Parking with density[1]

Jean-François Marckert

Université de Versailles
45 Avenue des Etats Unis
78035 Versailles Cedex
marckert@math.uvsq.fr

ABSTRACT: We study parkings with n places, where $m(n)$ cars are placed according to a nonuniform probability. The aim of this paper is to show a threshold function (depending on the distribution) for the emergence of a giant component. The size of the largest blocks of consecutive occupied places and the total displacement of the cars are also studied.

1 Introduction

The *usual parking problem*, introduced by Knuth [11], p. 545, is the following one: $m(n)$ cars, $(m(n) \leq n)$ try to park on n places labeled $\{p_1, \cdots, p_n\}$. The ith car **chooses** a place $p_{l(i)}$ equally likely among the n places (independently from the other cars). Then, it **parks** on the first empty place among $p_{l(i)}, p_{l(i)+1}, p_{l(i)+2}, \cdots$ (with the convention $p_{n+1} = p_1$, that is the parking is a circle).

In this paper, we consider parkings in which the place of each car is chosen with a probability distribution which is not uniform. Consider $(X_i)_{i=1\cdots m(n)}$ a sequence of i.i.d. random variables on $[0,1]$ with common density f. With these variables we construct a parking denoted

$$Par(n, m(n), f):$$

n is the number of places, $m(n)$ the number of cars and in $Par(n, m(n), f)$ the ith car chooses place $p_{l(i)}$ with

$$l(i) = \lceil n\, X_i \rceil.$$

See Figure (1.1). In $Par(n, m(n), f)$, the ith car chooses place p_k with probability

$$\mathbb{P}_k = \mathbb{P}\Big(X_i \in \Big]\frac{k-1}{n}, \frac{k}{n}\Big]\Big) = \int_{(k-1)/n}^{k/n} f(x)\, dx.$$

Our model can also be interpreted as follows: the interval $[0,1]$ is split into n subintervals $(I_i^{(n)})_{1\leq i\leq n}$ with $I_i^{(n)} =](i-1)/n, i/n]$. $I_i^{(n)}$ is interpreted as the ith parking place, of width $1/n$. Car k chooses place p_i with probability $\mathbb{P}(X_k \in I_i)$. The *usual parking problem* described above appears to be the *uniform* parking, that is $Par(n, m(n), \mathbb{I}_{[0,1]}(x))$.

Several parameters are useful to study and to understand the behavior of the parking: the displacement of a car, that is the number of places it visits before it finds an empty place (at

[1] *AMS* 1991 *subject classifications.* Primary: 68Q25; Secondary: 68P10, 90B40

Key words and phrases: analysis of algorithms, hashing with linear probing, giant component, threshold function, parking

least 1) and the total displacement, that is the sum of all the displacements; more formally, if V_i is the place on which the ith car parks at the end, the displacement of car i is:

$$\begin{array}{ll} V_i - l(i) + 1 & \text{if } l(i) \le V(i) \\ n - l(i) + 1 + V(i) & \text{else.} \end{array}$$

For example, in Figure (1.1), the displacement of car 1 is 1, of car 5 is 4, and the total displacement is 16. Once $m(n)$ cars are parked in the parking, one focuses on the following questions:

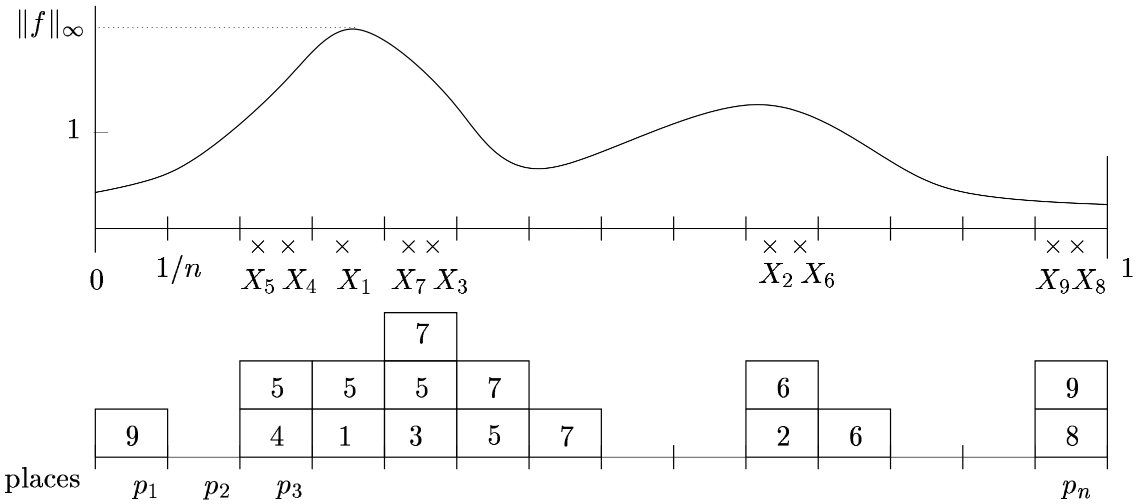

Figure 1.1 : parking with density

- What is the worst case displacement for a $m(n)+1^{th}$ car?
- What is the law of the displacement for a $m(n)+1^{th}$ car?
- What is the law, the mean of the total displacement for the $m(n)$ first cars?

The first two questions are closely related to the size of the **blocks** of consecutive occupied places on the parking after the placement of the first $m(n)$ cars (the size of the blocks of the parking in Figure (1.1) are 5, 2 and 2). The third question (seemingly more complex in appearance) depends on the evolution of the size of the blocks during the "building" of the parking.

The aim of this paper is to study the influence of f on the behavior of each of these variables and to show that given a nonuniform density f, there exists a threshold function for $m(n)$ for the appearance of a giant block (of order n).

In computer science, the parking problem with density corresponds to the study of hashing with linear probing (see following paragraph) on nonuniform data (see Devroye [6], Gonnet [10], Pflug & Kessler [14] and Aldous [1] for similar assumptions). The profound difference between uniform and nonuniform case justifies also a study.

Hashing with linear probing (HLP)

HLP is an important algorithm used (in computer science) to store and to recover stored data on a disk in an economic way (Knuth [11]). According to HLP, a record x is hashed to the cell $h(x)$. If the cell $h(x)$ is not empty, x is hashed in the first empty place among $h(x)+1, h(x)+2, \cdots$ (with the convention $n+1 \equiv 1$). h is called the hashing function, n the size of the table. To recover x one just has to probe position $h(x), h(x)+1, \cdots$. The correspondence with parking is obvious. When m records are hashed, the size of the largest block of consecutive occupied cells (plus one) is the worst case cost (the maximal number of cells to visit) to place a $m(n)+1^{\text{th}}$ item, the total displacement (that is the sum of the number of cells visited), divided by $m(n)$ gives the average cost analysis of HLP.

In order to analyse HLP, one generally assumes that the sequence of records $x_1,\cdots,x_m$ and the hashing function are such that $h(x_1),\cdots,h(x_m)$ are independent, uniformly distributed over $\{1,\cdots,n\}$. For a survey on this topic one can consult Chassaing & Louchard [3]. For a study of hashing under nonuniform probabilities, we refer to Aldous [1]. The point of view of Aldous is different from the one of the present paper since he "adapts" HLP to the nonuniform case; the probings are no linear but randomized in some sense. This way to proceed is very efficient (since he shows that the asymptotic behavior is in the same range as in the uniform case) but not at all realistic in terms of parking (the successive places where cars try to park would be given by a random cycle of the n places).

In the following paragraph, we mention some results about uniform parkings that are significant and/or that are useful for the sequel of this paper.

Results on uniform parking

Denote by $B_1,B_2,\cdots$ the sequence of the block sizes sorted in decreasing order. Denote by $d_{n,m(n)}$ the total displacement for the $m(n)$ first cars. In figure 1.1, $B_1=5, B_2=B_3=2, d_{14,9}=16$

• Full parking ($m(n)=n$). We have $B_1=n$ and

$$\frac{d_{n,n}}{n^{3/2}} \xrightarrow[n\longrightarrow+\infty]{(d),Moments} \int_0^1 e(s)ds$$

where $(e_s)_{0\le s\le 1}$ is a normalized Brownian excursion. This convergence is a moment convergence, and a law convergence. The limit law is the Airy law. See Flajolet & al. [7] for a fine analysis and Chassaing & Marckert [4] for a probabilistic explanation of the presence of the normalized Brownian excursion.

• Parking under construction, sparse parking ($m(n)=\alpha\, n, 0<\alpha<1$).

$$B_1=\frac{1}{c(\alpha)}\big(\log n-\frac{3}{2}\log\log n+C_n\big)-1$$

where C_n is a r.v. which is asymptotically double exponentially distributed (Pittel [13]) and

$$c(x)=x-1-\log x,\quad 0<x<1. \tag{1.1}$$

The total displacement is asymptotically normally distributed [7]:

$$\frac{d_{n,m(n)}-\mu_{n,m(n)}}{\sigma_{n,m(n)}} \xrightarrow{(d)} \mathcal{N}(0,1),$$

where $\mathcal{N}(0,1)$ is a centered normal distributed r.v. with variance 1 and

$$\begin{aligned}\mu_{n,m(n)} &= \mathbb{E}(d_{n,m(n)})=\frac{\alpha}{2(1-\alpha)}m(n)-\frac{\alpha}{2(1-\alpha)^3}+O(n^{-1}), \qquad (1.2)\\ \sigma^2_{n,m(n)} &= Var(d_{n,m(n)})=\frac{6\alpha-6\alpha^2+4\alpha^3-\alpha^4}{12(1-\alpha)^4}m(n)-\frac{6\alpha^3+24\alpha^2+6\alpha}{12(1-\alpha)^6}+O(n^{-1}).\end{aligned}$$

• Phase transition ($m(n) = n - \lambda\sqrt{n}, \lambda > 0$). One finds in Chassaing & Louchard [3] a complete and dynamic (when λ changes) probabilistic study in this case, and its relation with coalescence (see also Chassaing & Janson [2]).

$$\frac{B_1}{n} \xrightarrow{(d)} B(\lambda)$$

where $B(\lambda)$ is the width of the largest excursion of the process

$$\big(\Psi_\lambda e\big)(x) = e(x) - \lambda(x) + \sup_{o\leq y\leq x}\{\lambda y - e(y)\},$$

where e is a normalized Brownian excursion. The joint limit law of $(B_1, B_2, \cdots, B_k)$ is expressed in term of Gaussian law.

2 Parking with density

The first part of this section is devoted to show a threshold phenomena for the appearance of a giant component of consecutively occupied places in the parking. The second part of the section is devoted to the study of the behavior of the parking in the subcritical and in the supercritical case.

2.1 A threshold function for the appearance of a giant component

For a parking Par, denote by $B_1(Par), B_2(Par), \cdots$ the size of the largest blocks sorted in decreasing order. In order to simplify the notations, in the sequel of the paper, we suppose that αn is an integer. We do not assume that f is continuous in this theorem. We define $\|f\|_\infty = \text{ess sup } f$ (resp. $\inf f = \text{ess inf } f$), the essential supremum (resp. infimum) of f.

Theorem 2.1 *(Threshold)*
Suppose that $m(n) = \alpha n$ with $0 < \alpha < 1$. Par denoting the parking $Par(n, \alpha n, f)$, we have:
i) If $\alpha\|f\|_\infty < 1$ then:

$$\frac{1}{c(\alpha \inf f)} \leq \liminf_n \frac{\mathbb{E}(B_1(Par))}{\log n} \leq \limsup_n \frac{\mathbb{E}(B_1(Par))}{\log n} \leq \frac{1}{c(\alpha\,\|f\|_\infty)}.$$

where the function c is defined in (1.1).
ii) If $\alpha\|f\|_\infty > 1$ then there exists a constant $\gamma > 0$ such that

$$\mathbb{P}(B_1(Par)) \geq \gamma\, n) \xrightarrow[n\longrightarrow +\infty]{} 1.$$

The sizes of these components (and their locations) are given in Theorem 2.12.

Case (ii) is to be compared to the uniform case where the size of the largest block is close to $(\log n)/c(\alpha)$.

2.2 The subcritical case: $m(n) = \alpha n$ with $\alpha\|f\|_\infty < 1$

The study of the subcritical parking that we give here, relies on a coupling argument. We construct two auxiliar uniform parkings ($Par1$ and $Par3$) on the same probability space as $Par2 = Par(n, m(n), f)$. Thus, we can compare, realization by realization, the largest block of $Par2$ to $Par1$ and $Par3$ largest blocks (cf. Lemma 2.2 and 2.5, Proposition 2.3 and Corollary 2.6).

Construction of the three parkings $Par1, Par2, Par3$

Consider the sequence of independent random variables $(U_i)_{i\geq 1} = \big((x_i, y_i)\big)_{i\geq 1}$, uniform on $[0,1] \times [0, \|f\|_\infty]$. With this sequence, we define three parkings $Par1, Par2$ and $Par3$, in the following way:

- $Par1, Par2$ and $Par3$ have n places labeled $\{p_1, \cdots, p_n\}$ (on a circle).
- If $\lceil n\, x_i \rceil = k$ then
  - In $Par1$, car i chooses place p_k.
  - In $Par2$, if $y_i \leq f(x_i)$, car i chooses place p_k else it disappears.
  - In $Par3$, if $y_i \leq \inf\{f(x) | 0 \leq x \leq 1\}$, car i chooses place p_k else it disappears.

Cars park until the number of cars in $Par2$ is $\alpha\, n$. If $Par1$ is full before $Par2$ contains αn cars, then we go on with the construction of $Par2$ but stop the one of $Par1$. One shall see that the probability that such an event happens is negligible.

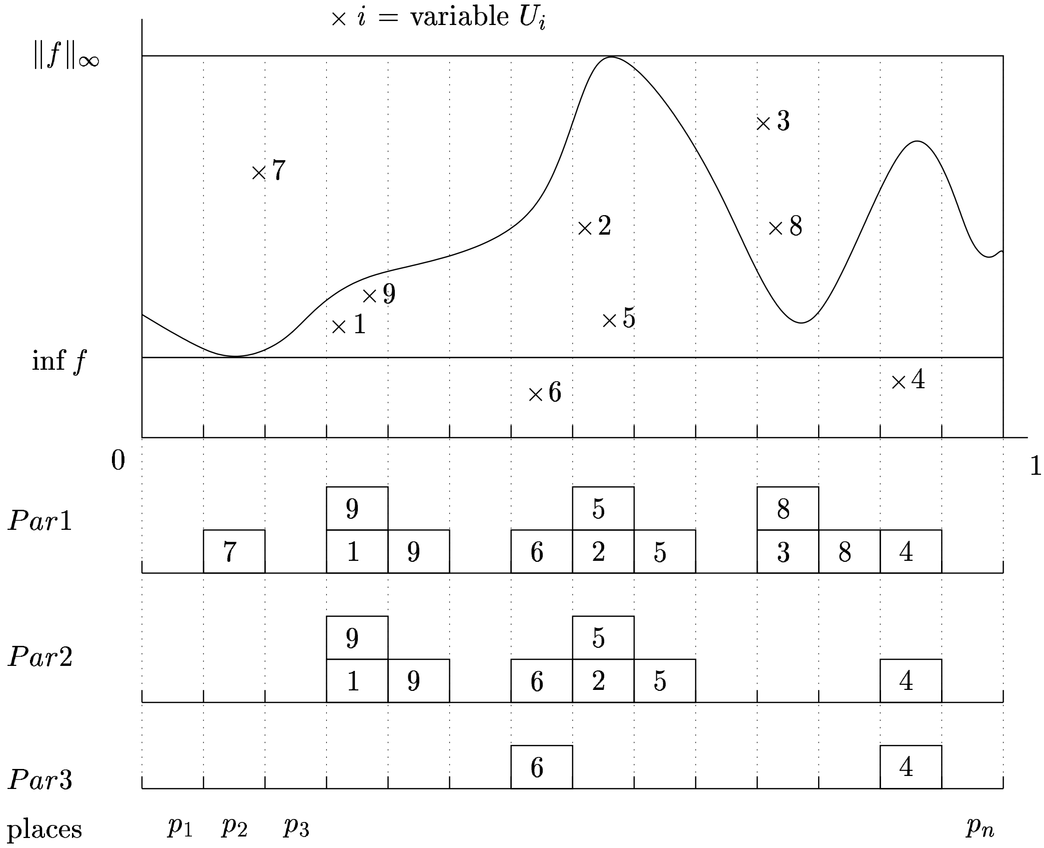

Figure 2.1 : The three parkings

We can compare the size of the largest block of the three parkings:

Lemma 2.2

$$B_1(Par1) \geq B_1(Par2) \geq B_1(Par3).$$

Proof: By construction, $Par1$ contains $Par2$ which contains $Par3$. That is, if place p_k is occupied in $Par(i+1)$ then it is also occupied in $Par(i)$ (for $i = 1$ or 2). □.
The next proposition identifies the constructed parkings: $Par2$ is $Par(n, m(n), f)$, $Par1$ and $Par2$ are uniform parking but contain random numbers of cars.

Proposition 2.3
i) $Par2 \overset{(d)}{=} Par(n, m(n), f)$, $m(n) = \alpha\, n$.

ii) $Par1 \overset{(d)}{=} Par(n, N, \mathbb{I}_{[0,1]}(x))$ *and*

$$N \overset{(d)}{=} \tilde{N}\, \mathbb{I}(\tilde{N} \leq n) + n\, \mathbb{I}(\tilde{N} > n)$$

where $\tilde{N}$ is a negative binomial r.v. with parameter $(\alpha\, n, \|f\|_\infty^{-1})$, that is

$$\mathbb{P}(\tilde{N} = j) = \binom{j-1}{\alpha\, n - 1} (\|f\|_\infty^{-1})^{\alpha\, n} (1 - \|f\|_\infty)^{j - \alpha\, n}\ \mathbb{I}(j \geq \alpha\, n).$$

iii) $Par3 \overset{(d)}{=} Par(n, N', \mathbb{I}_{[0,1]}(x))$ *where N' is binomial with parameter $(\alpha\, n, \inf f)$, that is*

$$\mathbb{P}(N' = j) = \binom{\alpha\, n}{j} (\inf f)^j (1 - \inf f)^{\alpha\, n - j}.$$

Proof : (i) $Par2$ contains αn cars which choose places independently from each other. A trite computation shows that a car that parks in $Par2$ chooses place p_k with probability $\mathbb{P}_k$.
(ii) By construction, $Par1$ is associated with the uniform distribution on $[0, 1]$. The number of cars N that park on $Par1$ is the number of cars that try to park on $Par2$ (successfully or not). For one car on $Par2$, there are $G(\|f\|_\infty^{-1})$ cars that try to park on $Par1$, where $G(\|f\|_\infty^{-1})$ denotes a classical geometric r.v. with parameter $\|f\|_\infty^{-1}$. Indeed

$$\mathbb{P}(\text{car 1 parks on Par2}) = \frac{\int_0^1 f(x)dx}{\|f\|_\infty} = \frac{1}{\|f\|_\infty}.$$

Thus, if N_1 denotes the label of the first car that parks on $Par2$:

$$\mathbb{P}(N_1 = j) = (1/\|f\|_\infty)(1 - 1/\|f\|_\infty)^{j-1}.$$

Define $\tilde{N}$, the sum of $\alpha\, n$ i.i.d. r.v. with the same law as N_1. We obtain:

$$N \overset{(d)}{=} \tilde{N}\, \mathbb{I}(\tilde{N} \leq n) + n\, \mathbb{I}(\tilde{N} > n)$$

since the number of cars in $Par1$ is bounded by the number of places, n.
(iii) The probability for a car that parks in $Par2$ to also parks in $Par3$ is $\inf f$.□

Pittel's results apply when the number of cars is known (fixed). In $Par1$ and in $Par3$, the number of cars is random. Using large deviation techniques, we bound the number of cars in $Par1$ and in $Par3$ with probability exponentially close to 1. Then, we use the fact that the size of the largest block (in a given parking) is an increasing function of the number of the cars.

The following lemma, due to Cramér, gives a large deviation principle for N and N'.

Lemma 2.4 *(Cramér)*
(a) For $x \geq \|f\|_\infty$,

$$\mathbb{P}\big(\frac{\tilde{N}}{\alpha\, n} \geq x\big) \leq \exp(-n\,\Lambda_1^\star(x)).$$

(b) For $x \leq \inf f$,

$$\mathbb{P}\big(\frac{N'}{\alpha\, n} \leq x\big) \leq \exp(-n\,\Lambda_2^\star(x)).$$

where

$$\Lambda_1^\star(x) = \begin{cases} +\infty & \textit{if} \quad x \leq 1 \\ x\,\log\big(\frac{x-1}{x(1-\|f\|_\infty^{-1})}\big) - \log\big(\frac{\|f\|_\infty^{-1}(x-1)}{1-\|f\|_\infty^{-1}}\big) & \textit{if} \quad x > 1 \end{cases}$$

and

$$\Lambda_2^\star(x) = \begin{cases} +\infty & \textit{if} \quad x > 1 \textit{ or } x < 0 \\ x\log\big(\frac{1-\inf f}{\inf f}\frac{x}{1-x}\big) - \log\big(\frac{1-\inf f}{1-x}\big) & \textit{if} \quad x \in [0,1] \end{cases}$$

Proof: For a proof of Cramér's theorem see [5], chap. 2.2, proof of Theorem 2.2.3.. $\tilde{N}$ is the sum of αn independent geometric r.v. with parameter $1/\|f\|_\infty$ and N' the sum of αn independent Bernoulli r.v. with parameter $\inf f$. Let g_1 (resp. b_1) be a geometric r.v. with parameter $1/\|f\|_\infty$ (resp. a Bernoulli r.v. with parameter $\inf f$). Set

$$\Lambda_1(\lambda) = \log(\mathbb{E}(e^{\lambda g_1})), \qquad \Lambda_2(\lambda) = \log(\mathbb{E}(e^{\lambda b_1}))$$

and

$$\Lambda_1^\star(x) = \sup_{\lambda\in\mathbb{R}}\{\lambda\, x - \Lambda_1(\lambda)\}, \qquad \Lambda_2^\star(x) = \sup_{\lambda\in\mathbb{R}}\{\lambda\, x - \Lambda_2(\lambda)\}$$

the Fenchel-Legendre transform of Λ_1 and Λ_2. One has, $\Lambda_1^\star \geq 0$ and $\Lambda_1^\star(x) = 0 \Leftrightarrow x = \|f\|_\infty$; one also has, $\Lambda_2^\star \geq 0$ and $\Lambda_2^\star(x) = 0 \Leftrightarrow x = \inf f$.□
Cramèr's Lemma implies that (with probability that tends exponentially fast to 1), we have:

$$N \leq \alpha\, n(\|f\|_\infty + \varepsilon) \text{ and } N' \geq \alpha\, n(\inf f - \varepsilon).$$

The number of cars in the two uniform parkings $Par1$ and $Par3$ are then (roughly speaking) essentially bounded. Then, by Pittel's result [13],

$$\frac{\mathbb{E}(B_1(Par(n, \beta n, \mathbb{I}_{[0,1]}(x))))}{c(\beta)^{-1}\log n} \longrightarrow 1, \tag{2.3}$$

we obtain:

Lemma 2.5 *For any* $\varepsilon > 0$,

i) $\limsup_n \dfrac{\mathbb{E}(B_1(Par1))}{\mathbb{E}(B_1(Par(n, \alpha\, n(\|f\|_\infty + \varepsilon)), \mathbb{I}_{[0,1]}(x)))} \leq 1$ *(for* ε *such that* $\alpha(\|f\|_\infty + \varepsilon) < 1$*)*

ii) $\liminf_n \dfrac{\mathbb{E}(B_1(Par3))}{\mathbb{E}(B_1(Par(n,\alpha\, n(\inf f-\varepsilon)),\mathbb{I}_{[0,1]}(x)))} \geq 1$ *(for ε such that $\inf f-\varepsilon>0$)*

Proof : Consider A and B two integers such that $0<A\leq B\leq n$, then:

$$\mathbb{E}(B_1(Par(n,A,\mathbb{I}_{[0,1]}(x)))) \leq \mathbb{E}(Par(B_1(n,B,\mathbb{I}_{[0,1]}(x)))). \tag{2.4}$$

For any $\varepsilon>0$ such that

$$\alpha\,(\|f\|_\infty+\varepsilon)<1,$$

one has:

$$\begin{array}{rcl} N & \leq & \alpha\, n(\|f\|_\infty+\varepsilon)\mathbb{I}_{\tilde{N}\leq\alpha\, n(\|f\|_\infty+\varepsilon)}+n\mathbb{I}_{\tilde{N}\geq\alpha\, n(\|f\|_\infty+\varepsilon)} \\ N' & \geq & \alpha\, n(\inf f-\varepsilon)\mathbb{I}_{N'\geq\alpha\, n(\inf f-\varepsilon)}. \end{array}$$

Using (2.4), we obtain:

$$\begin{array}{rcl} \mathbb{E}(B_1(Par1)) & \leq & \mathbb{E}\big(B_1(n,\alpha n(\|f\|_\infty+\varepsilon),\mathbb{I}_{[0,1]}(x))\big)\,\mathbb{P}(\tilde{N}\leq\alpha n(\|f\|_\infty+\varepsilon)) \\ & & +n\;\mathbb{P}(N>\alpha n(\|f\|_\infty+\varepsilon)) \end{array}$$

and

$$\mathbb{E}(B_1(Par3))\geq\mathbb{E}\big(B_1(n,\alpha\, n(\inf f-\varepsilon),\mathbb{I}_{[0,1]}(x))\big)\;\mathbb{P}(N'>\alpha(\inf f-\varepsilon)).$$

Indeed, the largest block is smaller than n. Lemma 2.5 follows from these inequalities and from Lemma 2.4.□

The proof of Theorem 2.1 (i) is completed by using Lemma 2.5 and the continuity of the function c. The construction of the three parkings $Par1, Par2$ and $Par3$ and Lemma 2.4 allows us also to bound the mean total displacement in $Par(n,m(n),f)$ (using (1.2)):

Corollary 2.6 *(Total displacement in sparse parking)*
In $Par(n,\alpha n,f)$, for $0<\alpha<\|f\|_\infty$, one has:

$$\frac{\mu_{n,(\alpha\inf)n}}{n}\leq\liminf_n\frac{\mathbb{E}(d_{n,\alpha\, n})}{n}\leq\limsup_n\frac{\mathbb{E}(d_{n,\alpha\, n})}{n}\leq\frac{\mu_{n,(\alpha\|f\|_\infty)n}}{n}$$

where $\mu_{n,m(n)}$ is defined in equation (1.2).

Pflug and Kessler [14] obtained a better result since they proved that

$$\frac{\mathbb{E}(d_{n,\alpha\, n})}{n}\underset{n\longrightarrow+\infty}{\longrightarrow}\int_0^1\frac{1-(2\alpha-1)f(x)+\alpha^2f^2(x)}{2(1-\alpha f(x))^2}dx.$$

They also give a similar result for the average cost of the successful search in the table.

2.3 The supercritical case: $m(n) = \alpha n$ with $\alpha\|f\|_\infty > 1$

Let Y_k be the number of cars that choose place p_k and H_k the number of cars that try (at first or later) to park on place p_k. In Figure (1.1), $H_1 = 1, H_2 = 0, H_3 = 2, H_4 = 2, \cdots$ and $Y_1 = 0, Y_2 = 0, Y_3 = 2, Y_4 = 1, \cdots$. The sequence $(H_k)_{k=1,\cdots,n}$ is called the **profile** of the parking (this notion is already used in Chassaing & Marckert [4]). The profile obviously contains all the information relative to the parking. A block of consecutive occupied places is equivalent to a block of consecutive indices k such that $H_k > 0$. We have:

$$Y_k = \sum_{i=1}^{\alpha n} \mathbb{I}_{X_i \in](k-1)/n, k/n]}.$$

The sequences $H.$ and $Y.$ are related by:

$$H_{k+1} = (H_k - 1)_+ + Y_{k+1}, \tag{2.5}$$

since all the cars that try to park on place p_k minus 1 (the ones which stays on place p_k) try also to park on place p_{k+1}. Relation (2.5) on its own does not define the parking profile since it is a "circular relation", but:

$$H_k = \max_{l \geq 0} \left\{ \dot{Y}_k + \cdots + \dot{Y}_{k-l} - l \right\} \tag{2.6}$$

where

$$\dot{Y}_k = Y_{k \bmod n} \text{ for } k \neq 0 \text{ and } Y_0 = Y_n.$$

Indeed, similarly, we define:

$$\dot{p}_k = p_{k \bmod n} \text{ for } k \neq 0 \text{ and } p_0 = p_n.$$

One observes then, that on the places $\{\dot{p}_{k-1}, \dot{p}_{k-2}, \cdots, \dot{p}_{k-l}\}$, at most l cars can park; the ones in excess, plus Y_k try to park on place p_k.

Proof of Theorem 2.1 (ii). For any bounded f, with $\|f\|_\infty = \text{ess sup}_{[0,1]} f$, $\alpha\|f\|_\infty > 1$, there exists an $x \in]0,1[$ and a constant $\beta > 0$ such that

$$\frac{1}{\delta} \int_{I_\delta(x)} f(y)dy > \frac{1}{\alpha}(1+\beta)$$

for all δ small enough, where $I_\delta(x)$ is any interval of length δ containing x. We apply the Lebesgue density theorem. Let $[a,b]$ denotes one particular such interval. The number of cars $Y_{\lceil an+1 \rceil} + \cdots + Y_{\lfloor bn \rfloor}$ that arrive on the places $\{p_{\lceil an+1 \rceil}, \cdots, p_{\lfloor bn \rfloor}\}$ is stochastically larger than a binomial $B\big(\alpha n, \delta(1+\beta)/\alpha\big)$. Since

$$\alpha n \frac{\delta(1+\beta)}{\alpha} = \delta n(1+\beta) = n(b-a)(1+\beta),$$

we obtain

$$\mathbb{P}\Big(Y_{\lceil an+1 \rceil} + \cdots + Y_{\lfloor bn \rfloor} - n(b-a) \geq \frac{\beta}{2} n(b-a)\Big) \longrightarrow 1$$

The statement (ii) of Theorem 2.1 holds with $\gamma = \frac{\beta}{2}(b-a)$. $\square$

Limit of the profile

The aim of this paragraph is to study the limit profile in order to obtain some results about the total displacement in the parking. We shall describe more precisely the asymptotic width of the largest blocks and where they are placed. We introduce the same notations and objects as in [3, 4]. Denote by F_m the empirical cumulative function associated to the r.v. $(X_k)_{1\leq k\leq m}$:

$$F_m(t) = \frac{\#\{\, i \,|\, X_i \leq t,\, 1 \leq i \leq m\}}{m}.$$

The knowledge of the function F_m is obviously sufficient to build the parking, since we have:

$$Y_k = m\big(F_m(k/n) - F_m((k-1)/n)\big).$$

Equality (2.6) can be reformulated as:

$$H_k = \max_{l\geq 0}\Big\{m\Big(F_m\big(\frac{k}{n}\big) - F_m\big(\overline{\frac{k-l-1}{n}}\big)\Big) - l\Big\}, \tag{2.7}$$

where the symbol $\overline{x}$ denotes the fractional part of x. Now, we use the strong convergence of F_m towards F. Let α_m be the function defined on $[0,1]$ by:

$$F_m(t) = F(t) + \frac{\alpha_m(F(t))}{\sqrt{m}}. \tag{2.8}$$

Theorem 2.7 *(Komlos, Major & Tusnady[12])*
Given X_1, X_2, .. i.i.d. r.v. on $[0,1]$ with cumulative function F, there exists a sequence $(br_n)_{n\geq 1}$ of normalized Brownian bridges such that for all n and x,

$$\mathbb{P}\Big(\sup_{0\leq t\leq 1}|\alpha_n(t) - br_n(t)| \geq \frac{A\log n + x}{\sqrt{n}}\Big) \leq Me^{-\mu x},$$

where A, M and μ are positive absolute constants.

In other terms, we can write:

$$F_m(t) = F(t) + \frac{br_n(F(t))}{\sqrt{m}} + \frac{c_n(t)}{m},$$

where $c_n(t)$ denotes $\sqrt{n}(\alpha_n(t) - br_n(t))$, and satisfies:

$$\mathbb{P}(\sup_{0\leq t\leq 1}|c_n(t)| \geq A\log n + x) \leq Me^{-\mu x}, \tag{2.9}$$

By (2.7), (2.8) and Theorem 2.7, one has:

$$H_k = \max_{l\geq 0}\Big\{m\int_{\overline{\frac{k-l-1}{n}}\curvearrowright\frac{k}{n}} f(x)dx + \sqrt{m}\Big(\alpha_m\big(F\big(\frac{k}{n}\big)\big) - \alpha_m\big(F\big(\overline{\frac{k-l-1}{n}}\big)\big)\Big) - l\Big\} \tag{2.10}$$

where

$$\int_{a\curvearrowright b} f(x)dx = \begin{cases} \displaystyle\int_a^b f(x)\,dx & \text{if } 0\leq a\leq b \\ \displaystyle\int_a^1 f(x)\,dx + \int_0^b f(x)\,dx & \text{if } 0\leq b\leq a \end{cases}$$

From (2.7) one has for any $k \in \{1, \cdots, n\}$:

$$\left| \frac{H_k}{m} - \max_{l \geq 0} \left\{ \int_{\frac{k-l-1}{n} \curvearrowright \frac{k}{n}} f(x)\, dx - \frac{l}{m} \right\} \right| \leq \frac{2\|\alpha_m\|_\infty}{\sqrt{m}}. \tag{2.11}$$

Theorem 2.8 *We define the (normalized) piecewise constant function h_n which interpolates the profile of the parking $(H_k)_{k=0,\cdots,n}$ by:*

$$h_n(t) = \frac{H_{\lfloor nt \rfloor}}{m(n)}.$$

One has:

$$h_n \xrightarrow[n \longrightarrow +\infty]{a.s.} h$$

where $(h(t))_{t\in[0,1]}$ is the deterministic function defined by:

$$h(t) = \max_{x \geq 0} \left\{ \int_{(t-x) \curvearrowright t} f(u)\ du - \frac{x}{\alpha} \right\}.$$

This convergence is almost sure for $\|.\|_\infty$ on $[0,1]$.

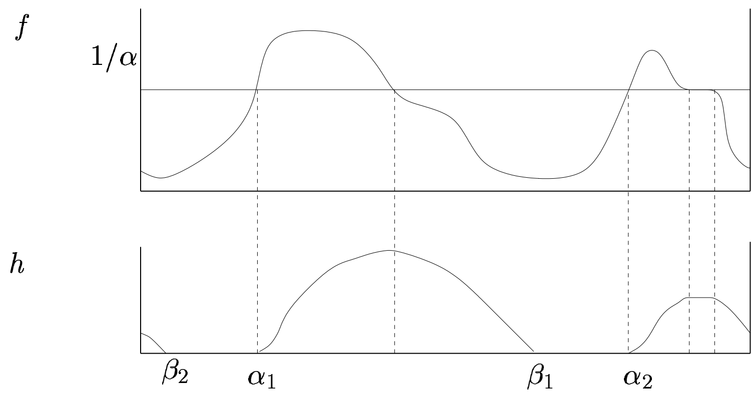

Figure 2.2 : Relation between density and limit profile

Proof : Using (2.11) and Theorem 2.7 we obtain:

$$\mathbb{P}\left(\left\|h_n - h\right\|_\infty \geq \varepsilon\right) \quad \leq \quad \mathbb{P}\left(\frac{2\|\alpha_m\|_\infty}{\sqrt{m}} \geq \varepsilon\right)$$

which is summable, since for any i:

$$\mathbb{P}(\|br_i\|_\infty \geq x) = e^{-2x^2}\mathbb{I}_{x>0},$$

and by (2.9).Theorem 2.8 follows from Borel-Cantelli Lemma. □

One deduces from this theorem the asymptotics of the total displacement in this parking:

Corollary 2.9 *(Total displacement)*
$TD(n, \alpha n, f)$, the total displacement in $Par(n, \alpha n, f)$ satisfies:

$$\frac{TD(n, \alpha\, n, f)}{\alpha n^2} \xrightarrow{a.s.} \int_0^1 h(t)dt.$$

Proof: We have

$$\frac{TD(n, \alpha\, n, f)}{\alpha n^2} = \frac{1}{n}\sum_{k=1}^{n} \frac{H_k}{\alpha n} = \frac{1}{n}\sum_{k=1}^{n} \frac{H_{\lfloor n\frac{k}{n}\rfloor}}{\alpha n} = \frac{1}{n}\sum_{k=1}^{n} h_n(k/n).$$

The results follows from Theorem 2.8 and Riemann sums.□

Corollary 2.10 *(The maximal number of visits on a given place)*

$$\frac{\max H_k}{m(n)} = \max_{1\le k\le n} h_n \xrightarrow{a.s.} \max h.$$

The most visited places are in the neighborhood of the place $p_{\lceil n\, argmax\, (h)\rceil}$.

Note that these places are not the places where f reaches its maximum.

Limit of the block sizes

The weak convergence in Theorem 2.8 does not imply the convergence of the intervals where h_n is positive to the intervals where h is positive. To give the limit size of the blocks, we shall impose some constraints on the density f. We identify each block of the parking by the two positions of the "bounding cars". Hence, in Figure (2.3), the set of blocks is

$$I(H) = \{[4, 7[, [11, 2[, [8, 9[\}.$$

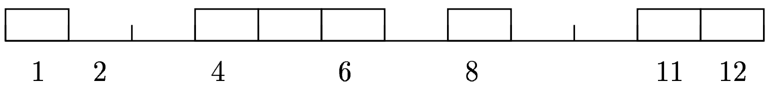

Figure 2.3 : Blocks of the parkings

We define the *blocks* of h with (also) a cyclical point of view: $[a, b]$ is a block of h if $h(a) = h(b) = 0$ and $h(x) > 0$ on $]a, b[$ if $a < b$ and on $]a, 1] \cup [0, b[$ if $b \le a$ (the blocks are called excursions in probability theory). The **width** $w([a, b])$ of a block $[a, b]$ is $b - a$ if $0 \le a < b$ and $1 - a + b$ else. We note that

$$I(h) = \big\{[a, b] \,\big|\, [a, b] \text{ is a block of } h\big\}.$$

In Figure (2.2), we have $I(h) = \big\{[\alpha_1, \beta_1], [\alpha_2, \beta_2]\big\}$. We denote by $I(h_n)$ (resp. $I^y(h_n)$) the normalized blocks of H (resp. with width larger than y):

$$\begin{aligned} I(h_n) &= \big\{[a/n, b/n] \text{ for } [a, b[\in I(H)\big\}, \\ I^y(h_n) &= \{[a/n, b/n] \text{ for } [a, b[\in I(H) \text{ such that } w(b-a) \ge y\}. \end{aligned}$$

In order to avoid problems related to the *critical case* (that is $h(x)=0$ and $f(x)=1/\alpha$) we set

$$Z=\{x\,|\,f(x)=1/\alpha\}$$

and

$$Z'=\{x\,|\,h(x)=0\},$$

the set of zeros of h. We denote by (Hyp) the set of hypothesis:

$$(Hyp):\ \begin{cases} f \text{ is continuous} \\ Z\cap Z' \text{ has no accumulation points}\end{cases}$$

For a given continuous function f (Hyp) holds for almost all α in [0,1].
Under Assumptions (Hyp), the "one car per place" rate ($f=1/\alpha$ and αn cars) happens only on isolated point (on $Z\cap Z'$) or when $h>0$; in this last case, the cars in excess coming from the left cancel the effects of this "one car per place" rate. The following theorem is not true without this kind of hypothesis (see the phenomenon described in Chassaing & Louchard: blocks of size of order n with $n-\lambda\sqrt{n}$ cars, for uniform parking).

Lemma 2.11 is used to prove that the "one car per place" rate (when not excess cars come from the left) holds only on isolated points of $[0,1]$.

Lemma 2.11 *Under the set of assumptions (Hyp), we have:*

$$Card(I(h))<+\infty.$$

Proof: If $[a_i,b_i]$ is a block of h then $h(a_i)=0$ and $f(a_i)=1/\alpha$ (f is continuous). So $a_i\in Z\cap Z'$ and since $Z\cap Z'$ has no accumulation points, $Card(Z\cap Z')<+\infty$. □
The result about the convergence of the blocks of h_n to the blocks of h is given by:

Theorem 2.12 *(Convergence of the blocks)*
Suppose that the blocks of h are disjoint (there are no two blocks $[a,b]$ and $[c,d]$ in $I(h)$ with $b=c$). Let η be any real positive number such that

$$\eta<\min\{w([a,b]),[a,b]\in I(h)\}. \tag{2.12}$$

Under (Hyp), we have:

$$\int_0^1 \left|\mathbb{I}_{I(h)}(x)-\mathbb{I}_{I^\eta(h_n)}(x)\right|dx=\int_0^1 \mathbb{I}_{I(h)\setminus I^\eta(h_n)}(x)+\mathbb{I}_{I^\eta(h_n)\setminus I(h)}(x)dx\ \xrightarrow[n]{a.s}\ 0.$$

The large blocks of $Par(n,m(n),f)$ (with size larger than ηn), suitably normalized, converge to the blocks of h. The small blocks of h_n with width smaller than η do not appear in the limit. It is easy to see that, roughly speaking, $(\alpha-\int_0^1\mathbb{I}_{I(h)}(x)dx)n$ cars in $Par(n,m(n),f)$ are in these small blocks.
The assumption that the blocks of h are disjoint is needed to prove the convergence of each large block of h_n to a corresponding block of h.
Proof of Theorem 2.12: Using the uniform convergence of h_n to h, we have

$$\int_0^1 \mathbb{I}_{I(h)\setminus I^\eta(h_n)}(x)\ \xrightarrow{a.s.}\ 0.$$

Let us study

$$\int_0^1 \mathbb{I}_{I^\eta(h_n)\backslash I(h)}(x)dx.$$

$I(h)$ and $I^\eta(h_n)$ contain at most $\lceil 1/\eta \rceil$ intervals. Consider the set $J = [0,1] - I(h)$. For an interval $[a,b]$ in J there exists a corresponding set of places on the parking, say:

$$p_{[\![a\,n, b\,n]\!]} = \{\dot{p}_{\lceil an \rceil}, \cdots, \dot{p}_{\lfloor bn \rfloor}\}.$$

We prove, in the following lemma, that each connex subset of the set of places $p_{[\![a\,n,b\,n]\!]}$ containing $\frac{\varepsilon}{2}n$ consecutive places contains at least an empty place (with a probability exponentially close to 1). It follows that at the limit, there does not exist any block with size larger than ϵn that are not expected.

Lemma 2.13 *Let ε such that $0 < \varepsilon < \eta$.*

$$\mathbb{P}(\exists [a,b] \in J | w(b-a) \geq \varepsilon/2 \textit{ and } p_{[\![a\,n,b\,n]\!]} \textit{ contains no empty place}) \leq \frac{6}{\varepsilon} \exp(-nl(\varepsilon))$$

where $l(\varepsilon) > 0$.

Hence, for sufficiently large n, there are no blocks of size larger than ε for h_n in the set $[0,1]\backslash I(h)$ (with probability tends to 1) and then in the set $I(h_n)\backslash I(h)$. Taking a small ε and since $I(h_n)\backslash I(h)$ is constituted by at most $2/\eta$ intervals, Theorem 2.12 appears to be a simple consequence of Lemma 2.13 and Borel-Cantelli Lemma.
Proof of Lemma 2.13: Consider K_ε a (fixed) set of disjoint intervals $(A_i = [a_i, b_i])_i$ - with width $\varepsilon/6$ - included in J and such that there are no interval of width larger than $\varepsilon/6$ in $J - (\cup_i A_i)$.

$$\begin{aligned} &\mathbb{P}(\exists [a_i,b_i] \in K_\varepsilon | p_{[\![a_i n, b_i n]\!]} \text{ contains no empty place}) \\ &\quad \leq \sum_{[a_i,b_i]\in K_\varepsilon} \mathbb{P}(p_{[\![a_i n, b_i n]\!]} \text{ contains no empty place}). \end{aligned} \tag{2.13}$$

The number of cars that choose one of the places $p_{[\![a_i\,n,b_i\,n]\!]}$ is Binomial $B(\alpha n, \int_{a_i}^{b_i} f(x)dx)$ distributed; since $a_i \in [0,1] - I(H)$, we have $H_{a_i n} = o(n)$. Since $f < 1/\alpha$ on $[a_i, b_i]$, by Cramér:

$$\mathbb{P}(p_{[\![a_i\,n,b_i\,n]\!]} \text{ contains no empty place}) \leq \exp(-nl_i(\varepsilon))$$

for a positive function l_i of ε. K_ε contains at most $\lfloor 6/\varepsilon \rfloor$ intervals and the summands in (2.13) goes to 0, for any ε, when n go to $+\infty$. $\square$

Conclusion

The study of the supercritical case relies on equation (2.10) which relates the profile of the parking to the random process α_m largely studied by mathematical statisticians. Taking $m = n$ and $f(x) = \mathbb{I}_{[0,1]}(x)$ in (2.10) yields formula

$$\frac{H_k}{\sqrt{n}} = \max_{l\geq 0} \Big\{ \alpha_n\big(\frac{k}{n}\big) - \alpha_n\big(\overline{\frac{k-l-1}{n}}\big) \Big\}$$

which is used in [4] to study the uniform parking. The Brownian excursion which is the limit of uniform parking profile appears in the present paper to be simply the second terms in the asymptotics of the profile of parkings with density.
Equation (2.10) allows also to easily study **linear parking with density on** $\mathbb{R}$. In linear parking with density f (defined on $\mathbb{R}$), car i parks on the first empty place among $p_{\lceil nX_i\rceil}, p_{\lceil nX_i\rceil+1}, \cdots$ (for i from 1 to νn, the number of cars). Then, the threshold phenomenon, depends on the values of ν. For example, taking $f(x) = \mathbb{I}_{[0,1]}(x)$, $\nu = 1$, we obtain the analogous of uniform parking, where the cars in excess on place p_n park on place $p_{n+1}, p_{n+2}, \cdots$ instead of parking on place $p_1, p_2, \cdots$. Using formula (2.10), one sees that the number of cars in excess on place n is about $-\sqrt{n}\min\{br(s)|0 \leq s \leq 1\}$. In terms of hashing with linear probing, this result means that if the table is linear and if the n data are hashed uniformally on the first n cells, the number of data that are at the end placed in the cells $n+1, n+2, \cdots$ are (roughly speaking) distributed as $-\sqrt{n}\min\{br(s)|0 \leq s \leq 1\} = o(n^{1/2+\epsilon})$.

Acknowledgments

I thank Philippe Chassaing who suggested this problem to me, L. Devroye, A. Mokkadem and J. Pian for useful comments.

References

[1] D. Aldous (1988). *Hashing with linear probing under nonuniform probabilities.* Prob. Eng. Inf. Sci. **2**, 1-14.

[2] P. Chassaing, S. Janson, (1999). *A Vervaat-like path transformation for the reflected Brownian bridge conditioned on its local time at 0.* Preprint 99/39, Institut Elie Cartan.

[3] P. Chassaing, G. Louchard, (1999) *Phase transition for parking blocks, Brownian excursion and coalescence.* Preliminary version.

[4] P. Chassaing, J.F. Marckert, (1999). *Parking functions, empirical processes and the width of rooted labeled trees.* To appear in Elec. Jour. Comb.

[5] A. Dembo, O. Zeitouni, (1998). *Large Deviations Techniques and Applications. 2nd ed..* [B] Applications of Mathematics. 38. New York: Springer.

[6] L. Devroye, (1985). *The expected length of the longest probe sequence for bucket searching when the distribution is not uniform.* J. Algorithms 6, 1-9.

[7] P. Flajolet, P. Poblete, A. Viola, (1998). *On the Analysis of Linear Probing Hashing.* Algoritmica Vol. **22** (4): 490-515.

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

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

[10] G.H. Gonnet, (1980) *Open-addressing hashing with unequal-probability keys.* J. Comput. Syst. Sci. 21, 354-367.

[11] D. E. Knuth, (1997). *The Art of Computer Programming. Vol. 3: Sorting and Searching.* 2nd ed., Addison-Wesley.

[12] J. Komlos, P. Major, G. Tusnady, (1976). *An approximation of partial sums of independent RV's and the sample DF. II.*, Z. Wahrs. verw. **34**, 33-58.

[13] B. Pittel, (1987). *Linear probing: the probable largest search time grows logarithmically with the number of records.* J. Algorithms, Vol. **8** (2): 236–249.

[14] G.CH. Pflug, H.W. Kessler, (1987). *Linear probing with a nonuniform address distribution* Journal of ACM, **34** (2), 397-410.