,

$$z_m(t_{m,i}) - z_m(t_{m,i+1}) = 1/\sqrt{m},$$

giving assumption (iii). The standard Brownian motion satisfies the assumption "almost surely, $\zeta(l_1) < \zeta(l_2)$ for any $l_1 < l_2$", and, due the Cameron-Martin-Girsanov formula, this extends to solutions of stochastic differential equations with smooth coefficients, including z (cf. [44, Chp. XI, Ex. 3.11]). Setting $O = \cup_{I \in E}(l, r)$, the Lebesgue measure of O^c is 0 for similar reasons (see [17, pp. 33-34] for details). $\diamondsuit$

4 Extension to finite-dimensional distributions

This Section is devoted to the proof of Theorem 1.3. Up to now, with the exception of Subsection 1.4, we only considered the parking process frozen at a given time n, that is, just after the arrival of car c_n. Theorem 1.3 is a result about the dependence between parking schemes, at successive times $n_1 < n_2 < \cdots < n_k$. Thus we shall need a two-parameters (time and place) analog of Theorem 3.1. For each m, ℓ, let $h_{m,\ell}(t)$ be the profile of the parking scheme of the $m - \ell$ first cars on the m places. Similarly, let $z_{m,\ell}(t)$ be the analog of z_m defined at Section 3. Finally, for $\lambda\sqrt{m} \leq m$, set

$$\psi_m(\lambda, t) = h_{m,\lceil \lambda\sqrt{m}\rceil}(t),$$

else let $\psi_m(\lambda, t) = 0$. The dependence between the m successive parking schemes, after the m successive arrivals on m places is captured by the two-parameters process

$$\psi_m = (\psi_m(\lambda, t))_{0\leq\lambda,\, 0\leq t\leq 1}\,.$$

Note that the time parameter, λ, decreases as time goes by and cars arrive, while t

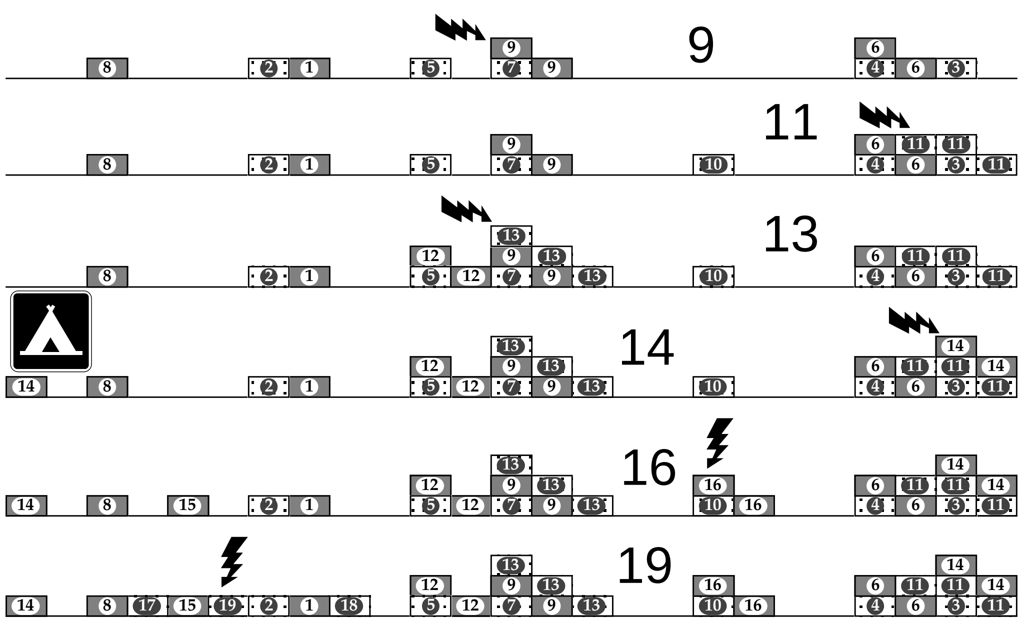

Figure 6: Parking schemes for $m = 25$ places and $n = 1, \cdots, 19$ cars.

is the location parameter: $\sqrt{m}\ \psi_m\left(\frac{\ell}{\sqrt{m}}, \frac{k}{m}\right)$ is the number of cars that tried to park, successfully or not, on place k, among the $m - \ell$ cars already arrived. We have:

Theorem 4.1 *There exists, on some probability space Ω, a uniform random variable v, and copies of ψ_m and of the normalized Brownian excursion e, such that, for $D_\Lambda = [0, \Lambda] \times [0, 1]$,*

$$\Pr\left(\forall\Lambda,\ \ \psi_m(\lambda, t) \ \xrightarrow[on\ D_\Lambda]{uniformly}\ \ h_\lambda(t)\right) = 1.$$

Set

$$Z_m(\lambda, t) = z_{m,\lceil\lambda\sqrt{m}\rceil}(t).$$

We shall actually prove that

$$\Pr\left(\forall\Lambda,\ \ Z_m(\lambda,t) \ \ \underset{on\ D_\Lambda}{\overset{uniformly}{\longrightarrow}} \ \ e(t)-\lambda t\right) = 1. \qquad (4.15)$$

Theorem 4.1 will follow, as well as a description of the asymptotic evolution, as cars arrive, of the whole sequence of sizes *and positions* of blocks.

We need more notations to give a precise statement. Let $\Theta^{(m)}(\lambda)$ denote the point process corresponding to the choice $\zeta_m = Z_m(\lambda, \cdot)$ in Lemma 3.8: the first components of points of $\Theta^{(m)}(\lambda)$ are the positions, *relative to* $V(m,\ell)$ and normalized by m, of the $\lceil\lambda\sqrt{m}\rceil$ empty places after the $\lfloor m-\lambda\sqrt{m}\rfloor^{th}$ arrival; the second components are the lengths, normalized by m, of blocks starting at these empty places (the length including also the initial empty place). We allow empty blocks, that is, empty places followed by another empty place: the corresponding length is $1/m$. Similarly, let $\Theta(\lambda)$ denote the point process corresponding to the choice $\zeta(t) = e(t) - \lambda t$ in Lemma 3.8: the first component of an element of $\Theta(\lambda)$ is the starting point of an excursion of $\Psi_\lambda e$, the second component of this element being the width of the same excursion. We have

Theorem 4.2 *The finite-dimensional distributions of* $\Theta^{(m)}(\lambda)$ *converges weakly to the finite-dimensional distributions of* $\Theta(\lambda)$.

This result is weaker than the weak convergence of $\Theta^{(m)}$ to Θ, that is, it does not insure the weak convergence of any continuous functional of $\Theta^{(m)}$ to the same functional applied to Θ, but it insures that if $\Theta^{(m)}$ has a weak limit, this limit can only be Θ.

Proof of Theorem 4.1. As in Subsection 3.2, we start, on some space Ω, with a sequence α_m of empirical processes that converges almost surely uniformly to a Brownian bridge $b(t) = e(\{t-v\}) - e(-v)$. For the proof of Theorem 3.1, there was no need to build a random parking scheme corresponding to α_m - but, maybe, for the mental picture. This task cannot be avoided now, as we need the *chronology* to deduce $h_{m,\ell}$, $z_{m,\ell}$, ψ_m, Z_m and Θ_m from α_m.

There is however a slight difficulty: α_m provides the total number $Y_k^{m,\ell}$ of cars whose first try was on place k, but it does not provide the chronology. Let us collect some basic facts concerning empirical processes: α_m has m positive jumps with height $\frac{1}{\sqrt{m}}$, at places that we call $\left(J_k^{(m)}\right)_{1\le k\le m}$. Between the jumps α_m has the slope - negative - $-\sqrt{m}$. The random vector $J^{(m)} = \left(J_k^{(m)}\right)_{1\le k\le m}$ is uniformly distributed on the simplex $\{0 < x_1 < x_2 < ... < x_m < 1\}$. Any random permutation σ_m of $J^{(m)}$'s components, with σ_m and $J^{(m)}$ independent, yields a sequence

$$\left(U_k^{(m)}\right)_{1\le k\le m} = \left(J_{\sigma_m(k)}^{(m)}\right)_{1\le k\le m}$$

of independent uniform random variables on $[0,1]$, whose empirical process is α_n.

Thus we can recover the chronology with the help of σ_m, assuming that car c_k tries to park first on place $\left\lceil nU_k^{(m)} \right\rceil$. Let us define $\alpha_{m,\ell}(t)$ (resp. $\tilde{\alpha}_{m,\ell}(t)$) as the empirical processes for the samples $(U_i^{(m)})_{1\le i\le m-\ell}$ (resp. $(U_i^{(m)})_{m-\ell+1\le i\le m}$). Both are samples of independent and uniform random variables. Now $\alpha_{m,\ell}(t)$ allows to define the profiles $h_{m,\ell}(t)$ of the successive parking schemes, and to define $z_{m,\ell}(t)$, ψ_m and Z_m as well, following the same lines as in Subsection 3.2.

We shall see now that any choice of the sequence $(\sigma_m)_{m\ge1}$ of uniform random permutations insures the convergence of Θ_m to Θ, provided that α_m and σ_m are independent for each m. We give at the end of the proof a construction of σ_m that will be useful in Section 6. We have:

$$\sqrt{m}\ \ \alpha_m(t) = \sqrt{m-\ell}\ \ \alpha_{m,\ell}(t) + \sqrt{\ell}\ \ \tilde{\alpha}_{m,\ell}(t),$$

with the consequence that:

$$|\alpha_m(t) - \alpha_{m,\ell}(t)| \le \left| -1 + \sqrt{1 - \frac{\ell}{m}} \right| |\alpha_{m,\ell}(t)| + \sqrt{\frac{\ell}{m}}\ |\tilde{\alpha}_{m,\ell}(t)|.$$

According to the DKW inequality [33], not depending on (m,ℓ),

$$\Pr(\sup_t |\tilde{\alpha}_{m,\ell}(t)| \ge x) \le 2\exp(-2x^2),$$

thus, for suitable K_1 and K_2, and for $\varepsilon > 0$,

$$\Pr\left(\sup_{0\le\ell\le\Lambda\sqrt{m}} \sup_t |\alpha_m(t) - \alpha_{m,\ell}(t)| \ge m^{-1/4+\varepsilon} \right) \le K_1\sqrt{m}\ e^{-K_2 m^{2\varepsilon}}.$$

Thus, using Borel-Cantelli lemma, we obtain

$$\Pr\left(\sup_{0\le\ell\le\Lambda\sqrt{m}} \sup_t |\alpha_m(t) - \alpha_{m,\ell}(t)| = O(m^{-1/4+\varepsilon}) \right) = 1. \tag{4.16}$$

Owing to (4.16), a simple glance at the proof of (3.11) show that the convergence of $V(m,\ell)/m$ to v, for $0 \le \ell \le \Lambda\sqrt{m}$, is uniform, almost surely. Slightly changing the definitions of $\varepsilon_{m,1}$, $\varepsilon_{m,2}$, $\varepsilon_{m,3}$ of Subsection 3.2 and defining also $\varepsilon_{m,5}$, $\varepsilon_{m,6}$, $v_m(\lambda)$ as follows:

$$\begin{aligned}
\varepsilon_{m,1} &= \sup_{0\le t\le1} |\alpha_m(\lfloor mt \rfloor/m) - \alpha_m(t)|, \\
\varepsilon_{m,2} &= \sup_{0\le t\le1} |b(t) - \alpha_m(t)|, \\
\varepsilon_{m,3} &= \sup_{0\le\ell\le\Lambda\sqrt{m}} \sup_{0\le t\le1} |b(t+v) - b(t + V(m,\ell)/m)|, \\
\varepsilon_{m,5} &= \sup_{0\le\ell\le\Lambda\sqrt{m}} \sup_{0\le t\le1} |\alpha_m(t) - \alpha_{m,\ell}(t)|, \\
\varepsilon_{m,6} &= \sup_{0\le\lambda\le\Lambda} \sup_{0\le t\le1} \left| \frac{\lceil \lambda\sqrt{m} \rceil}{\sqrt{m - \lceil \lambda\sqrt{m} \rceil}} \frac{\lfloor mt \rfloor}{m} - \lambda t \right| \\
v_m(\lambda) &= \frac{V(m, \lceil \lambda\sqrt{m} \rceil)}{n},
\end{aligned}$$

we have, for $i \in \{1,2,3,5,6\}$,

$$\lim_m \varepsilon_{m,i} = 0,$$

from Subsection 3.2 for $i = 1, 2$ and from (4.16) and uniform continuity of b for $i = 3$. Furthermore, we have

$$\begin{aligned} |Z_m(\lambda,t) - e(t) + \lambda t| &\leq \varepsilon_{m,1} + \varepsilon_{m,2} + \varepsilon_{m,5} + \varepsilon_{m,6} + |b(t + v_m(\lambda)) - b(t+v)| \\ &\leq \varepsilon_{m,1} + \varepsilon_{m,2} + \varepsilon_{m,3} + \varepsilon_{m,5} + \varepsilon_{m,6}. \end{aligned}$$

Finally let us give a construction of σ_m that will prove useful in Section 6. We can enlarge the probability space Ω, provided by the Skorohod representation Theorem, to $\Omega \times [0,1]^{\{1,2,3,\dots\}}$, obtaining a sequence of independent random variables $(u_k)_{k\geq 1}$, uniform on $[0,1]$ and independent of the sequence $(\alpha_m)_{m\geq 1}$, and we let

$$\sigma_m(k) = \#\left\{ i \mid 1 \leq i \leq m \text{ and } u_i \leq u_k \right\}.$$

Note that with this construction of the sequence $U^{(n)}$, $U^{(n)}$ cannot be obtained by erasing the last term of $U^{(n+1)}$, as usual. If it was the case, α_n would not converge uniformly to b, due to Finkelstein's law of the iterated logarithm [18, Theorem 5.1.2]. Incidentally, Finkelstein's law suggests that $\alpha_{m,\ell}(t)$ converges uniformly to b only if we choose $\ell = o(m)$. $\diamondsuit$

Proof of Theorem 4.2. We have seen that, almost surely, for a given λ, the assumptions of Lemma 3.8 hold true for $\zeta(t) = e(t) - \lambda t$, so, still almost surely, they hold true jointly for $0 \leq \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_k$, yielding that

$$\Pr\left(\lim_m \left(\Theta^{(m)}(\lambda_i)\right)_{1\leq i\leq k} = (\Theta(\lambda_i))_{1\leq i\leq k} \right) = 1.$$

Proof of Theorem 1.3.

$$\lim_m \left(\Theta^{(m)}(\lambda_i)\right)_{1\leq i\leq k} = (\Theta(\lambda_i))_{1\leq i\leq k}$$

entails that

$$\lim_m \left(B^{(m)}(\lambda_i)\right)_{1\leq i\leq k} = (B(\lambda_i))_{1\leq i\leq k}.$$

5 Distribution of components of $B(\lambda)$ and $R(\lambda)$

The proofs of Theorems 1.4 and 1.6, that we give in this Section, are more of a combinatorial nature.

5.1 Proof of Theorem 1.4

This proof reduces to explain a one-to-one correspondence between *confined* parking schemes with n cars and ℓ empty places and Pavlov's forests with ℓ rooted trees and n non-root vertices, correspondence in which the sizes of trees and the sizes of blocks are in correspondence too. Then Theorem 1.4 is just a restatement of [34, Theorem 4].

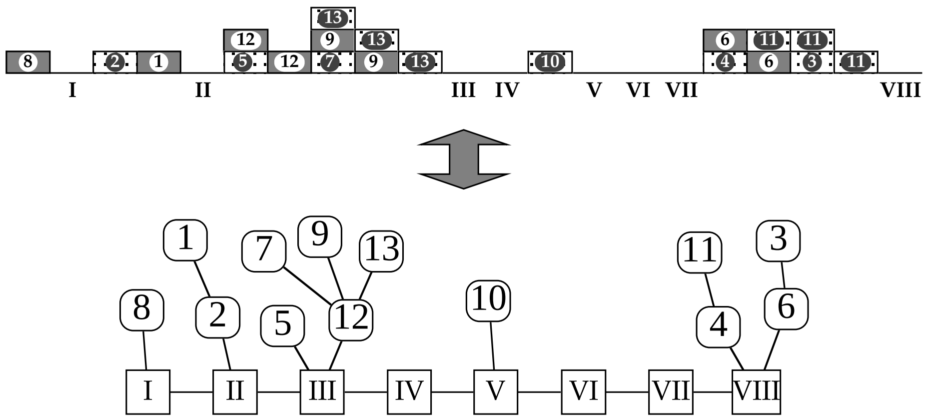

Figure 7: Correspondence parking $\leftrightarrow$ Pavlov's forests, empty places $\leftrightarrow$ roots.

In Pavlov forests, roots and non-roots are labeled separately, for instance the roots (resp. non-roots) are labeled $\{r_1, r_2, \ldots, r_\ell\}$ (resp. $\{v_1, v_2, \ldots, v_n\}$). The label of the root is also the label of the corresponding tree. Let us define the Pavlov forest T corresponding to a given confined parking scheme Π: the non-roots of the first tree of T (that is, of the tree rooted at r_1) are the cars parked before the first empty place of Π, and the non-root vertices of the k^{th} tree are the s_k cars parked between the $k-1^{th}$ and the k^{th} empty places. The way these s_k cars are parked can be described by a confined parking scheme of s_k cars on s_k+1 places: we define the k^{th} tree of T through one among the many one-to-one correspondences between rooted labeled trees with m nodes, and confined parking schemes of $m-1$ cars on m places [15, 25, 26, 49].

The following one-to-one correspondence will be specially useful at Section 8, to explain the relation between parking and the standard additive coalescent. Consider a random labeled tree t with k vertices $v_1, \ldots, v_k$ and let Π denote the corresponding confined parking scheme for $k-1$ cars on k places. The description of Π uses a variant

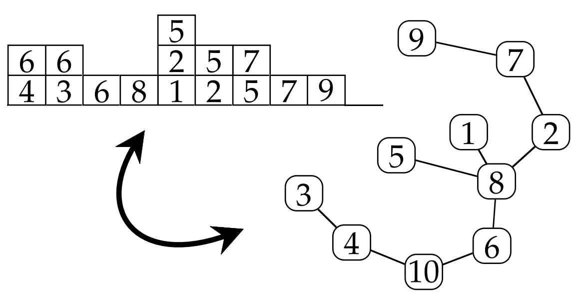

Figure 8: Correspondence parking $\leftrightarrow$ labeled tree

of the breadth first search of t_k: by convention v_k is the root of t ; at step 1, v_k's sons are stored in a queue, *the smallest labels at the head of the queue.* Then at each step the vertex at the head of the queue is removed from the queue, while its sons are added to the queue, and *the queue is reordered* (the smallest labels at the head) to be ready for the following step. The corresponding parking scheme Π is defined by specifying that the first try of car c_j is on place i if and only if the first appearance of v_j in the queue is at step i. In this correspondence, one checks easily that car c_j finally parks at place i if and only if v_j is at the head of the queue at step i, and also that the successive lengths of the queue just give Π's profile.

5.2 Proof of Theorem 1.6

We first provide a useful identity leading to the proof of Theorem 1.6. Set

$$d_i = k_1 + k_2 + ... + k_i.$$

We have

Proposition 5.1

$$\Pr(R_j^{m,\ell} = k_j,\ \ 1 \leq j \leq i) = \prod_{j=0}^{i-1} \varphi(m - d_j - j, n - d_j, k_{j+1}).$$

Proof : The choice of the elements in each of the blocks can be done in

$$\prod_{j=1}^{i} C_{n-d_{j-1}-1}^{k_j-1}$$

ways, and they can be arranged inside each of these blocks in

$$\prod_{j=1}^{i} (k_j+1)^{k_j-1}$$

ways.

It will be convenient to argue in terms of confined parking schemes, since rotations do not change the sizes of blocks. The total number of confined parking schemes is $m^{n-1}\ell$. We obtain a confined parking scheme with blocks' sizes k_1, k_2, etc ... , for the i first blocks, respectively, by inserting these i blocks successively, with an empty place attached to the right of them, insertion taking place at the front of the confined parking scheme Π for the remaining cars, or just after one of the empty places of the confined parking scheme for the remaining cars. There are $(m - d_i - i)^{n-d_i-1}(m - n - i)$ choices for Π, $m - n - i + 1$ possible insertions for the first block, $m - n - i + 2$ possible insertions for the second block, and so on ... Finally, the probability $p(k)$ on the left hand of Proposition 5.1 is given by

$$p(k) = \frac{(m - d_i - i)^{n-d_i-1}(m - n - i)}{m^{n-1}(m - n)} \prod_{j=1}^{i} (k_j + 1)^{k_j-1}(m - n + j - i)C_{n-d_{j-1}-1}^{k_j-1}.$$

It is not hard to check that this last expression is the same as the right hand of Proposition 5.1. $\diamondsuit$

Proof of Theorem 1.6. Set $S_0 = s_0 = 0$ and

$$\begin{aligned} S_j &= R_1(\lambda) + R_2(\lambda) + ... + R_j(\lambda), \\ s_i &= x_1 + x_2 + ... + x_i. \end{aligned}$$

According to Theorem 1.5 (that will be proved independently at Subsection 7.1), the law of $(R_1(\lambda), R_2(\lambda), ..., R_i(\lambda))$ has the following alternative characterization: for any k, conditionally, given $(R_1(\lambda), R_2(\lambda), ..., R_k(\lambda))$, $R_{k+1}(\lambda)$ is distributed as

$$(1 - S_k) \frac{N_{k+1}^2}{\frac{\lambda^2}{1-S_k} + N_{k+1}^2},$$

in which N_{k+1} is standard Gaussian and independent of S_k. In other terms, $R_{k+1}(\lambda)$ has the following conditional density:

$$\frac{1}{1 - S_k} f\left(\frac{\lambda}{\sqrt{1 - S_k}}, \frac{x_k}{1 - S_k}\right),$$

and $R_1(\lambda)$ has the unconditional density $f(\lambda, x)$. On the other hand, using the same line of proof as in Theorem 2.1, the approximations of Lemma 2.2 for $\varphi(m, n, k)$ and Proposition 5.1 lead, for $\left(R_1^{(m)}(\lambda), R_2^{(m)}(\lambda), ..., R_i^{(m)}(\lambda)\right)$, to the following limit density:

$$\prod_{j=0}^{i-1} \frac{1}{1 - s_j} f\left(\frac{\lambda}{\sqrt{1 - s_j}}, \frac{x_{j+1}}{1 - s_j}\right). \qquad \diamondsuit$$

6 Sampling excursions of $\Psi_\lambda e$

In this Section, we give the proofs of Theorems 1.7 and 1.10. They make essential use, to build the parking schemes, of the random permutation of jumps of α_m defined at Section 4.

6.1 Proof of Theorem 1.7.

We build a probability space where almost sure convergence of $R_1^{(m)}(\lambda)$ to $R_1(\lambda)$ holds for any λ. As a consequence, for any k and any $(\lambda_1, \lambda_2, \ \dots \ , \lambda_k)$,

$$\left(R_1^{(m)}(\lambda_1), R_1^{(m)}(\lambda_2), \ \dots \ , R_1^{(m)}(\lambda_k)\right) \xrightarrow{a.s.} (R_1(\lambda_1), R_1(\lambda_2), \ \dots \ , R_1(\lambda_k)),$$

entailing the result.

As in Section 4, we enlarge the probability space to $\Omega \times [0, 1]^{\{1,2,3,...\}}$, obtaining a sequence of i.i.d. random variables $(u_k)_{k\geq 1}$, uniform on $[0, 1]$ and independent of b and $(\alpha_m)_{m\geq 1}$. We let

$$\begin{aligned} \sigma_m(k) &= \#\{i \mid 1 \leq i \leq m \text{ and } u_i \leq u_k\}, \\ U_k^{(m)} &= J_{\sigma_m(k)}^{(m)}, \\ \pi_m &= \left\{U_1^{(m)} - v_m(\lambda)\right\}, \end{aligned}$$

and $\rho_1 = \{u_1 - v\}$, so that ρ_1 is uniform and independent of e. We still assume that first try of car c_k is place $\left\lceil mU_k^{(m)} \right\rceil$. Thus, counted from $V\left(m, \lceil \lambda\sqrt{m} \rceil\right)$, car c_1 parks at place $\lceil m\pi_m \rceil$. Borel-Cantelli Lemma yields that

Lemma 6.1 *Almost surely,*

$$\lim_m \left(U_1^{(m)}, \pi_m\right) = (u_1, \rho_1).$$

Let $g_m(\lambda)$ (resp. $d_m(\lambda)$) be the last zero of $t \to \tau_{v_m(\lambda)}\psi_m(\lambda, t)$ on the left of π_m (resp. the first zero on the right). That is, $mg_m(\lambda) - 1$ (resp. $md_m(\lambda)$) is the empty place at the beginning (resp. at the end) of the block containing car c_1, *counted from* $V(m, \lceil \lambda\sqrt{m} \rceil)$. Thus

$$d_m(\lambda) - g_m(\lambda) = R_1^{(m)}(\lambda).$$

Almost surely, due to Lemmata 3.7 and 6.1,

$$\lim_m Z_m(\lambda, \pi_m) = z(\rho_1).$$

Due to Lemma 3.4, the minimum value of $t \to Z_m(\lambda, t)$ on $(-\infty, \pi_m]$ is the value of $Z_m(\lambda, .)$ on the interval $\left[g_m(\lambda), g_m\left(\lambda + \frac{1}{m}\right)\right)$. On the other hand, due to the Cameron-Martin-Girsanov formula, almost surely, $t \to z(t)$ has only one minimum on the interval $(-\infty, \rho_1]$, but by definition of Ψ_λ this unique minimum is at $g(\lambda)$. Thus, by uniform convergence of $Z_m(\lambda, t)$ to $z(t)$, almost surely, $\lim g_m(\lambda) = g(\lambda)$. Still by definition of Ψ_λ, $d(\lambda)$ is the first hitting time of level $z\left(g(\lambda)\right)$ after ρ_1:

$$d(\lambda) = \inf\{t \geq \rho_1 \ : \ z(t) \geq z\left(g(\lambda)\right)\},$$

but due to Proposition 3.4,

$$\begin{aligned} Z_m\left(\lambda, g_m(\lambda)\right) &= Z_m\left(\lambda, d_m(\lambda)\right) + \frac{1}{\sqrt{m - \lceil \lambda\sqrt{m} \rceil}}, \\ \lim_m Z_m\left(\lambda, d_m(\lambda)\right) &= z\left(g(\lambda)\right), \end{aligned}$$

so that $\liminf d_m(\lambda) \geq d(\lambda)$. Because $d(\lambda)$ is a stopping time, almost surely there exists a sequence $t_k \downarrow d(\lambda)$ such that

$$z(t_k) < z(d(\lambda)) = z(g(\lambda)).$$

Thus

$$\lim_m Z_m\left(\lambda, t_k\right) < \lim_m Z_m\left(\lambda, d_m(\lambda)\right),$$

and $d_m(\lambda) < t_k$ for m large enough. Finally, almost surely,

$$\begin{aligned} \lim_m d_m(\lambda) &= d(\lambda), \\ \lim_m R_1^{(m)}(\lambda) &= R_1(\lambda). \end{aligned} \tag{6.17}$$

6.2 Proof of Theorem 1.10

We essentially do the same surgery on $h_m(t) = \psi_m(\lambda, t)$ as we did on $\Psi_\lambda e$ at Figure 4: the analogs, for h_m, of properties (i), (ii) and (iii) of Theorem 1.10 are combinatorial properties of parking schemes. Going to the limit with the help of Theorem 3.1 then yields Theorem 1.10. We set

$$n = \lfloor m - \lambda\sqrt{m} \rfloor \text{ and } \ell = \lceil \lambda\sqrt{m} \rceil,$$

and we consider the probability space of Subsection 6.1, enlarged to obtain a uniform random variable w, independent of $((\alpha_m)_{m\geq 1}, b, e, v, (u_i)_{i\geq 1})$. From a parking scheme of n cars on m places, generated with the help of $\alpha_{m,\ell}$ as in Section 4, we obtain a profile $h_m(t) = \psi_m(\lambda, t)$, and we have in mind to decompose it as shown on Figure 9: extend h_m to a periodic function on the line, and set

$$\begin{aligned} q_m &= \left(\tau_{v_m(\lambda)} h_m\right)^{[g_m(\lambda), d_m(\lambda)]}, \\ r_m(t) &= \left(\tau_{v_m(\lambda)} h_m\right)^{[d_m(\lambda), 1+g_m(\lambda)]} \\ w(m) &= \frac{\left\lceil \left(m - 1 - R_1^{m,\ell}\right) w \right\rceil}{m - 1 - R_1^{m,\ell}}. \end{aligned}$$

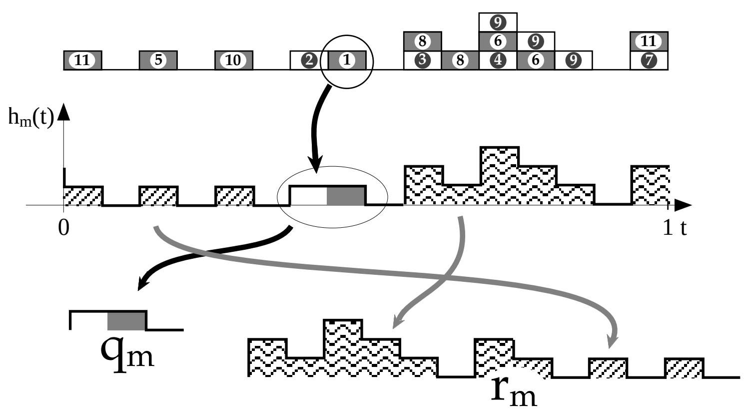

Figure 9: Decomposition of the profile h_m

From relation (2.6) $(n \leq m - 2)$, there are

$$C_{n-1}^{k-1} m (k+1)^{k-1} (m-k-1)^{n-k-1} (\ell - 1)$$

parking schemes such that the block containing c_1 has k cars ; for $C_{n-1}^{k-1} m$ of these parking schemes, the k cars of the block containing c_1 are parked according to a given parking scheme Π, and the remaining $n - k$ cars are parked according to another given parking scheme $\tilde{\Pi}$: C_{n-1}^{k-1} choices for the elements of the block containing c_1, m choices for the position of this block. Thus, according to (2.6), the conditional probability of the parking schemes $(\Pi, \tilde{\Pi})$, given that $mR_1^{(m)}(\lambda) = k$, is

$$\frac{1}{(k+1)^{k-1}} \times \frac{1}{(m-k-1)^{n-k-1}(m-n-1)}.$$

That is, we have:

Proposition 6.2 *Given $R_1^{(m)}(\lambda)$, r_m and q_m are the profiles of independent random uniform confined parking schemes.*

This Proposition is the discrete analog of Theorem 1.10, so, to end the proof, we just have to go (carefully) to the limit. In order to do that we need additional notations: let C be the space of continuous functions on $[0,1]$, with the topology of uniform convergence, and let D be the space of cadlag functions on the same interval, embedded with the Skorohod topology (see [13, Ch. 3]). The triplet of independent random variables $(\Psi_\lambda e, \rho_1, w)$ with value in $C \times [0,1]^2$ defines the random variable $(R_1(\lambda), q, \tau_w r)$ and its law Q, that is a probability measure on the space $[0,1] \times C^2 \subset [0,1] \times D^2$. The normalized Brownian excursion e $\left(\text{resp. } \tau_w \left(\Psi_{\frac{\lambda}{\sqrt{1-x}}} e\right)\right)$ defines the probability measure ν (resp. μ_x) on C. Now Theorem 1.10 is equivalent to:

$$\begin{aligned}\int_0^1 f(\lambda,x) \int_{C^2} \Phi(x,y,z)\mu_x(dz)\nu(dy)dx &= E[\Phi(R_1(\lambda), q, \tau_w r)] \qquad (6.18)\\ &= \int_{R\times C^2} \Phi(x,y,z) Q(dx,dy,dz),\end{aligned}$$

for any bounded uniformly continuous function Φ on the space $[0,1] \times D^2$. It is harmless to assume that $\Phi = 0$ outside $[a,b] \times D^2$, for some choice $0 < a < b < 1$. On a probability space (Ω, A, P), we already exhibited a triplet (e, ρ_1, w) and a sequence of $[0,1] \times D^2$-valued random variables $\left(R_1^{(m)}(\lambda), q_m, \tau_{w(m)} r_m\right)$, satisfying:

(1) almost surely in Ω, $\left(R_1^{(m)}(\lambda), q_m, \tau_{w(m)} r_m\right)$ converges to $(R_1(\lambda), q, \tau_w r)$ for the product topology of $[0,1] \times D^2$;

(2) $\Pr\left(R_1^{(m)}(\lambda) = \frac{k}{m}\right) = \varphi(m,n,k)$, in which $n = \lceil m - \lambda\sqrt{m}\rceil$;

(3) the conditional law, ν_k, of q_m given that $R_1^{(m)}(\lambda) = \frac{k}{m}$, does not depend on m and satisfies:
$$\nu_k \overset{weakly}{\longrightarrow} \nu;$$

(4) the conditional law, $\mu_{m,k}$, of $\tau_{w(m)} r_m$ given that $R_1^{(m)}(\lambda) = \frac{k}{m}$, satisfies:
$$\mu_{m,k} \overset{weakly}{\longrightarrow} \mu_x \qquad \text{if } m \longrightarrow \infty \text{ and } k/m \longrightarrow x \in]0,1[;$$

(5) conditionally, given that $R_1^{(m)}(\lambda) = \frac{k}{m}$, $\tau_{w(m)} r_m$ and q_m are independent.

As $R_1^{(m)}(\lambda) = d_m(\lambda) - g_m(\lambda)$, convergence of the first component in point (1) is just (6.17). Uniform convergence of h_m to $\tau_v \Psi_\lambda e$, uniform continuity of $\Psi_\lambda e(t)$ and (6.17) entails the uniform convergence of q_m to q and the uniform convergence of $\tau_{w(m)} r_m$ to $\tau_w r$, completing point (1). Point (2) is just relation (2.6). As a consequence of Proposition 6.2, given that $R_1^{(m)}(\lambda) = k$, q_m is the profile of a random uniform confined parking scheme of k cars on $k+1$ places, so q_m converges weakly to a normalized Brownian excursion, as a special case $\lambda = 0$ of Theorem 3.1 (see also [15, Section 4]). That is, ν_k converges weakly to ν, giving point (3). Similarly, given that $R_1^{(m)}(\lambda) = k$, r_m is the profile of a random uniform confined parking scheme for the $n-k$ remaining cars on the $m-k-1$ remaining places, and $(m-k-1)w(m)$ is random uniform on $\{1, 2, \ \ldots \ , m-k-1\}$, so $\tau_{w(m)} r_m$ is the profile of a random uniform *non-confined* parking scheme of $n-k$ cars on $m-k-1$ places. If $k \simeq xm$, this parking scheme has

$$\ell - 1 = m - n - 1 \simeq \lambda\sqrt{m} \simeq \frac{\lambda}{\sqrt{1-x}}\ \sqrt{m-k-1}$$

empty places, thus Theorem 3.1 applied to the conditional law $\mu_{m,k}$ of $\tau_{w(m)} r_m$ yields point (4). Point (5) is already contained in Proposition 6.2.

As a consequence of (1):

$$\lim_m E\left[\Phi(R_1^{(m)}(\lambda), q_m, \tau_{w(m)} r_m)\right] \quad = \quad E\left[\Phi(R_1(\lambda), q, \tau_w r)\right],$$

for any bounded uniformly continuous function Φ. We shall prove now that properties (2) to (5) are sufficient to insure that, for any choice $0 < a < b < 1$, and for any bounded uniformly continuous function Φ satisfying $\Phi = 0$ outside $[a, b] \times C^2$, we have

$$\lim_m E[\Phi(R_1^{(m)}(\lambda), q_m, \tau_{w(m)} r_m)] \quad = \quad \int_0^1 f(\lambda, x) \int_{C^2} \Phi(x, y, z)\mu_x(dz)\nu(dy)dx,$$

entailing (6.18).

Let M be a bound for $|\Phi|$. Set:

$$\begin{aligned}
E &= \int_0^1 f(\lambda, x) \int_C \int_C \Phi(x, y, z)\mu_x(dz)\nu(dy)dx \\
&= E\left[\Phi(R_1(\lambda), q, \tau_w r)\right] \\
E_{1,m} &= \int_a^b f(\lambda, x) \int_C \int_C \Phi(x, y, z)\mu_{m,\lceil mx\rceil}(dz)\nu_{\lceil mx\rceil}(dy)dx \\
E_{2,m} &= \int_a^b f(\lambda, \lceil mx\rceil/m) \int_C \int_C \Phi(\lceil mx\rceil/m, y, z)\mu_{n,\lceil mx\rceil}(dz)\nu_{\lceil mx\rceil}(dy)dx \\
&= \frac{1}{m}\sum_{k=1}^{m} f(\lambda, k/m) \int_C \int_C \Phi(k/m, y, z)\mu_{m,k}(dz)\nu_k(dy)dx \\
E_{3,m} &= \sum_{k=1}^{m} \varphi(m, n, k) \int_C \int_C \Phi(k/m, y, z)\mu_{m,k}(dz)\nu_k(dy)dx \\
&= E[\Phi(R_1^{(m)}(\lambda), q_m, \tau_{w(m)} r_m)].
\end{aligned}$$

The last equality is a consequence of point (5). By dominated convergence, owing to (3) and (4), $\lim_m E_{1,m} = E$. By uniform continuity of q and Φ, $\lim_m E_{1,m} - E_{2,m} = 0$. Finally $\lim_m E_{2,m} - E_{3,m} = 0$ due to Lemma 2.2. $\diamondsuit$

7 Parking, fragmentation processes and the standard additive coalescent

In this Section, we give the proofs of Theorems 1.5, 1.8 and 1.9. These results are consequences of Theorem 1.10.

7.1 Proof of Theorem 1.5

It should be possible, following the line of proof of Subsection 6.2, to exhibit a space (Ω, A, P) on which there is almost sure convergence of $\frac{1}{m}(R_1^m, R_2^m, \ \dots\ , R_k^m)$ to $(R_1(\lambda), R_2(\lambda), \ \dots\ , R_k(\lambda))$ for each k, therefore yielding Theorem 1.5. We rather borrow the clever idea of [41, Section 6.4], that uses the decomposition of sample paths of a Brownian bridge to compute the distribution of the sequence of widths of its excursions (in that case a Poisson-Dirichlet distribution).

We introduce, as in [41], a sequence $\rho = (\rho_k)_{k\geq 0}$ of uniform random variables, ρ being independent of e. With probability 1, $\Psi_\lambda e(\rho_k) > 0$: if the excursion containing ρ_k has width $B_j(\lambda)$, we define

$$I_k = j,$$

yielding a size-biased permutation of $B(\lambda)$, as explained in the introduction. Set:

$$\begin{aligned} T(1) &= \inf \{i \geq 2 \mid \rho_i \notin [g(\lambda), d(\lambda)]\} \\ T(k+1) &= \inf \{i \geq T(k)+1 \mid \rho_i \notin [g(\lambda), d(\lambda)]\}. \end{aligned}$$

The random variables $\rho_{T(k)}$ are independent and uniformly distributed on $[0, g(\lambda)] \cup [d(\lambda), 1]$, and, almost surely, there exist a unique number $\tilde{\rho}_k \in]0,1[$ such that

$$\rho_{T(k)} = \{d(\lambda) + \tilde{\rho}_k(1 - R_1(\lambda))\} \ ;$$

$\tilde{\rho} = (\tilde{\rho}_k)_{k\geq 1}$ is a sequence of independent random variables, uniform on $[0,1]$, and independent of (e, ρ_1). Set $\theta R(\lambda) = (R_k(\lambda))_{k\geq 2}$: (ρ_1, e) defines $R_1(\lambda)$, but among the $(\rho_i)_{i\geq 2}$, only the $\rho_{T(k)}$ are useful to determine $\theta R(\lambda)$. Actually, up to a multiplicative factor $1 - R_1(\lambda)$, $\theta R(\lambda)$ is the size-biased permutation, built with the help of the sequence $\tilde{\rho}$, of the sequence of widths of excursions of r, or, equivalently, the size-biased permutation, built with the help of the sequence $\hat{\rho} = (\{\tilde{\rho}_k - w\})_{k\geq 1}$, of the sequence of widths of excursions of $\tau_w r$. Clearly $\hat{\rho}$ is a sequence of independent and uniform random variables, independent of (r, w). In view of Theorem 1.10(iii), this leads to

Lemma 7.1 *Given that $R_1(\lambda) = x$, the sequence $\theta R(\lambda) = (R_k(\lambda))_{k\geq 2}$ is distributed as $(1-x)R\left(\frac{\lambda}{\sqrt{1-x}}\right)$.*

Set:

$$s_k = x_1 + \cdots + x_k.$$

Using Lemma 7.1, an easy induction on k gives the two following properties

(1) $(R_j(\lambda))_{1\leq j\leq k}$ has the distribution asserted in Theorem 1.5 ;

(2) conditionally, given that $(R_j(\lambda))_{1\le j\le k} = (x_j)_{1\le j\le k}$, $\theta^k R(\lambda)$ is distributed as $(1-s_k)R\left(\frac{\lambda}{\sqrt{1-s_k}}\right)$,

ending the proof. $\diamondsuit$

7.2 Proof of Theorem 1.8

The operator Ψ_λ has the semigroup property, and, if a and $a+x$ are two zeroes of a nonnegative function f, due to the Brownian scaling,

$$(\Psi_\lambda f)^{[a,a+x]} = \Psi_{\lambda\sqrt{x}}\left(f^{[a,a+x]}\right).$$

Thus, conditionally, given that $(g(\lambda), d(\lambda)) = (a, a+x)$, we have:

$$\begin{aligned}(\Psi_{\lambda+\mu} e)^{[g(\lambda),d(\lambda)]} &= (\Psi_\mu(\Psi_\lambda e))^{[g(\lambda),d(\lambda)]} \\ &= \Psi_{\mu\sqrt{x}}\left((\Psi_\lambda e)^{[g(\lambda),d(\lambda)]}\right) \\ &= \Psi_{\mu\sqrt{x}}\, q.\end{aligned}$$

Still conditionally, ρ_1 is distributed as $a + x\rho_2$ in which ρ_2 is uniform on $]0,1[$, not depending on a. Thus, Theorem 1.10 (ii), with $(q, \mu\sqrt{x}, \rho_2)$ replacing (e, λ, ρ_1), entails that, given $R_1(\lambda) = x$, $(R_1(\lambda+\mu))_{\mu\ge 0}$ is distributed as $(x\, R_1(\mu\sqrt{x}))_{\mu\ge 0}$. Equivalently, by change of variables, the conditional distribution of $(\Sigma(\lambda+\mu))_{\mu\ge 0}$, given that $\Sigma(\lambda) = y$, is the same as the unconditional distribution of

$$\left((1+y)\Sigma\left(\frac{\mu}{\sqrt{1+y}}\right) + y\right)_{\mu\ge 0}.$$

This last statement yields (1.4), by induction on k: assuming that property at rank $k-1$ holds, we see that, given that $\Sigma(\lambda_1) = y$,

$$\begin{aligned}(\Sigma(\lambda_1+\lambda_2+\ldots+\lambda_i))_{2\le i\le k} &\overset{law}{=} \left((1+y)\Sigma\left(\frac{\lambda_2+\ldots+\lambda_i}{\sqrt{1+y}}\right) + y\right)_{2\le i\le k} \\ &\overset{law}{=} \left(y + \frac{\lambda_2^2}{N_2^2} + \ldots + \frac{\lambda_i^2}{N_i^2}\right)_{2\le i\le k}.\end{aligned}$$

Owing to Theorem 1.10,

$$\Sigma(\lambda_1) \overset{law}{=} \frac{\lambda_1^2}{N_1^2}. \quad \diamondsuit$$

7.3 Parking and the additive coalescent

In this subsection, we give an alternative proof of Bertoin's Theorem 1.9: the coalescence of excursions of Ψ_λ, as $\lambda \searrow 0$, has the same law as the coalescence of continuum random trees in the standard additive coalescent of Aldous and Pitman [8]. As we do not claim novelty, our proof will be sketchy at some points, but we hope to show that some properties of the additive coalescent seem natural, once translated to parking schemes.

First let us prove that B has the Markov property, that is:

$$\Pr\left(B(\lambda+\tilde{\lambda}) \in A \mid \mathcal{G}_\lambda\right) = \Pr\left(B(\lambda+\tilde{\lambda}) \in A \mid B(\lambda)\right), \tag{7.19}$$

in which $\mathcal{G}_\lambda$ is a sigma-field that contains all the information about $(B(\lambda))_{0\le\mu<\lambda}$ (the past of the process). Following closely [12, Section 2], let $\mathcal{G}_\lambda$ stand for the P-completed sigma-field generated by $\left(S_t^\lambda\right)_{0\le t\le 1}$, in which

$$S_t^\lambda = \sup_{0\le s\le t} \lambda s - e(s).$$

Bertoin [12] argues that the complement of the support of the Stieltjes measure dS^λ is the union of nonoverlapping open intervals whose lengths are given by $B(\lambda)$, making $B(\lambda)$ $\mathcal{G}_\lambda$-measurable. Bertoin gives furthermore the following Skorohod-like formula, for $0 \le \mu < \lambda$,

$$S_t^\mu = \sup_{0\le s\le t} S_s^\lambda - (\lambda-\mu)s,$$

with the consequence that $B(\mu)$ is $\mathcal{G}_\lambda$-measurable too. Chassaing and Janson [16] give a construction of $\Psi_\lambda e$ from the three independent sequences U, η and $B(\lambda)$, where $\eta = (\eta_k)_{k\ge1}$ is a sequence of independent Brownian excursions and $U = (U_k)_{k\ge1}$ is a sequence of independent random variables uniform on $[0,1]$. In this construction, S^λ depends only on U and $B(\lambda)$, while the way $B(\lambda)$ is fragmented to give $B(\lambda+\tilde{\lambda})$ depends only on $B(\lambda)$ and η, yielding (7.19).

We give some details, because the construction of [16] describes the distribution of the point process $\Theta(\lambda)$ (see Section 4) that keeps track of positions of excursions, and as such, this construction gives some light on the distribution of positions of blocks: in the limit model, the set of zeroes of $\Psi_\lambda e$ (empty places) is so to say Cantor-like, as there infinitely many excursions (blocks) between any given pair of zeroes. First we build a copy $(X_\lambda(t))_{0\le t\le 1}$ of the reflected Brownian bridge $|b|$ conditioned on its local time at zero $L_1(b) = \lambda$, following [40, Section 6]: we place side by side excursions with shape η_i and width $B_i(\lambda)$, the order of excursions being dictated by U, that is, the excursion with width $B_i(\lambda)$ and shape η_i is on the left of the excursion $(B_j(\lambda), \eta_j)$ if $U_i < U_j$. To be formal, set

$$\begin{aligned} g_i &= \sum_{j, U_j < U_i} B_j(\lambda) \\ d_i &= \sum_{j, U_j \le U_i} B_j(\lambda) \\ &= g_i + B_i(\lambda), \end{aligned}$$

and let X_λ be defined, on a dense subset of $[0,1]$, by

$$X_\lambda^{[g_i,d_i]} = \eta_i.$$

Then X_λ is extended by continuity to $[0,1]$. Incidentally, this random ordering of excursions is analog to the random insertion of blocks in a confined parking scheme (cf. Subsection 5.2): the analogy is used in [16] to prove that the limit of profiles of confined parking schemes is X_λ. Then, in the same way as the random rotation of a random *confined* parking scheme gives a random parking scheme, $\Psi_\lambda e$ is obtained by random rotation of X_λ.

The proof of the Markov property requires that this random rotation depends only on U and $B(\lambda)$, not on η, so let us define it. According to [40, Section 6], the local time at 0 of X_λ, denoted $(L_t(X_\lambda))_{0\le t\le 1}$, is defined for $t \in [g_i, d_i]$, by

$$L_t(X_\lambda) = \lambda U_i,$$

thus L_t depends only on U and $B(\lambda)$. According to [16, Theorem 2.6 (i)], almost surely, there exists a unique point v in $[0, 1)$ such that

$$L_v(X_\lambda) - \lambda v = \max_{0\le t\le 1} L_t(X_\lambda) - \lambda t,$$

and $t \to X_\lambda(\{v + t\})$ is a copy of $\Psi_\lambda e$, that is, $\Psi_\lambda e$ is obtained from X_λ through a random rotation v that depends only on U and $B(\lambda)$, not on η. Thus the local time at 0 of $\Psi_\lambda e$, $(L_t(\Psi_\lambda e))_{0\le t\le 1}$, is deduced from $(L_t(X_\lambda))_{0\le t\le 1}$ through the shift v as well, and depends only on U and $B(\lambda)$. Finally, according to [16, Proposition 8.2],

$$L_t(\Psi_\lambda e) = S_t^\lambda,$$

so that $\mathcal{G}_\lambda$ is a subset of $\sigma(B(\lambda), U)$, the P-completed sigma-field generated by $B(\lambda)$ and U. As a by-product of the second part of the proof, we shall see that

$$\sigma\left(B(\lambda + \tilde{\lambda})\right) \subset \sigma(B(\lambda), \eta).$$

These two inclusions, with the independence between $B(\lambda)$, η and U, entail the Markov property (7.19).

Once we know that both $B(\lambda)$ and $Y(\lambda)$ have the Markov property, we just have to check that they have the same transition probabilities, that is:

$$\Pr\left(B(\lambda + \tilde{\lambda}) \in A \mid B(\lambda) = x\right) = \Pr\left(Y(\lambda + \tilde{\lambda}) \in A \mid Y(\lambda) = x\right). \tag{7.20}$$

Let us describe the conditional distribution of $Y(\lambda+\tilde{\lambda})$, given $Y(\lambda) = x = (x_1, x_2, ..)$. To this aim, let Δ_s denote the space of nondecreasing sequences x of nonnegative numbers with $\sum_{i\ge 1} x_i = s$; $B(\lambda)$ and $Y(\lambda)$ are Δ_1-valued random variables. Let $(y_k)_{k\ge 1}$ denote a sequence of independent random variables, y_k being a Δ_{x_k}-valued random variable with the same distribution as $x_k B\left(\tilde{\lambda}\sqrt{x_k}\right)$ or as $x_k Y\left(\tilde{\lambda}\sqrt{x_k}\right)$. According to [12, Section 1],

Proposition 7.2 *Given $Y(\lambda) = x = (x_1, x_2, \cdots)$, $Y(\lambda + \tilde{\lambda})$ is distributed as the decreasing rearrangement of the elements of sequences $y_1, y_2, \ldots$.*

Let us prove that Proposition 7.2 holds also true with B replacing Y, meaning, informally, that each of the clusters x_i of the fragmentation process $B(\lambda)$ starts anew a fragmentation process distributed as $\left(x_i B(\tilde{\lambda}\sqrt{x_i})\right)_{\lambda\ge 0}$. We shall see that, in the case of B, the scaling factors x_i and $\sqrt{x_i}$ come from the Brownian scaling in the definition of $f^{[a,b]}$. These scaling factors can also be foreseen on parking schemes: the time unit for the discrete fragmentation process associated with parking m cars on m places, is the departure of $\sqrt{m}$ cars. Due to the law of large numbers, during one time unit, a given block of cars with size $x_i m$ loses approximately $x_i\sqrt{m} = \sqrt{x_i}\sqrt{x_i m}$ cars, meaning that, for the internal clock of this block, $\sqrt{x_i}$ time units elapsed.

In order to give a formal proof, let g_k (resp. $d_k = g_k + x_k$, η_k) denote the beginning (resp. the end, the shape) of the excursion of $\Psi_\lambda e$ whose length is $B_k(\lambda) = x_k$:

$$\eta_k = (\Psi_\lambda e)^{[g_k,d_k]}.$$

As in Subsection 7.2, we have:

$$\begin{aligned} \left(\Psi_{\lambda+\tilde{\lambda}} e\right)^{[g_k,d_k]} &= \Psi_{\tilde{\lambda}\sqrt{x_k}}\left((\Psi_\lambda e)^{[g_k,d_k]}\right) \\ &= \Psi_{\tilde{\lambda}\sqrt{x_k}}\,\eta_k. \end{aligned}$$

Let y_k denote the decreasing sequence of widths of those excursions of $\Psi_{\lambda+\tilde{\lambda}} e$ that belong to the interval $[g_k, d_k]$: y_k is also, after normalisation by x_k, the decreasing sequence of widths of excursions of $\Psi_{\tilde{\lambda}\sqrt{x_k}}\,\eta_k$. As consequences of Theorem 1.10, or of [16, 40], $B(\lambda)$ and η are independent, and η is a sequence of independent normalized Brownian excursions. Thus, given $B(\lambda)$, the sequences y_k are independent and respectively distributed as $x_k B\left(\tilde{\lambda}\sqrt{x_k}\right)$. We also have clearly

$$\sigma\left(B(\lambda+\tilde{\lambda})\right) \subset \sigma(B(\lambda), \eta).$$

8 Concluding remarks

The additive coalescent has at least two constructions, seemingly quite different, one by Aldous and Pitman, through the continuum random tree, the other by Bertoin through excursions of the family of stochastic processes $(\Psi_\lambda e)_{\lambda\geq 0}$. Actually Aldous and Pitman [8] build the standard additive coalescent as the limit of a discrete model of coalescence-fragmentation: they reverse the time of a discrete fragmentation process that starts with a random unrooted labeled tree (the discrete analog of the continuum random tree), whose edges are erased at random, one after the other (the discrete analog of Poisson cuts). In this paper we show that, asymptotically, parking schemes lead to Bertoin's construction of the additive coalescent. As a first step towards a better understanding of the connection between these two different constructions of the additive coalescent, we show below an explicit connection between the discrete approximations for the additive coalescent, by random forests on one hand, and by parking schemes in the other hand.

Erasing edges, Aldous and Pitman [8] split the tree in smallest subtrees, and obtain a forest-valued stochastic process, but in [8] (as opposed to Pavlov's forests) the forests are *unordered* sets of *unrooted* trees. Then, focusing on the process of sizes of subtrees, and reversing time, Aldous and Pitman obtain a discrete Markovian coalescent, with the following transition probability: when the forest has m nodes and ℓ subtrees the probability that two clusters (subtrees) with sizes x and y merge in a larger subtree with size $x+y$ is [8, Lemma 1]

$$\frac{x+y}{m(\ell-1)}$$

(as opposed to the fragmentation process where the edges are deleted uniformly at random, in the time-reversed process, edges are not added *uniformly*).

There exists a striking similarity between the previous transition probability and relation (1.5), that gives the probability of aggregation of two parking blocks with sizes x and y: if we assume that the mass of a parking block with x cars is actually $x+1$, for instance counting the empty place on the right of this block, then the process of sizes of blocks has the same distribution as the discrete Markovian coalescent considered in [8]. We already knew that these two processes had the same *asymptotic* distribution (the standard additive coalescent), but this is even better, and suggest the following question: are the two underlying richer structures, the forest-valued stochastic process of on one hand, and the parking schemes-valued process on the other hand, isomorphic in any sense ?

The answer is positive, up to a slight change in Aldous & Pitman' s model, that will be explained later. The description of the relation between the two discrete models has several steps: let $(\Pi_k)_{0\le k\le m-2}$ denote the intermediate parking schemes leading to a confined parking scheme Π_{m-1} for $m-1$ cars on m places, through successive arrivals of cars. For $0\le k\le m-1$, let t_k denote the Pavlov's forest associated with Π_k through the one-to-one correspondence described at Subsection 5.1. Then t_k can be obtained from t_{m-1} by erasing edges in a natural but *deterministic* way. To describe it, consider a random labeled tree t_m with m vertices $v_1, \ldots, v_m$. Relabel v_m as r_1: we get a Pavlov's tree $\tilde{t}_{m-1}$ with $m-1$ non-roots. Then erase the edge between v_{m-1} and his father, and relabel v_{m-1} as r_2: we get a Pavlov's tree $\tilde{t}_{m-2}$ with $m-2$ non-roots. In $\tilde{t}_{m-k}$, v_{m-k} belongs to a subtree with root r_i, $i\le k$. To obtain $\tilde{t}_{m-k-1}$ from $\tilde{t}_{m-k}$, first, relabel roots $r_{i+1}, \ldots, r_k$ as $r_{i+2}, \ldots, r_{k+1}$, then erase the edge between v_{m-k} and his father and relabel v_{m-k} as r_{i+1}. It is now easy to check that $\tilde{t}_j = t_j$. In other terms, erasing at random edges of a random rooted labeled tree, as in [8], or erasing successively the edge between v_{m-1} (resp. v_{m-2}, v_{m-3}, ...) and its father, produces discrete Markovian coalescents with the same law, though the underlying forest-valued stochastic processes are different. We do not know if a different one-to-one correspondence forests-parking would produce a nicer description of the forest-valued process associated to parking, but very likely, and as opposed to [8], given the random tree, the cuts would be deterministic.

In order to circumvent this problem, instead of drawing a random labeled tree, as in [8], we draw the unlabeled *random shape* st_m of a rooted labeled tree (that is, st_m is a Galton-Watson tree with Poisson progeny, conditioned to have size m), then we delete st_m's edges one by one in uniform random order. Compared with [8], the change is not fundamental. But this last model is isomorphic to parking: label v_m the root of the random shape, then label v_{m-k} the vertice at the end (starting from v_m) of the k^{th} deleted edge, and we obtain a uniform random labeled tree t_m. With him, comes the associated random parking scheme Π_{m-1}. Focusing on sizes, and eventually reversing time, we get three Markovian coalescents, from

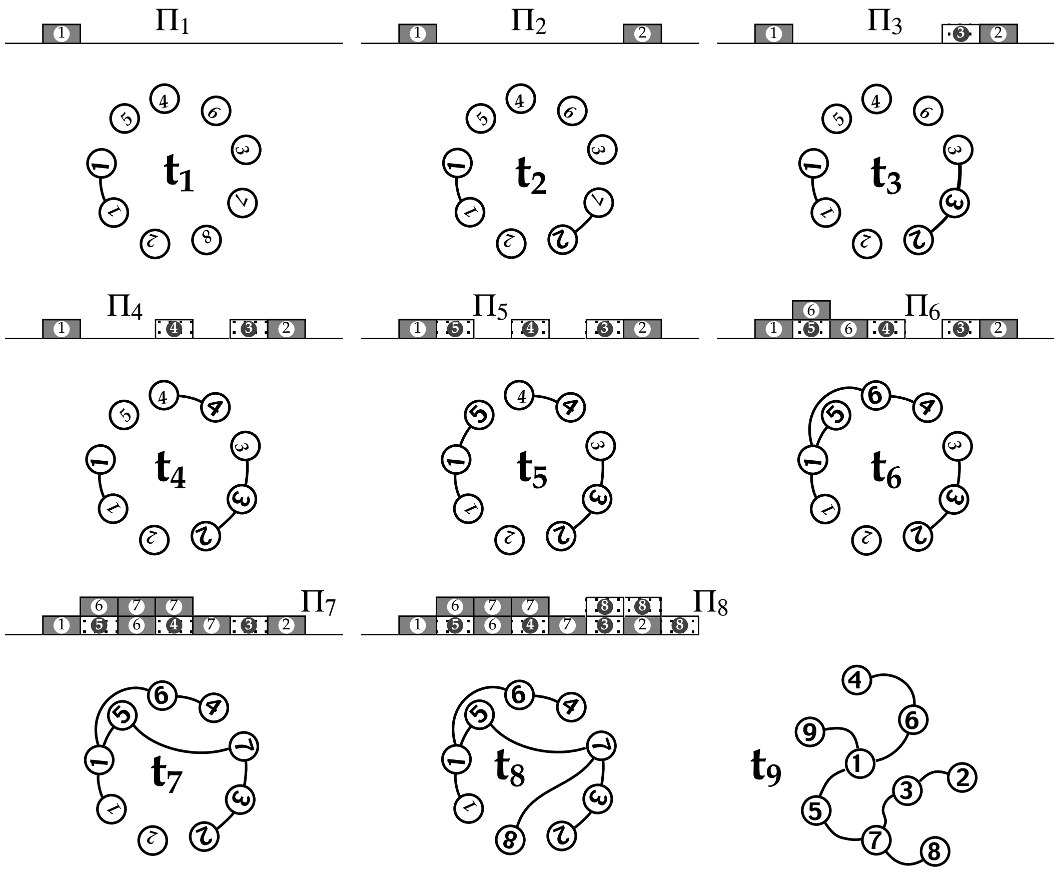

Figure 10: Simultaneous fragmentation of trees and blocks.

st_m, deleting edges at random, from t_{m-1}, erasing successively the edge between v_{m-1} (resp. v_{m-2}, v_{m-3}, ...) and its father, and reversing time, or, finally, from $(\Pi_k)_{0\le k\le m-1}$: these three coalescents turn out to be equal, not only in distribution, but also ω by ω.

A (problematic) translation of this construction to the continuous model would open the way to a direct proof of Theorem 1.9, in which the Poisson process of cuts of the continuum random tree, once translated to Brownian excursion, would give the same cuts of excursions as those obtained from the operators Ψ_λ.

Another possible direction for further research would be to explore connections between this limit model for parking and the reflected Brownian motion with drift, or the Brownian storage process, that appear as heavy traffic limits of queuing or storage systems [27, 32].

Acknowledgements

The starting point for this paper was the talk of Philippe Flajolet at the meeting ALEA in February 1998 at Asnelles (concerning his paper with Viola & Poblete), and a discussion that Philippe and the authors had in a small cafe of the French Riviera after the SMAI meeting of September 1998. Some discussions of the first author with Marc Yor, and also with Uwe Rösler, were quite fruitful. We thank Philippe Laurençot for calling our attention to the works of Aldous, Pitman & Evans on coalescence models. The papers of Perman, Pitman, & Yor about random probability measures and Poisson-Dirichlet processes were also of a great help. Finally joint work [16] with Svante Janson lead to substantial changes to a previous version of this paper. We are specially indebted to Svante for improvements to the proofs of Subsection 3.1, and he also pointed to us the connection with [34]. Finally we thank Régine Marchand and two referees, whose careful reading led to substantial improvements.

References

[1] D. J. Aldous, *Exchangeability and related topics.* Ecole d'ete de probabilites de Saint-Flour XIII - 1983, Lect. Notes Math. 1117, 1-198 (1985).

[2] D. Aldous, *Hashing with linear probing, under non-uniform probabilities.* Probab. Eng. Inform. Sci., 2, 1-14 (1988).

[3] D. Aldous, *The continuum random tree III.* Ann. Probab. 21, 248-289 (1993).

[4] D. J. Aldous, *Brownian excursions, critical random graphs and the multiplicative coalescent.* Ann. Probab. 25, No.2, 812-854 (1997).

[5] D. Aldous, *Stochastic coalescence.* Documenta Mathematica, Extra Volume I.C.M. 1998, III, 205-211 (1998).

[6] D. Aldous, *Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists.* Bernoulli 5, No. 1, 3-48 (1999).

[7] D. J. Aldous & Jim Pitman, *Brownian bridge asymptotics for random mappings.* Random Structures Algorithms 5, No. 4, 487-512 (1994).

[8] D. J. Aldous & Jim Pitman, *The standard additive coalescent.* Ann. of Proba. 26, 1703-1726 (1998).

[9] N. Alon & J.H. Spencer, *The probabilistic method.* With an appendix by Paul Erdos. John Wiley & Sons, Inc., New York-London-Sydney (1992).

[10] R. Arratia, A. D. Barbour & S. Tavaré, *Poisson process approximations for the Ewens sampling formula.* Ann. Appl. Probab. 2, No. 3, 519-535 (1992).

[11] R. Arratia, A. D. Barbour & S. Tavaré, *On Poisson-Dirichlet limits for random decomposable combinatorial structures.* Combin. Probab. Comput. 8, No. 3, 193-208 (1999).

[12] J. Bertoin, *A fragmentation process connected to Brownian motion.* Probab. Theory Relat. Fields 117, 289-301 (2000).

[13] P. Billingsley, *Convergence of Probability Measures.* John Wiley & Sons, Inc., New York-London-Sydney (1968).

[14] B. Bollobás, *Random graphs.* Academic Press, London-New York (1985).

[15] P. Chassaing & J.F. Marckert, *Parking functions, empirical processes and the width of rooted labeled trees.* Electronic J. Combinatorics 8(1), R14 (2001).

[16] P. Chassaing & S. Janson, *A Vervaat-like path transformation for the reflected Brownian bridge conditioned on its local time at* 0, Ann. Probab., October 2001.

[17] P. Chassaing & G. Louchard, *Phase transition for parking blocks, Brownian excursion and coalescence.* Prépublication 99/34 de l'Institut Elie Cartan, Sep. 1999, available at: http://www.iecn.u-nancy.fr/∼chassain

[18] M. Csörgö, P. Révész, *Strong approximations in probability and statistics.* Probability and Mathematical Statistics. Academic Press, Inc., New York-London (1981).

[19] K. Dickman, *On the frequency of numbers containing prime factors of a certain relative magnitude.* Ark. Mat. Astronomi och Fysik 22, 1-14 (1930).

[20] M. D. Donsker, *Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems.* Ann. Math. Statistics 23, 277-281 (1952).

[21] S. N. Evans & J. Pitman, *Construction of Markovian coalescents.* Ann. Inst. Henri Poincaré 34, No. 3, 339-383 (1998).

[22] P. Flajolet, D.E. Knuth & B. Pittel, *The first cycles in an evolving graph.*, Graph theory and combinatorics (Cambridge, 1988). Discrete Math. 75, No. 1-3, 167-215 (1989).

[23] P. Flajolet, P. Poblete, A. Viola, *On the Analysis of Linear Probing Hashing*, Algoritmica Vol. 22, No. 4, 490-515 (1998).

[24] Ph. Flajolet & B. Salvy, *private communication.*

[25] D. Foata, & J. Riordan, *Mappings of acyclic and parking functions.* Aequationes math. 10, 10-22 (1974).

[26] J. Françon, *Acyclic and parking functions.* J. Combinat. Theory Ser. A 18, 27-35 (1975).

[27] J. M. Harrison, *Brownian motion and stochastic flow systems.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York (1985).

[28] S. Janson, D.E. Knuth, T. Łuczak, & B. Pittel, *The birth of the giant component.* Random Structures Algorithms 4, No. 3, 231-358 (1993).

[29] D.E. Knuth,*The art of computer programming. Vol. 3: sorting and searching*, 2nd ed., Addison-Wesley (1997).

[30] D. E. Knuth, *Linear Probing and Graphs*, Algoritmica 22, No. 4, 561-568 (1998).

[31] A.G. Konheim, B. Weiss, *An occupancy discipline and applications*, SIAM J. Appl. Math. 14, No. 6, 1266-1274 (1966).

[32] G. Louchard, *Large finite population queuing systems: the finite server model*, Stoch. Proc. Appl. 55, 117-145 (1994).

[33] P. Massart, *The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality*, Ann. Probab. 18, No.3, 1269-1283 (1990).

[34] Yu. L. Pavlov, *The asymptotic distribution of maximum tree size in a random forest.* Th. Probab. Appl. 22, 509-520 (1977).

[35] M. Perman, *Order statistics for jumps of normalised subordinators.* Stochastic Process. Appl. 46, No. 2, 267-281 (1993).

[36] M. Perman, J. Pitman, M. Yor, *Size-biased sampling of Poisson point processes and excursions.* Probab. Theory Related Fields 92, No. 1, 21-39 (1992).

[37] J. Pitman, *Exchangeable and partially exchangeable random partitions.* Probab. Theory Relat. Fields 102, No.2, 145-158 (1995).

[38] J. Pitman, *Random discrete distributions invariant under size-biased permutation.* Adv. in Appl. Probab. 28, No.2, 525-539 (1996).

[39] J. Pitman & R. Stanley, *A polytope related to empirical distributions, plane trees, parking functions, and the associahedron.* Discrete and Computational Geometry 27, 603-634 (2002).

[40] J. Pitman & M. Yor, *Arcsine laws and interval partitions derived from a stable subordinator.* Proc. London Math. Soc.(3) 65, No. 2, 326-356 (1992).

[41] J. Pitman, M. Yor, *The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator.* Ann. Probab. 25, No. 2, 855-900 (1997).

[42] B. Pittel, *Linear probing: the probable largest search time grows logarithmically with the number of records.* J. Algorithms 8, No. 2, 236-249 (1987).

[43] D. Pollard, *Convergence of stochastic processes.* Springer Series in Statistics. Springer-Verlag, New York-Berlin (1984).

[44] D. Revuz, M.Yor, *Continuous Martingales and Brownian Motion*, Springer, 2^{nd} edition (1994).

[45] J. Riordan, *Ballots and trees.* J. Combinatorial Theory 6, 408-411 (1969).

[46] L. C. G. Rogers, D. Williams, *Diffusions, Markov processes, and martingales. Vol. 1. Foundations.* Second edition. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Ltd., Chichester (1994).

[47] L. A. Shepp, S. P. Lloyd, *Ordered cycle lengths in a random permutation.*, Trans. Amer. Math. Soc. 121, 340-357 (1966).

[48] G. R. Shorack, J. A. Wellner, *Empirical processes with applications to statistics.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York (1986).

[49] M. P. Schutzenberger, *On an enumeration problem.* J. Combinatorial Theory 4, 219-221 (1968).

[50] R. Stanley, *Parking functions and noncrossing partitions.* Electron. J. Comb. 4, No. 2 (1997).

[51] R. Stanley, *Hyperplane arrangements, parking functions, and tree inversions,* in Mathematical Essays in Honor of Gian-Carlo Rota (B. Sagan and R. Stanley, eds.), Birkhäuser, Boston/Basel/Berlin, 259-375 (1998).

[52] G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres.*, Second edition. Cours Spécialisés, Société Mathématique de France, Paris (1995).

[53] W. Vervaat, *A relation between Brownian bridge and Brownian excursion*, Ann. Probab. 7, no. 1, 143-149 (1979).

[54] M. Yor, *Local times and excursions for Brownian motion: a concise introduction*, Lecciones en Matemáticas, Universidad Central de Venezuela (1995).