Hammersley's Path Process

Steven Finch

April 5, 2004

The following is a generalization of a process introduced by Hammersley [1, 2]. Fix three parameters $\lambda > 0$, $\alpha_+ \geq 0$ and $\alpha_- \geq 0$. Let $P(\mu)$ denote a Poisson random variable with mean μ. Baik & Rains [3] constructed a set of points S in the unit square $[0,1] \times [0,1]$ according to three rules:

- $P(\lambda^2)$ points are selected uniformly inside $(0,1) \times (0,1)$
- $P(\alpha_+\lambda)$ points are selected uniformly on the open bottom edge $(0,1) \times \{0\}$
- $P(\alpha_-\lambda)$ points are selected uniformly on the open left edge $\{0\} \times (0,1)$.

These rules are independently executed. No points are selected from the closed top and right edges, nor is the origin $(0,0)$ allowed.

Consider any sequence of distinct points of the form

$$(0,0),\ (s_1,t_1),\ (s_2,t_2),\ \ldots\ ,\ (s_n,t_n),\ (1,1)$$

where each $(s_k,t_k) \in S$, $1 \leq k \leq n$, and n is arbitrary. For convenience, define $(s_0,t_0) = (0,0)$ and $(s_{n+1},t_{n+1}) = (1,1)$. Define such a point sequence to be an **up/right path** if, for any $k \geq 1$, we have $s_{k-1} \leq s_k$ and $t_{k-1} \leq t_k$. Hence an up/right path joins points of S in a continuous, piecewise linear manner with line segments of slope m_k, $0 \leq m_k \leq \infty$, attaching (s_{k-1},t_{k-1}) and (s_k,t_k) for all k.

Of all up/right paths determined by S, there is (at least) one with a maximum number n of points. Call this number N_λ. (This is usually referred to as a *length* in the literature. Of course, it also depends implicitly on α_+ and α_-.) What can be said about the probability distribution of N_λ as $\lambda \to \infty$?

A special case of the above is the longest increasing subsequence problem [4], achieved when $\alpha_+ = \alpha_- = 0$. Its solution will be folded into the formulas we give shortly for the general problem. This turns out to be related to the polynuclear growth (PNG) model in physics due to Prähofer & Spohn [5, 6, 7], but we cannot discuss such topics now.

When $0 \le \alpha_+ \le 1$ and $0 \le \alpha_- \le 1$ are fixed, the following formulas hold [3]:

$$\lim_{\lambda\to\infty} \mathrm{P}\left(\frac{N_\lambda - 2\lambda}{\lambda^{1/2}} \le x\right) = \begin{cases} F_{\mathrm{GUE}}(x) & \text{if } \alpha_+ < 1 \text{ and } \alpha_- < 1, \\ F_{\mathrm{GOE}}(x)^2 & \text{if } \alpha_+ = 1,\ \alpha_- < 1 \text{ or } \alpha_+ < 1,\ \alpha_- = 1, \\ F_0(x) & \text{if } \alpha_+ = 1 \text{ and } \alpha_- = 1, \end{cases}$$

where the distribution functions $F_{\mathrm{GUE}}(x)$, $F_{\mathrm{GOE}}(x)$ and $F_0(x)$ will be defined shortly. Also, when $\alpha_+ > 1$ or $\alpha_- > 1$, we have

$$\lim_{\lambda\to\infty} \mathrm{P}\left(\frac{N_\lambda - (\alpha+\alpha^{-1})\lambda}{\sqrt{\alpha-\alpha^{-1}}\lambda^{1/2}} \le x\right) = \begin{cases} \Phi(x) & \text{if } \alpha_+ \ne \alpha_-, \\ \Phi(x)^2 & \text{if } \alpha_+ = \alpha_-, \end{cases}$$

where $\alpha = \max\{\alpha_+, \alpha_-\}$ and $\Phi(x)$ is the standard normal distribution function:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{r^2}{2}\right) dr = \frac{1}{2}\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right) + \frac{1}{2}.$$

We provide moments corresponding to these distributions (and more) in Tables 1 and 2; computations were performed by Prähofer [8]. The functions $F_{\mathrm{GUE}}(x)$, $F_{\mathrm{GOE}}(x)$ and $F_{\mathrm{GSE}}(x)$ were first discovered by Tracy & Widom [9, 10, 11], whereas $F_0(x)$ arose more recently [3]. See Figure 1 for the associated density plots.

Table 1. *Moments of GUE, GOE, GUE^2 and GOE^2*

	F_{GUE}
mean	$-1.7710868074...$
variance	$0.8131947928... = (0.9017731382...)^2$
skewness	$0.2240842036...$
kurtosis	$0.0934480876...$
	F_{GOE}
mean	$-1.2065335745... = 2^{2/3}(-0.7600685240...)$
variance	$1.6077810345... = (1.2679830576...)^2 = 2^{4/3}(0.6380483264...)$
skewness	$0.2934645240...$
kurtosis	$0.1652429384...$
	F_{GUE}^2
mean	$-1.2633181526...$
variance	$0.6066887541... = (0.7789022750...)^2$
skewness	$0.3290093382...$
kurtosis	$0.2254319482...$
	F_{GOE}^2
mean	$-0.4936399332...$
variance	$1.2320144032... = (1.1099614422...)^2$
skewness	$0.3917246784...$
kurtosis	$0.3086329720...$

Table 2. *Moments of GSE and Other Distributions*

	F_{GSE}	
mean	$-2.3068848932... = \frac{1}{\sqrt{2}}(-3.2624279028...)$	
variance	$0.5177237207... = (0.7195302083...)^2 = \frac{1}{2}(1.0354474415...) = \frac{1}{2}(1.0175693792...)^2$	
skewness	0.1655094943	
kurtosis	0.0491951565	
	F_0	
mean	0	
variance	$1.1503944782... = 2^{2/3}(0.7247031094...) = (0.8104567006...)^{-2/3}$	
skewness	$0.3594116897...$	
kurtosis	$0.2891570248...$	
	Φ	Φ^2
mean	0	$\frac{1}{\sqrt{\pi}} = 0.5641895835...$
variance	1	$1 - \frac{1}{\pi} = 0.6816901138... = (0.8256452711...)^2$
skewness	0	$\frac{4-\pi}{2(\pi-1)^{3/2}} = 0.1369487673...$
kurtosis	0	$\frac{2(\pi-3)}{(\pi-1)^2} = 0.0617443154...$

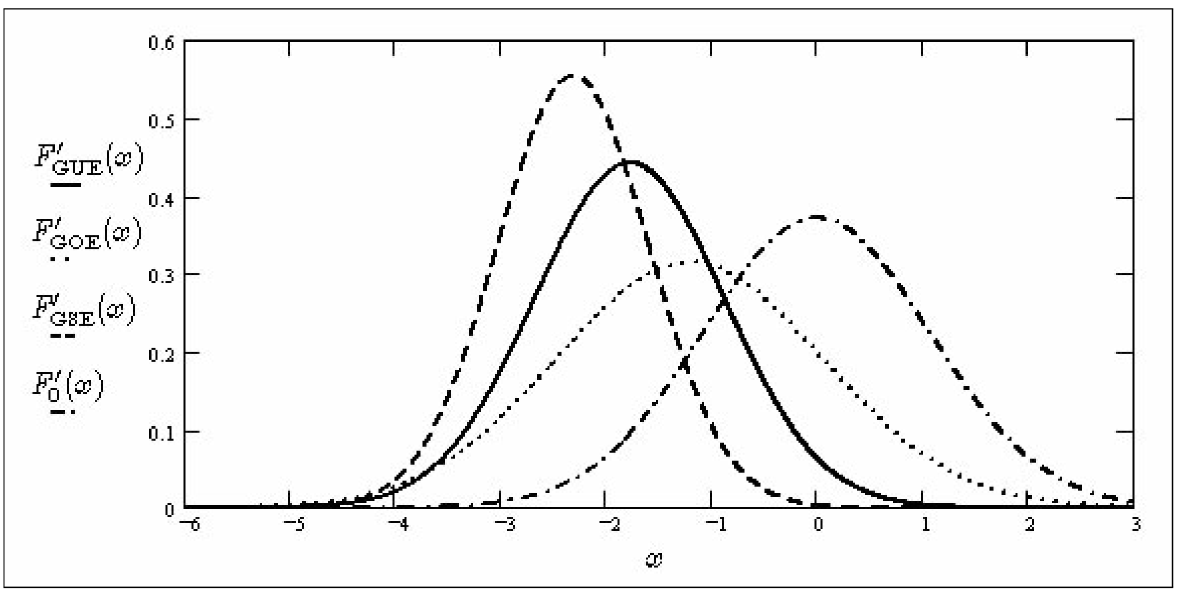

Figure 1: The Tracy-Widom density functions, as well as $F_0'(x)$.

Let $u(x)$ be the solution of the Painlevé II differential equation:

$$u''(x) = 2u(x)^3 + xu(x), \qquad u(x) \sim -\frac{1}{2\sqrt{\pi}} x^{-1/4} \exp\left(-\frac{2}{3}x^{3/2}\right) \quad \text{as } x \to \infty,$$

and define

$$U(x) = -\int_x^\infty u(r)\,dr, \qquad V(x) = -\int_x^\infty v(r)\,dr$$

where

$$v(x) = -\int_x^\infty u(r)^2\,dr.$$

The largest eigenvalue of a random complex Hermitian matrix, when generated according to the Gaussian Unitary Ensemble (GUE) probability law and properly normalized, has distribution function

$$F_{\mathrm{GUE}}(x) = \exp(-V(x)) \qquad \text{(often denoted as the case } \beta = 2\text{)}.$$

More details appear in [0.1]. Replacing Hermitian matrices by real symmetric matrices, we obtain the Gaussian Orthogonal Ensemble (GOE) and corresponding distribution function

$$F_{\mathrm{GOE}}(x) = \exp\left(-\frac{U(x)+V(x)}{2}\right) \qquad \text{(often denoted as the case } \beta = 1\text{)}.$$

Likewise, for the Gaussian Symplectic Ensemble (GSE), we have

$$F_{\mathrm{GSE}}(x) = \cosh\left(\frac{U(x)}{2}\right)\exp\left(-\frac{V(x)}{2}\right) \qquad \text{(the case } \beta = 4\text{)}.$$

Define also

$$F_0(x) = \left[1 - \left(x + 2u'(x) + 2u(x)^2\right) v(x)\right] \exp\left(-2U(x) - V(x)\right),$$

which does not yet seem to possess a random matrix interpretation. These formulas serve as the computational basis for Tables 1 and 2, where skewness and kurtotis of a random variable Y are given as

$$\mathrm{Skew}(Y) = \frac{\mathrm{E}\left[(Y - \mathrm{E}(Y))^3\right]}{\mathrm{Var}(Y)^{3/2}}, \qquad \mathrm{Kurt}(Y) = \frac{\mathrm{E}\left[(Y - \mathrm{E}(Y))^4\right]}{\mathrm{Var}(Y)^2} - 3.$$

For example, if $\alpha_+ < 1$ and $\alpha_- < 1$, it follows that

$$\lim_{\lambda\to\infty} \lambda^{-1/3}(\mathrm{E}(N_\lambda) - 2\lambda) = -1.7710868074..., \qquad \lim_{\lambda\to\infty} \lambda^{-2/3}\,\mathrm{Var}(N_\lambda) = 0.8131947928...$$

which generalize results given earlier by Tracy & Widom and Baik, Deift & Johansson [4]. If instead $\alpha_+ = 1$ and $\alpha_- = 1$, we have

$$\lim_{\lambda\to\infty} \lambda^{-1/3}\,\mathrm{Var}(N_\lambda) = 1.1503944782...$$

which is called the Baik-Rains constant in [7].

0.1. GUE/GOE/GSE. A random complex Hermitian $N \times N$ matrix X belongs to GUE if its (real) diagonal elements x_{jj} and (complex) upper triangular elements $x_{jk} = \xi_{jk} + i\eta_{jk}$ are independently chosen from zero-mean Gaussian distributions with $\mathrm{Var}(x_{jj}) = 2$ for $1 \le j \le N$ and $\mathrm{Var}(\xi_{jk}) = \mathrm{Var}(\eta_{jk}) = 1$ for $1 \le j < k \le N$. Let λ denote the largest (real) eigenvalue of X and define the normalization [11]

$$\tilde{\lambda} = \frac{N^{1/6}(\lambda - 2\sigma\sqrt{N})}{\sigma}$$

where $\sigma = \sqrt{\mathrm{Var}(x_{jk})} = \sqrt{2}$. Then the distribution of $\tilde{\lambda}$ has the moments indicated for GUE in Table 1. A related discussion, involving spacings between adjacent eigenvalues of X and featuring connections to the Riemann hypothesis, appears in [12].

A random real symmetric $N \times N$ matrix X belongs to GOE if its diagonal elements x_{jj} and upper triangular elements x_{jk} are independently chosen from zero-mean Gaussian distributions with $\mathrm{Var}(x_{jj}) = 2$ and $\mathrm{Var}(x_{jk}) = 1$. Let $\tilde{\lambda}$ denote the largest (real) eigenvalue of X, normalized as before with $\sigma = 1$ in this case. Then the distribution of $\tilde{\lambda}$ has the moments indicated for GOE in Table 1.

A complex Hermitian $2N \times 2N$ matrix is said to be **real quaternionic** [13] if, when viewed as an $N \times N$ matrix X consisting of 2×2 blocks, the diagonal blocks X_{jj} look like

$$X_{jj} = \begin{pmatrix} x_{jj} & 0 \\ 0 & x_{jj} \end{pmatrix}, \qquad x_{jj} \in \mathbb{R}$$

and the upper triangular blocks X_{jk} look like

$$X_{jk} = \begin{pmatrix} \xi_{jk} + i\eta_{jk} & \xi'_{jk} + i\eta'_{jk} \\ -\xi'_{jk} + i\eta'_{jk} & \xi_{jk} - i\eta_{jk} \end{pmatrix}.$$

A random real quaternionic matrix X belongs to GSE if the nonzero distinct elements of its diagonal and upper triangular blocks are independently chosen from zero-mean Gaussian distributions with $\mathrm{Var}(x_{jj}) = 2$ and $\mathrm{Var}(\xi_{jk}) = \mathrm{Var}(\eta_{jk}) = \mathrm{Var}(\xi'_{jk}) = \mathrm{Var}(\eta'_{jk}) = 1$. Let $\tilde{\lambda}$ denote the largest (real) eigenvalue of X, normalized as before with $\sigma = 2$ in this case. Then the distribution of $\tilde{\lambda}$ has the moments indicated for GSE in Table 2.

Here is an occurrence of $F_{\mathrm{GUE}}(x)^2$: Define a **signed permutation** π to be a bijection from $\{-n, -n+1, \ldots, -2, -1, 1, 2, \ldots, n-1, n\}$ onto itself which satisfies $\pi(-k) = -\pi(k)$ for all k. Tracy & Widom [14, 15] proved that the length L_{2n} of the longest increasing subsequence of a random signed permutation π satisfies

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{L_{2n} - 2\sqrt{2n}}{2^{2/3}(2n)^{1/6}} \le x \right) = F_{\mathrm{GUE}}(x)^2.$$

A nice combinatorial application involving $F_{\mathrm{GSE}}(x)$ or $F_{\mathrm{GSE}}(x)^2$, especially one as simple as this, would be good to find.

Other applications appear in [11, 16, 17]. A d-dimensional analog of Hammersley's original process (with $\alpha_+ = \alpha_- = 0$) appears in [18]: Let S denote a set of $P(\lambda^d)$ points selected uniformly inside the d-dimensional unit cube and N_λ denote the number of points in a maximal chain (totally ordered subset) of S. Define c_d to be $\limsup_{\lambda\to\infty} \mathrm{E}(N_\lambda)/\lambda$. Then it is known that $c_2 = 2$ and $c_\infty = e$, but $2.363 \le c_3 \le 2.366$, $2.514 \le c_4 \le 2.521$, $2.583 \le c_5 \le 2.589$ and $2.607 \le c_6 \le 2.617$. We draw attention finally to the obvious identity [2]:

$$F_{\mathrm{GSE}}(x) = \frac{1}{2}\left(F_{\mathrm{GOE}}(x) + \frac{F_{\mathrm{GUE}}(x)}{F_{\mathrm{GOE}}(x)}\right)$$

and wonder whether a similar identity relating F_0 to other distributions can ever be found.

0.2. Positive Definite/Indefinite. Among many possible questions, we ask for the probability that a random $N\times N$ matrix, distributed according to GOE, is positive definite. Since

$$\begin{aligned} \mathrm{P}\,(\text{indefinite}) &= 1 - \mathrm{P}\,(\text{positive definite}) - \mathrm{P}\,(\text{negative definite}) \\ &= 1 - 2\,\mathrm{P}\,(\text{positive definite}), \end{aligned}$$

the answer for indefinite matrices is clear once it is found for positive definite matrices. The joint density for the N unordered (real) eigenvalues of a GOE matrix is [19]

$$\frac{1}{C_N}\prod_{1\le i<j\le N} |\lambda_i - \lambda_j| \cdot \exp\left(-\frac{1}{4}\sum_{k=1}^{N}\lambda_k^2\right)$$

where

$$C_N = N!(2\pi)^{N/2}2^{N(N+1)/4}\prod_{\ell=1}^{N}\frac{\Gamma(\ell/2)}{\Gamma(1/2)}.$$

A complicated formula for the density of the smallest eigenvalue follows, as do the results in Table 3 for small N.

Table 3. *Probabilities that an $N \times N$ GOE Matrix is Positive Definite/Indefinite*

N	positive definite	indefinite
1	$1/2 = 0.5$	0
2	$1/2 - \sqrt{2}/4 \approx 0.1464$	$\sqrt{2}/2 \approx 0.7071$
3	$1/4 - \left(\sqrt{2}/2\right)\pi^{-1} \approx 0.0249$	$1/2 + \sqrt{2}\pi^{-1} \approx 0.9502$
4	$1/4 - \sqrt{2}/16 - (1/2)\pi^{-1} \approx 0.0025$	$1/2 + \sqrt{2}/8 + \pi^{-1} \approx 0.9951$
5	$1/8 - \left(1/3 + \sqrt{2}/24\right)\pi^{-1} \approx 0.0001$	$3/4 + \left(2/3 + \sqrt{2}/12\right)\pi^{-1} \approx 0.9997$

Consider now the quadratic form

$$Q(x_1, x_2, \ldots, x_N) = \sum_{1 \le i \le j \le N} m_{ij} x_i x_j$$

where the coefficients m_{ij} form the upper triangular portion of a GOE matrix M. Another way of saying M is indefinite is that $Q = 0$ possesses a nonzero solution in $\mathbb{R}^N$. If we constrain the m_{ij} to be integers, what is the probability that $Q = 0$ possesses a nonzero solution in $\mathbb{Z}^N$? The answer is 0 for $1 \le N \le 3$, is the same as the real indefinite case for $N \ge 5$, but is miraculously [20, 21]

$$\left(\frac{1}{2} + \frac{\sqrt{2}}{8} + \frac{1}{\pi}\right) \prod_p \left(1 - \frac{p^3}{4(p+1)^2\,(p^4 + p^3 + p^2 + p + 1)}\right) = 0.9825845607...$$

for $N = 4$. If we replace the GOE distribution by, say, a uniform distribution on $[-1/2, 1/2]$ for each m_{ij}, then the probability becomes 0.97... instead. The structure of the formula – leading coefficient multipled by prime product – is similar. While the prime product 0.9874362482... remains identical, no closed-form expression is known for the leading coefficient.

REFERENCES

[1] J. M. Hammersley, A few seedlings of research, *Proc. Sixth Berkeley Symp. Math. Stat. Probab.*, v. 1, ed. L. M. Le Cam, J. Neyman and E. L. Scott, Univ. of Calif. Press, 1972, pp. 345-394; MR0405665 (53 #9457).

[2] P. J. Forrester, Growth models, random matrices and Painlevé transcendents, *Nonlinearity* 16 (2003) R27–R49; math-ph/0309058; MR2012845 (2004k:82071).

[3] J. Baik and E. M. Rains, Limiting distributions for a polynuclear growth model with external sources, *J. Stat. Phys.* 100 (2000) 523–541; math.PR/0003130; MR1788477 (2001h:82067).

[4] S. R. Finch, Longest subsequence constants, *Mathematical Constants*, Cambridge Univ. Press, 2003, pp. 382–387.

[5] M. Prähofer and H. Spohn, Statistical self-similarity of one-dimensional growth processes, *Physica A* 279 (2000) 342–352; cond-mat/9910273; MR1797145 (2001j:82098).

[6] M. Prähofer and H. Spohn, Universal distributions for growth processes in $1+1$ dimensions and random matrices, *Phys. Rev. Lett.* 84 (2000) 4882–4885; cond-mat/9912264.

[7] M. Prähofer and H. Spohn, Exact scaling functions for one-dimensional stationary KPZ growth, *J. Stat. Phys.* 115 (2004) 255-279; cond-mat/0212519; numerical tables at http://www-m5.ma.tum.de/KPZ/; MR2070096 (2005d:82097).

[8] M. Prähofer, High precision moments corresponding to GUE/GOE/GSE, unpublished note (2004).

[9] C. A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, *Comm. Math. Phys.* 159 (1994) 151-174; hep-th/9211141; MR1257246 (95e:82003).

[10] C. A. Tracy and H. Widom, On orthogonal and symplectic matrix ensembles, *Comm. Math. Phys.* 177 (1996) 727–754; solv-int/9509007; MR1385083 (97a:82055).

[11] C. A. Tracy and H. Widom, Distribution functions for largest eigenvalues and their applications, *International Congress of Mathematicians*, v. 1, Proc. 2002 Beijing conf., Higher Ed. Press, pp. 587–596; math-ph/0210034; MR1989209 (2004f:82034).

[12] S. R. Finch, Glaisher-Kinkelin constant: GUE hypothesis, *Mathematical Constants*, Cambridge Univ. Press, 2003, pp. 138–145.

[13] P. J. Forrester, *Log-Gases and Random Matrices*, Princeton University Press, 2010; MR2641363.

[14] C. A. Tracy and H. Widom, Random unitary matrices, permutations and Painlevé, *Comm. Math. Phys.* 207 (1999) 665–685; MR1727236 (2001h:15019).

[15] A. Borodin, Longest increasing subsequences of random colored permutations, *Elec. J. Combin.* 6 (1999) R13; MR1667453 (2000a:05014).

[16] J. Baik and E. M. Rains, Symmetrized random permutations, *Random Matrix Models and their Applications*, ed. P. Bleher and A. Its, Cambridge Univ. Press, 2001, pp. 1–19; math.CO/9910019; http://www.msri.org/publications/books/Book40/contents.html; MR1842780 (2002i:82038).

[17] P. L. Ferrari, Random matrices and determinantal processes, *Note di Matematica e Fisica (CERFIM)* 12 (2003) 67–83; http://wt.iam.uni-bonn.de/ferrari/publications/.

[18] E. Breimer, M. Goldberg, B. Kolstad and M. Magdon-Ismail, On the height of a random set of points in a d-dimensional unit cube, *Experim. Math.* 10 (2001) 583–597; *Algorithm Engineering and Experimentation (ALENEX)*, Proc. 2001 Washington DC conf., ed. A. L. Buchsbaum and J. Snoeyink, Lect. Notes in Comp. Sci. 2153, Springer-Verlag, 2001, pp. 161–171; MR1881759 (2002j:60013).

[19] G. W. Anderson, A. Guionnet and O. Zeitouni, *An Introduction to Random Matrices*, Cambridge Univ. Press, 2010, pp. 51–53; MR2760897 (2011m:60016).

[20] M. Bhargava, J. E. Cremona, T. A. Fisher, N. G. Jones and J. P. Keating, What is the probability that a random integral quadratic form in n variables has an integral zero?, *Internat. Math. Research Notices* (2016); arXiv:1502.05992.

[21] P. Sebah, Another infinite product over primes, unpublished note (2015).