Extreme value statistics in models with logarithmic correlations

Yan V Fyodorov[1]

School of Mathematical Sciences
The University of Nottingham, UK

Random Matrices and Integrability, Yad Hashmona, March 25-30, 2009

[1]**collaborators:** **Jean-Philippe Bouchaud**,**Pierre Le Doussal**, **Alberto Rosso** (Paris)

Informal Basics of Extreme Value Theory: Let $\mathbf{X_1, X_2, ..., X_M}$ be i.i.d. real random variables with a (continuous) probability density $\mathbf{P(X)}$ and cumulative distribution $\mathbf{F(X)} = \int_{-\infty}^{\mathbf{X}} \mathbf{P(Y)dY}$. When the moments $\langle|\mathbf{X}|^{\alpha}\rangle$ are finite for all $\mathbf{0} \le \alpha < \mathbf{1}/\gamma$ and infinite for $\alpha > \mathbf{1}/\gamma$ one can find such M-dependent constants $\mathbf{a_M, b_M}$ that the variable $\mathbf{x} = \frac{\max\{\mathbf{X_1, X_2, ..., X_M}\} - \mathbf{b_M}}{\mathbf{a_M}}$ has a well-defined distribution in the limit $\mathbf{M} \to \infty$, that is

$$\lim_{\mathbf{M}\to\infty} \mathbf{F^M(a_M\, x + b_M) = G(x)} = \exp -(\mathbf{1} + \gamma\mathbf{x})^{-\mathbf{1}/\gamma}$$

In particular, if all positive moments exist then $\gamma = \mathbf{0}$ and $\mathbf{G(x)} = \exp\{-\mathbf{e^x}\}$ which is known as the **Gumbel** distribution.

Gumbel law is rather robust if variables are short-range correlated. In particular, for Gaussian-distributed variables with $<\mathbf{X_i}>= \mathbf{0}$ the **Gumbel** distribution is known to be valid as long as $\mathbf{C(|i-j|)} = \langle\mathbf{X_i X_j}\rangle \gtrsim \mathbf{const}/\ln|\mathbf{i-j}|$.

Little is known about statistics of extremes of strongly correlated variables.

Notable exceptions are, e.g. **Tracy-Widom** distribution for min/max eigenvalues of random matrices and (ii) **"Airy law"** for distribution of maximal relative height in 1d random walks $\&$ solid-on-solid models (**S Majumdar et al**).

Problem:

- Given an instance of the $2D$ **Gaussian free field**:

$$\mathcal{P}\left[V(\mathbf{x})\right] \propto \exp -\frac{1}{g^2}\int \left[\nabla V(\mathbf{x})\right]^2 \, d^2\mathbf{x}$$

characterized by the covariance

$$\langle V(\mathbf{x_1})V(\mathbf{x_2})\rangle = -2g^2 \ln|\mathbf{x_1}-\mathbf{x_2}|$$

we wish to understand the statistics of its **minima/maxima** along various curves in the plane, and ultimately in various planar domains.

- The problem turns out to be intimately connected to the mechanism of **freezing transitions** in disordered systems theory (Random Energy Models, Dirac fermions in random magnetic field). It has also interesting relations to **Liouville Quantum Gravity** $\&$ conformal field theory, to **multifractal** random measures, $1/f$ **noises**, and processes arising in turbulence and mathematical finance, as well as to various aspects of **Random Matrix Theory**.

Idea of the method: We concentrate on considering samples of the Gaussian Free Field (**GFF**) along planar curves $\mathcal{C}$ parametrised by $\mathbf{x}(t) = (x(t), y(t))$ with real $t \in [a, b]$. To this end we introduce the regularized version $V_\epsilon(\mathbf{x})$ of the **GFF** with a short scale cutoff $\epsilon \ll 1$, i.e. zero mean and the covariance

$$\langle V_\epsilon(\mathbf{x}) V_\epsilon(\mathbf{x}') \rangle = -2g^2 \ln |\mathbf{x} - \mathbf{x}'|_\epsilon = \begin{cases} -2g^2 \ln |\mathbf{x} - \mathbf{x}'|, & |\mathbf{x} - \mathbf{x}'| > \epsilon \\ 2g^2 \ln(1/\epsilon), & |\mathbf{x} - \mathbf{x}'| < \epsilon \end{cases}$$

Given some measure $d\mu_\rho(t) = \rho(t)dt$ with the density $\rho(t) > 0$ for $t \in [a, b]$ our main object of study is the integral

$$Z = \epsilon^{\beta^2 g^2} \int_a^b e^{-\beta V_\epsilon(\mathbf{x}(t))} d\mu_\rho(t), \quad \beta > 0$$

which is to be interpreted as the **partition function** of the associated **Random Energy Model** at the temperature $T = \beta^{-1}$. This is to be studied in the limit $\epsilon \to 0$.

Note: For $\rho(t) = 1$ the Gibbs measure of the disordered system identifies with the random Liouville measure, and such Z can be interpreted as the (fluctuating) length of a curve in (critical) Liouville quantum gravity, cf. (**Duplantier** $\&$ **Sheffield**) .

Observation: The positive integer moments $\langle Z^n \rangle$, $n = 1, 2, \ldots$ turn out to be given by **statistical mechanics of Dyson Coulomb Gas with attraction, with effective coupling constant** $\gamma = \beta^2 g^2$.
In particular, choose the curve $\mathcal{C}$ to be
(i) **unit circle** $x(t) = \cos t, y(t) = \sin t,\ t \in [0, 2\pi)$ with the uniform weight $\rho(t) = 1$
(ii) **interval of the real axis** $x \in [0, 1]$ with the density $\rho(x) = x^a(1-x)^b$. Then

$$\langle Z^n_{circ} \rangle = \frac{1}{(2\pi)^n} \int_0^{2\pi} d\theta_1 \ldots \int_0^{2\pi} d\theta_n \prod_{a<b} |e^{i\theta_a} - e^{i\theta_b}|_\epsilon^{-2\gamma}$$

$$\left\langle Z^n_{[0,1]} \right\rangle = \int_0^1 \ldots \int_0^1 \prod_{1\le i<j\le n} |x_i - x_j|_\epsilon^{-2\gamma} \prod_{i=1}^{n} x_i^a (1-x_i)^b \, dx_i$$

For a fixed $n = 1, 2, ..$, a well defined and universal $\epsilon \to 0$ limits exists whenever the integrals are convergent, that is for $\gamma < 1/n$, in which case they are given by the famous **Selberg integral** formula. Defining $z = \Gamma(1-\gamma) Z = e^{-\beta f}$ we obtain

$$\langle z^n_{circ} \rangle = \Gamma(1 - n\gamma), \quad 1 \le n < 1/\gamma$$

$$\left\langle z^n_{[0,1]} \right\rangle = \prod_{j=1}^{j=n} \frac{\Gamma[1+a-(j-1)\gamma]\Gamma[1+b-(j-1)\gamma]\Gamma(1-j\gamma)}{\Gamma[2+a+b-(n+j-2)\gamma]}$$

Aim: to reconstruct the distribution $P(Z)$ from its moments in the high temperature phase $\gamma \leq 1$.

Difficulty: Only **finite** number of positive integer moments exists, as $\overline{Z^n}$ become infinite for $T > T_c^{(n)} = g\sqrt{n}$. The true transition in the full Gibbs measure happens however only at $T_c = g$ i.e. $\gamma = \gamma_c = 1$. Above T_c the distribution $P(Z)$ exists in the limit $\epsilon = 0$ and develops an algebraic tail. The formally divergent moments start depending on the cut-off parameter ϵ.

Way out: To continue analytically to **negative** integer moments which are all well-defined, and reconstruct the distribution $P(Z)$ from that end.

Implementation: Simple for the **unit circle** case where we just can change $n \to -n$ implying $\left\langle z_{circ}^{-n} \right\rangle = \Gamma(1 + n\gamma)$ which yields the probability density

$$P(z) = \gamma^{-1} z^{-1/\gamma - 1} \exp - z^{-1/\gamma}$$

$\Longrightarrow$ the free energy $f = -\beta^{-1} \ln z$ is distributed according to the **Gumbel law** $\mathcal{P}(f) = A \exp(Af - e^{Af}), \quad A = T/T_C^2$ for $T \geq T_c$.

Freezing scenario: Consider the generating function

$$g_\beta(x) = \left\langle \exp(-e^{\beta x} z) \right\rangle , \quad \beta = 1/T$$

For REM-type models the phase transition is believed to be described by the following freezing scenario: $g_\beta(x)$ depends on β for $T > T_c$ but **freezes** to the **temperature independent** profile everywhere in the low-temperature glassy phase $T \leq T_c$.

The picture is based on a heuristic **real-space renormalization group arguments** for the logarithmic models (**Carpentier, Le Doussal ’01**) revealing an analogy to the **travelling wave** analysis of polymers on disordered trees (**Derrida, Spohn 1989**).

Assuming validity of such scenario for the problem in hand, one finds the frozen profile for the circular model:

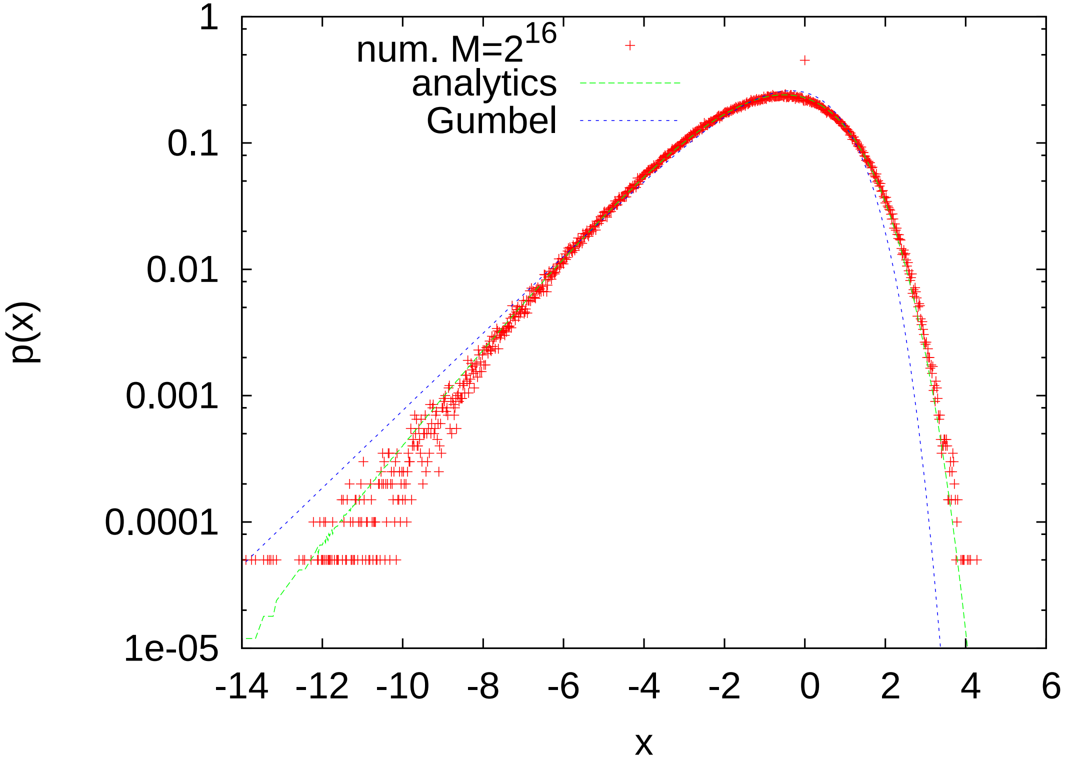
$$g_{\beta_c}^{circ}(x) = 2e^{x/2} K_1(2e^{x/2})$$

where $K_1(z)$ is the Macdonald function. This allows to reconstruct the distribution of the free energy $f = -\beta^{-1} \ln z$ for any $T < T_c$.

The minimum of the random potential is simply given by $V_{min} = -\lim_{T\to 0} f = const + x$, with known $const$ and the probability density of x related to the frozen profile $g_{\beta_c}(x)$ by

$$p(x) = -g'_{\beta_c}(x) = -\frac{d}{dx}\left[2e^{x/2}K_1(2e^{x/2})\right] \quad (1)$$

This is different from **Gumbel** distribution $p_{Gum}(x) = -\frac{d}{dx}\left[\exp -Be^{Ax}\right]$.



Distribution of extremes: we compare three distributions: (i) the histrogram for ensemble of 10^6 realizations of the Gaussian free field sampled at $M = 2^{16}$ points equispaced along the unit circle, (ii) the analytical prediction (1), and (iii) the Gumbel distribution for the mean $\&$ variance given by (1)

Numerics by A.Rosso

Unfortunately, such direct methods do not work for the **interval** case. However, the following **recursion relation**

$$\frac{z_n}{z_{n-1}} = \frac{\Gamma[1-n\gamma]\,\Gamma^2[1-(n-1)\gamma]\,\Gamma[2-(n-2)\gamma]}{\Gamma[2-(2n-3)\gamma]\,\Gamma[2-(2n-2)\gamma]}, \quad z_1 = \Gamma(1-\gamma)$$

holds for $a = b = 0$ and a similar formula for any a, b. We then perform a formal **continuation to negative integer moments** $m_k \equiv z_{-k}$ in the above recursion as

$$m_{k+1}/m_k \equiv z_{n-1}/z_n|_{n\to -k}, \quad m_0 = z_0 = 1$$

Solving the recursion and restoring a, b we find:

$$z_{-k} = \prod_{j=1}^{k} \frac{\Gamma[2+a+b+(k+j+1)\gamma]}{\Gamma[1+(j-1)\gamma]\,\Gamma[1+a+j\gamma]\Gamma[1+b+j\gamma]}$$

These expressions satisfy the convexity property $z_n^{p-m} z_p^{m-n} \geq z_m^{p-n}$ for any integers $n < m < p$ of arbitrary sign, which is a necessary condition for positivity of a probability. For $a = b = 0$ the negative moments were announced independently in **Ostrovsky D** *2008, Lett. Math. Phys* **83** *265*.

To restore the corresponding **probability density** we define the generic moments $M(s) = \langle z^{1-s} \rangle$, $M(1) = 1$ for any **complex** s, at fixed inverse temperature β. In particular, for the **critical** temperature $\beta = \beta_c$ and $a = b = 0$ we are able to find an analytical continuation for the moments which is given by

$$M(s) = \frac{2^{2s^2+s-2}}{G(5/2)^2 \pi^{s-1}} \frac{\Gamma(s+\frac{1}{2})^2}{\Gamma(s)\Gamma(s+2)} \frac{G(s+\frac{1}{2})^2}{G(s)^2} \tag{2}$$

where $G(s)$ is the *Barnes function*, which under some mild conditions is the only solution of:

$$G(s+1) = G(s)\Gamma(s), \quad \text{with} \quad G(1) = 1. \tag{3}$$

To guarantee that this is the correct continuation, we have checked
(i) **positivity**: $M(s)$ given above is finite and positive on the interval $s \in [0, +\infty[$ that is all real moments $n = 1 - s < 1$ exist.
(ii) **convexity**: on this interval $\partial_s^2 \ln M(s) > 0$.

The **probability density** $P(z)$ of the scaled partition function z and the **frozen profile** of the generation function $g_{\beta_c}(x)$ can be related to the function $M(s)$ via the contour integrals:

$$P\left(e^{-t}\right) = e^{2t}\frac{1}{2i\pi}\int_{s_0-i\infty}^{s_0+i\infty} e^{-st}M(s)ds, \quad g_{\beta_c}(x) = e^{x}\frac{1}{2i\pi}\int_{s_0-i\infty}^{s_0+i\infty} e^{-sx}M(s)\Gamma(s-1)\,ds$$

where the integration goes along the imaginary axis and the integral is convergent for $s_0 > 1$. Deforming the integration contour one obtains $g(x)$ as a sum of residues over the (multiple) poles of $M(s)$ at $s = -n$, which generates the expansion in powers of e^x:

$$g_{\beta_c}(x \to -\infty) = 1 + (x + A')e^{x} + (A + By + Cx^2 + \frac{1}{6}x^3)e^{2x} + \ldots \tag{4}$$

with $A' = 2\gamma_E + \ln(2\pi) - 1$ and $C = -0.253846$, $B = 1.25388$, $A = -5.09728$. Although this expansion is different from the circle case, the universal **Carpentier-Le Doussal** tail predicted by the renormalization group analysis is shared by both distributions: $p(x \to -\infty) = -g'_{\beta_c}(x \to -\infty) \sim -xe^{x}$. It has its origin in the $1/z^2$ forward tail which $P(z)$ of z develops at critical $\beta = \beta_c$, with the first moment $< z >$ becoming infinite.

Conclusions & Discussions:

- Using the methods of statistical mechanics we were able to conjecture the explicit expressions for distributions of minima of the Gaussian Free Field sampled along **(i)** circles of unit radius and **(ii)** intervals of unit length. The distributions are manifestly non-Gumbel. The results are expected to describe extreme value statistics for $1/f$ signals, and in this way could be relevant for spectral fluctuations of random matrices and chaotic systems (cf. e.g. $Relaño\ et\,al$ PRL **89** (2002) 244102).

- **Our method is based on a few assumptions, most importantly**
  (i) freezing scenario for REM models, and (ii) ability to continue analytically moments given by Selberg integrals away from positive integers. It remains a challenge:
  (1) to verify/justify/extend the assumptions/steps of the derivation; e.g. the continuation fails for the Gaussian density $\rho(t) = e^{-\frac{t^2}{2}}, t \in [-\infty, \infty]$ **(2) to understand universality of the results for other** $1d$ **curves (3) access extreme value statistics of GFF in 2D domains.**

Y V Fyodorov, J-P Bouchaud : J. Phys.A: Math.Theor 41 (2008), 372001 (12pp)

YV Fyodorov, P Le Doussal , and A Rosso , under preparation