

## LE PAYSAGE DIRIGÉ

[d'après Duncan Dauvergne, Janosch Ortmann et Bálint Virág]

par Guillaume Barraquand

### 1. Introduction

Soit  $\mathcal{S}_n$  le groupe des permutations de  $\{1, \dots, n\}$  et  $\sigma \in \mathcal{S}_n$ . On dit que  $\sigma(i_1), \dots, \sigma(i_L)$  est une sous-suite croissante de  $\sigma$  si  $i_1 < \dots < i_L$  et  $\sigma(i_1) < \dots < \sigma(i_L)$ . Notons  $L_n(\sigma)$  l'entier maximal  $L$  pour lequel il existe une sous-suite croissante de longueur  $L$ . Lorsque  $\sigma$  est choisie au hasard uniformément dans  $\mathcal{S}_n$ , et que  $n$  tend vers l'infini, comment se comporte  $L_n(\sigma)$ , la longueur de la plus longue sous-suite croissante? ULAM (1961) fut le premier à poser cette question. En s'appuyant sur des simulations numériques, il conjectura qu'il existe une constante  $c$  telle que

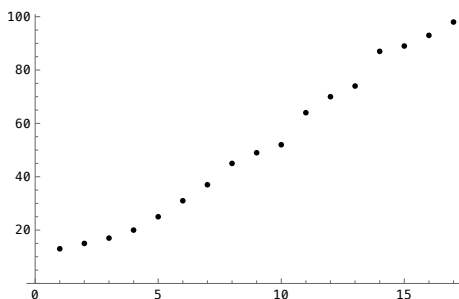
$$\frac{\mathbb{E}[L_n(\sigma)]}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{} c.$$

L'existence de la limite fut démontrée rigoureusement une dizaine d'années plus tard par HAMMERSLEY (1972). Quant à la valeur de la constante, on sait que  $c = 2$  depuis les travaux de LOGAN et SHEPP (1977) et de VERSHIK et KEROV (1977). La question de la loi des fluctuations de  $L_n(\sigma)$  autour de la moyenne a été résolue vingt ans plus tard par BAIK, DEIFT et JOHANSSON (1999). Les fluctuations sont d'ordre  $n^{1/6}$  et ne sont pas gaussiennes, mais

$$\frac{L_n(\sigma) - 2\sqrt{n}}{n^{1/6}}$$

converge vers la distribution de Tracy–Widom, introduite pour décrire les fluctuations de la plus grande valeur propre de matrices aléatoires hermitiennes (TRACY et WIDOM, 1994). La preuve de ce résultat, qui est particulièrement élégante, a débouché sur de nombreux développements mathématiques, constituant même un nouveau sous-domaine des probabilités : les *probabilités intégrables*.

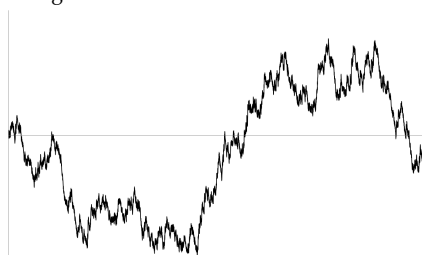
Au-delà de la plus longue sous-suite croissante, on peut aussi s'intéresser au comportement de la, ou des, plus longues sous-suites croissantes  $\sigma(i_1), \dots, \sigma(i_L)$ . Puisque



(A) Valeurs des indices  $i_1, \dots, i_k$  de la plus longue sous-suite croissante d'une permutation aléatoire uniforme de longueur  $n = 100$ .



(B) Le graphe de  $t \mapsto J_n(t)$  pour  $n = 10^7$ .



(C) Le graphe de  $t \mapsto J_n(t)$  pour  $n = 10^7$ .

**FIGURE 1** – Comparaison entre la géodésique du paysage dirigé et un pont brownien.

$\sigma$  et  $\sigma^{-1}$  ont la même loi, la plus longue sous-suite croissante suit la même loi de probabilité que la suite de ses indices  $i_1 < \dots < i_L$ . La figure 1(A) montre une simulation de cette suite d'entiers aléatoires pour  $n = 100$ . En reliant les points, on obtient une courbe aléatoire continue dont on voudrait connaître la loi lorsque  $n$  tend vers l'infini. En première approximation, les points semblent s'aligner le long d'une droite, mais après centrage et renormalisation, on obtient une trajectoire erratique dont la figure 1(B) montre une simulation.

À titre de comparaison, considérons un ensemble de  $2\sqrt{n}$  entiers  $j_1 < \dots < j_{2\sqrt{n}}$  choisis au hasard uniformément parmi les entiers de 1 à  $n$ . En définissant

$$J_n(t) = \frac{j_{\lfloor 2t\sqrt{n} \rfloor} - tn}{n^{3/4}}, \quad (1)$$

un résultat de DONSKER (1952) montre que  $J_n$  converge vers un pont brownien (c'est-à-dire un mouvement brownien défini sur l'intervalle  $[0, 1]$  démarrant en  $B(0) = 0$  et conditionné à terminer en  $B(1) = 0$ , voir figure 1(c)).

Revenons à la plus longue sous-suite croissante, ou plutôt la suite de ses indices  $i_1 < \dots < i_L$ . Ces entiers ne ressemblent pas du tout à  $L$  entiers tirés uniformément.

En effet, la manière correcte de renormaliser cette suite est différente de (1). En définissant

$$I_n(t) = \frac{i_{\lfloor 2t\sqrt{n} \rfloor} - tn}{2n^{5/6}}, \quad (2)$$

DAUVERGNE, ORTMANN et VIRÁG (2022) ont démontré que  $I_n$  converge vers une courbe fractale bien particulière, illustrée pour  $n = 10^7$  sur la figure 1(b). Fort différente d'un pont brownien, elle est définie comme la géodésique associée à une géométrie planaire aléatoire appelée le *paysage dirigé*. L'objectif de cet exposé est d'expliquer comment construire cet objet.

## 2. Contexte

Loin de concerner seulement les permutations aléatoires, le paysage dirigé décrit la limite d'échelle de nombreux modèles en physique statistique : percolation de premier ou dernier passage, modèles de croissance d'interface, systèmes de particules en interaction, etc.

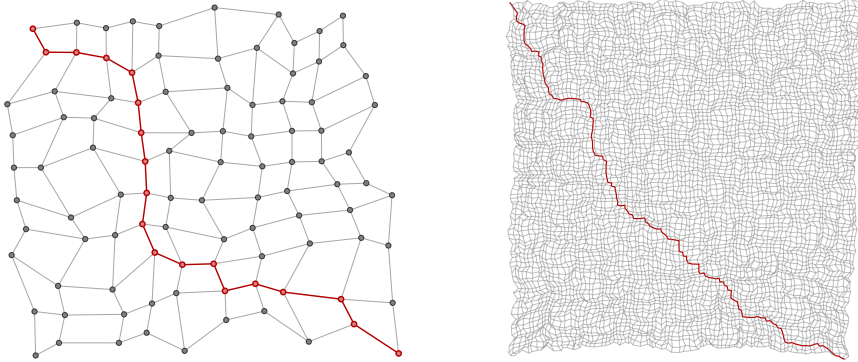
### 2.1. Géométrie planaire aléatoire

Le paysage dirigé décrit la limite d'échelle de distances aléatoires. Pour comprendre ce que cela signifie, considérons d'abord une perturbation aléatoire de la distance euclidienne sur  $\mathbb{Z}$ . Supposons par exemple que la distance entre les points  $i$  et  $i + 1$  est égale à  $w_i$ , où les  $w_i$  sont des variables aléatoires indépendantes prenant la valeur 1 ou 3 avec égale probabilité. Comme pour n'importe quel graphe, on définit la distance entre  $x$  et  $y$  comme le minimum sur les chemins de la somme des distances parcourues le long du chemin. Ainsi, pour  $x \leq y \in \mathbb{Z}$ ,  $d(x; y) = \sum_{i=x}^{y-1} w_i$ , et à grande échelle, lorsque  $n$  tend vers l'infini,

$$d(xn^2; yn^2) = 2n^2(y - x) + n(B(y) - B(x)) + o(n), \quad (3)$$

où  $B$  est un mouvement brownien standard. À quelques constantes près, le résultat reste vrai quelle que soit la loi des variables  $w_i$ , pourvu qu'elles aient une variance finie. Le mouvement brownien est donc la limite d'échelle universelle de distances aléatoires en dimension 1. La notion d'« universalité », empruntée à la physique statistique, signifie que la limite d'échelle ne dépend pas des détails du modèle.

Le paysage dirigé est la limite d'échelle universelle de géométries aléatoires en dimension 2 définies à partir de variables indépendantes. Contrairement au cas de la dimension 1, l'universalité en dimension 2 est seulement conjecturée. Plus précisément, considérons des variables  $w_e$  de même loi que précédemment, mais maintenant



**FIGURE 2** – Gauche : le graphe  $\mathbb{Z}^2$  muni de distances aléatoires sur les arêtes et une géodésique entre deux points. Droite : Idem, à plus grande échelle.

indexées par les arêtes de  $\mathbb{Z}^2$ . La distance est définie comme précédemment :

$$d(p; q) = \min_{\gamma: p \leftrightarrow q} \left\{ \sum_{e \in \gamma} w_e \right\}, \quad (4)$$

où le minimum est pris sur les chemins dans  $\mathbb{Z}^2$  reliant  $p$  à  $q$ . La figure 2 est une illustration de cette géométrie.

Comment renormaliser cette distance aléatoire afin d'obtenir un objet universel ? Puisqu'on s'intéresse au comportement de géodésiques dans une direction, fixons un vecteur  $\vec{v}$  et posons  $p = n^{1/\chi} s \vec{v}$  et  $q = n^{1/\chi} t \vec{v}$  où  $s, t \in \mathbb{R}$  et  $\chi$  est un exposant à déterminer de sorte que les fluctuations de  $d(p, q)$  soient d'ordre  $n$ , comme dans (3). L'analyse de certains modèles (BAIK, DEIFT et JOHANSSON, 1999; JOHANSSON, 2000, 2001a) montre que l'exposant correct est  $\chi = 1/3$ . Mais il ne suffit pas de caractériser les distances le long d'une direction, on voudrait aussi comprendre les fluctuations des distances lorsqu'on fait varier les points  $p$  et  $q$  dans la direction orthogonale à  $\vec{v}$ . Posons

$$p = n^{1/\chi} s \vec{v} + n^{\xi/\chi} x \vec{u} \text{ et } q = n^{1/\chi} t \vec{v} + n^{\xi/\chi} y \vec{u},$$

où le vecteur  $\vec{u}$  est orthogonal à  $\vec{v}$  et  $\xi$  est un autre exposant à choisir afin que les distances  $d(p; q)$  aient des corrélations non-triviales lorsque  $x$  et  $y$  varient.

En géométrie euclidienne, un chemin entre deux points à distance  $n^\alpha$  dont la longueur est  $n^\alpha + n$  dévie par rapport à une ligne droite d'une distance  $O(n^\beta)$  avec  $\beta = \frac{\alpha+1}{2}$ . Cela suggère que les exposants  $\chi$  et  $\xi$  vérifient la relation  $\chi = 2\xi - 1$ , appelée relation KPZ (prouvée par CHATTERJEE (2013) sous certaines hypothèses).

Ainsi,  $\chi = 1/3$ ,  $\xi = 2/3$ , et on s'attend à ce que pour toute direction  $\vec{v}_o$ , il existe des vecteurs  $\vec{v} \parallel \vec{v}_o$ ,  $\vec{u} \perp \vec{v}_o$  et une constante  $c$  telle que lorsque  $n \rightarrow \infty$ ,

$$d\left(n^3 s \vec{v} + n^2 x \vec{u}; n^3 t \vec{v} + n^2 y \vec{u}\right) = (t - s)n^3 c - n \mathcal{L}(x, s; y, t) + o(n). \quad (5)$$

En outre, on s'attend à ce que le processus  $\mathcal{L}(x, s; y, t)$ , le paysage dirigé, soit universel : non seulement  $\mathcal{L}$  ne doit pas dépendre de la loi des variables  $w_e$ , mais on pourrait aussi remplacer  $\mathbb{Z}^2$  par un autre graphe planaire. Comme pour le théorème central limite, qui requiert seulement que les variables aléatoires soient indépendantes et qu'elles aient une variance finie, l'universalité est attendue ici sous des hypothèses générales. On sait qu'il faut notamment que le cinquième moment de la loi de  $w_e$  existe :  $\mathbb{E}[w_e^5] < \infty$  (voir l'étude numérique de GUEUDRÉ, LE DOUSSAL, BOUCHAUD et Rosso (2015)). Aujourd'hui, la convergence (5) n'est établie que pour certains modèles particuliers.

Comme le mouvement brownien, le paysage dirigé obéit à une propriété d'invariance d'échelle : pour tout  $\sigma > 0$ ,

$$\mathcal{L}(\sigma^2 x, \sigma^3 s; \sigma^2 y, \sigma^3 t) \stackrel{(\text{loi})}{=} \sigma \mathcal{L}(x, s; y, t).$$

La présence du signe  $-$  dans (5) peut paraître surprenante. C'est une simple convention qui vient du fait que le paysage dirigé a été à l'origine introduit comme limite d'un modèle défini à partir d'un maximum au lieu d'un minimum comme dans (4). Le paysage dirigé est « dirigé » au sens où il correspond à la limite des distances le long d'une direction  $v_\circ$ . De plus, dans la suite de ce texte,  $\mathcal{L}(x, s; y, t)$  sera défini de telle sorte que (5) soit vraie pour  $s \leq t$ , et lorsque  $s > t$ , on posera  $\mathcal{L}(x, s; y, t) = -\infty$ . Cette brisure de symétrie dans la définition permet à  $-\mathcal{L}$  de vérifier l'inégalité triangulaire.

## 2.2. La classe d'universalité de Kardar–Parisi–Zhang

De nombreuses situations physiques décrites par la croissance d'une interface, par exemple dans des phénomènes de propagation (du feu, d'une épidémie, ...), ou des processus d'aggrégation de matière (par exemple la chute de neige sur une surface), obéissent à des lois similaires. Ce surprenant phénomène d'universalité a été étudié par KARDAR, PARISI et ZHANG (1986), qui ont calculé les exposants caractéristiques décrivant les corrélations des fluctuations en espace et en temps, à partir de l'étude d'un modèle jouet : une équation aux dérivées partielles stochastique aujourd'hui appelé équation KPZ. L'interface est modélisée par une fonction  $h(x, t)$ , dépendant du temps  $t$  et d'une position en espace  $x$ , qui obéit à

$$\partial_t h(x, t) = \frac{1}{2} \partial_{xx} h(x, t) + \frac{1}{2} (\partial_x h(x, t))^2 + \beta \zeta(x, t) \quad (6)$$

où  $\zeta(x, t)$  désigne un bruit blanc espace-temps. Nous ne définirons pas précisément le bruit blanc ici, mais rappelons seulement qu'il s'agit d'un terme de bruit aléatoire, décorréolé en espace et en temps. De plus, nous nous restreignons dans ce texte à un espace de dimension 1, de sorte que l'espace-temps est de dimension 2.