Asymptotic statistics of cycles in Surrogate-Spatial Random Permutations

Dirk Zeindler
(Joint work with Leonid Bogachev)

University of Bielefeld SFB 701, Germany

15 November 2012

Outline

Spatial random permutations are motivated by a model from quantum mechanics and have recently studied by V. Betz and D. Ueltschi.

Spatial random permutations are motivated by a model from quantum mechanics and have recently studied by V. Betz and D. Ueltschi.

Unfortunately, the computations with spatial random permutation have technical difficulties.

Spatial random permutations are motivated by a model from quantum mechanics and have recently studied by V. Betz and D. Ueltschi.

Unfortunately, the computations with spatial random permutation have technical difficulties.

We propose a natural approximation of spatial random permutations on the symmetric group S_n.

Spatial random permutations are motivated by a model from quantum mechanics and have recently studied by V. Betz and D. Ueltschi.

Unfortunately, the computations with spatial random permutation have technical difficulties.

We propose a natural approximation of spatial random permutations on the symmetric group S_n.

This approximation suggested here has a simpler structure and is thus more tractable.

This two measure share many properties, for instance

- Two possible behaviours (only small cycles or some fraction of long cycles)

This two measure share many properties, for instance

- Two possible behaviours (only small cycles or some fraction of long cycles)
- the same critical density

This two measure share many properties, for instance

- Two possible behaviours (only small cycles or some fraction of long cycles)
- the same critical density
- the same splitting into small and long cycles

This two measure share many properties, for instance

- Two possible behaviours (only small cycles or some fraction of long cycles)
- the same critical density
- the same splitting into small and long cycles
- similar behaviour of long cycles

This two measure share many properties, for instance

- Two possible behaviours (only small cycles or some fraction of long cycles)
- the same critical density
- the same splitting into small and long cycles
- similar behaviour of long cycles

Furthermore, we obtain a few new results about the distribution of the cycle lengths.

The structure of this talk is as follows

- Introduction.

The structure of this talk is as follows

- Introduction.
- Ewens measure

The structure of this talk is as follows

- Introduction.
- Ewens measure
- Spatial permutations

The structure of this talk is as follows

- Introduction.
- Ewens measure
- Spatial permutations
- Surrogate-spatial permutation

Outline

Ewens measure

An important and special case of the spatial measure and the surrogate spatial measure is the Ewens measure.

Ewens measure

An important and special case of the spatial measure and the surrogate spatial measure is the Ewens measure.
This was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

Ewens measure

An important and special case of the spatial measure and the surrogate spatial measure is the Ewens measure.
This was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

- It has a connection with Kingman's coalescent process (1982).

Ewens measure

An important and special case of the spatial measure and the surrogate spatial measure is the Ewens measure.
This was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

- It has a connection with Kingman's coalescent process (1982).
- It has been used to model the dynamics of tumour evolution. (Barbour and Tavaré (2010))

Ewens measure

An important and special case of the spatial measure and the surrogate spatial measure is the Ewens measure.
This was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

- It has a connection with Kingman's coalescent process (1982).
- It has been used to model the dynamics of tumour evolution. (Barbour and Tavaré (2010))
- It appears in a Bayesian non parametric statistics stetting. (Antoniak (1974))

Ewens measure

An important and special case of the spatial measure and the surrogate spatial measure is the Ewens measure.
This was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

- It has a connection with Kingman's coalescent process (1982).
- It has been used to model the dynamics of tumour evolution. (Barbour and Tavaré (2010))
- It appears in a Bayesian non parametric statistics stetting. (Antoniak (1974))
- It plays a crucial role for virtual permutations since it is central and stable under the restriction $S_n \to S_{n-1}$

Ewens measure

An important and special case of the spatial measure and the surrogate spatial measure is the Ewens measure.
This was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

- It has a connection with Kingman's coalescent process (1982).
- It has been used to model the dynamics of tumour evolution. (Barbour and Tavaré (2010))
- It appears in a Bayesian non parametric statistics stetting. (Antoniak (1974))
- It plays a crucial role for virtual permutations since it is central and stable under the restriction $S_n \to S_{n-1}$
- ....

A cycle $(s_0\ s_1 \ldots s_{k-1})$ is a permutation which maps

$$s_0 \mapsto s_1 \mapsto s_2 \mapsto \cdots \mapsto s_{k-1} \mapsto s_k = s_0.$$

and agrees with the identity on the remaining points.

A cycle $(s_0\ s_1 \dots s_{k-1})$ is a permutation which maps

$$s_0 \mapsto s_1 \mapsto s_2 \mapsto \cdots \mapsto s_{k-1} \mapsto s_k = s_0.$$

and agrees with the identity on the remaining points.
Two cycles $(s_0 \dots s_{k-1})$ and $(t_0 \dots t_{m-1})$ are called disjoint if the sets $\{s_0, \dots, s_{k-1}\}$ and $\{t_0, \dots, t_{m-1}\}$ are disjoint.

If $\sigma \in S_n$ is given, then it can be written as

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell$$

where $\sigma_1, \ldots, \sigma_\ell$ are disjoint cycles.

If $\sigma \in S_n$ is given, then it can be written as

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell$$

where $\sigma_1, \ldots, \sigma_\ell$ are disjoint cycles.
The Ewens measure is defined for $\vartheta > 0$ as

$$\mathbb{P}[\sigma] := \frac{\vartheta^\ell}{K_n}$$

If $\sigma \in S_n$ is given, then it can be written as

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell$$

where $\sigma_1, \ldots, \sigma_\ell$ are disjoint cycles.
The Ewens measure is defined for $\vartheta > 0$ as

$$\mathbb{P}[\sigma] := \frac{\vartheta^\ell}{\vartheta(\vartheta+1)\cdots(\vartheta+n-1)}$$

Small Cycles

Let

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell .$$

Small Cycles

Let

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell.$$

We write λ_j for the length of the cycle σ_j.

Small Cycles

Let

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell.$$

We write λ_j for the length of the cycle σ_j.
We define the cycle counts as

$$C_k := \# \{j; \lambda_j = k\}.$$

Small Cycles

Let

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell.$$

We write λ_j for the length of the cycle σ_j.
We define the cycle counts as

$$C_k := \# \{j; \lambda_j = k\}.$$

Theorem (Shepp,Loyd (1966) $\vartheta = 1$, Watterson (1974) general ϑ)

$$(C_1, \dots, C_b) \xrightarrow{d} (Y_1, \dots, Y_b)$$

with Y_k independent Poisson distributed with $\mathbb{E}[Y_k] = \frac{\vartheta}{k}$

Total Number of Cycles

The Total number of cycles is defined as $T_n := C_1 + \cdots + C_n$.

Total Number of Cycles

The Total number of cycles is defined as $T_n := C_1 + \cdots + C_n$.

Theorem (Goncharov (1942) $\vartheta = 1$, Watterson (1974) general ϑ)

$$\frac{T_n - \vartheta \log(n)}{\sqrt{\vartheta \log(n)}} \xrightarrow{d} N(0,1)$$

Long Cycles

Let

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell.$$

We write λ_j for the length of the cycle σ_j.

Long Cycles

Let

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell.$$

We write λ_j for the length of the cycle σ_j.
W.l.o.g. we can assume $\lambda_1 \geq \lambda_2 \geq \ldots$.

Long Cycles

Let

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell .$$

We write λ_j for the length of the cycle σ_j.
W.l.o.g. we can assume $\lambda_1 \geq \lambda_2 \geq \ldots$.

Theorem (Vershik and Shmidt (1977) resp. Kingman (1977))

$$\left(\frac{\lambda_1}{n}, \frac{\lambda_2}{n}, \ldots\right) \xrightarrow{d} \mathcal{PD}(\vartheta), \qquad (n \to \infty)$$

with $\mathcal{PD}(\vartheta)$ the Poisson–Dirichlet distribution with parameter ϑ.

What is the Poisson–Dirichlet distribution?

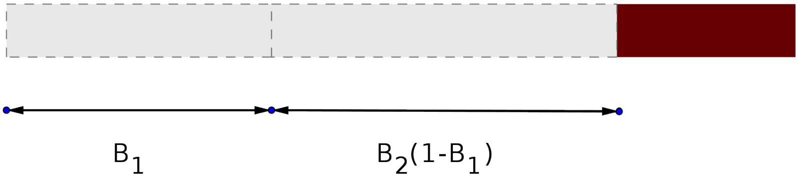
What is the Poisson–Dirichlet distribution?
The best way to describe this is the stick breaking process with size ordering.

What is the Poisson–Dirichlet distribution?
The best way to describe this is the stick breaking process with size ordering.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1,\vartheta)$ and consider a stick of length 1

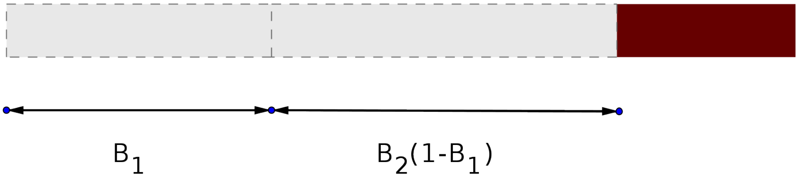
What is the Poisson–Dirichlet distribution?
The best way to describe this is the stick breaking process with size ordering.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1, \vartheta)$ and consider a stick of length 1

What is the Poisson–Dirichlet distribution?
The best way to describe this is the stick breaking process with size ordering.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1, \vartheta)$ and consider a stick of length 1

What is the Poisson–Dirichlet distribution?
The best way to describe this is the stick breaking process with size ordering.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1,\vartheta)$ and consider a stick of length 1



What is the Poisson–Dirichlet distribution?
The best way to describe this is the stick breaking process with size ordering.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1, \vartheta)$ and consider a stick of length 1



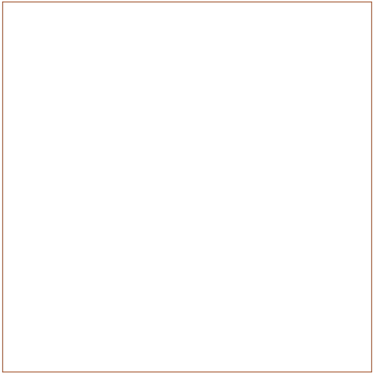
Ordering the sticks obtained by this process by size then has a Poisson–Dirichlet distribution.

Outline

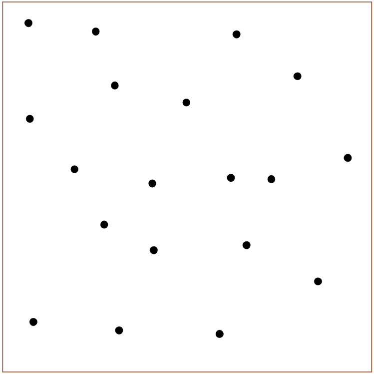
Spatial permutations occur as a model in quantum mechanics. More precisely, as a model for the Feynman–Kac representation of the dilute Bose gas.

Spatial permutations occur as a model in quantum mechanics. More precisely, as a model for the Feynman–Kac representation of the dilute Bose gas.

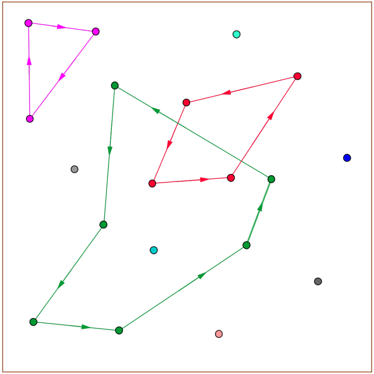
Let us first describe the idea of the model.



- A cube $\Lambda \subset \mathbb{R}^d$ with side length $L > 0$ $(d \geq 3)$



- A cube $\Lambda \subset \mathbb{R}^d$ with side length $L > 0$ $(d \geq 3)$
- n particles in the cube Λ



- A cube $\Lambda \subset \mathbb{R}^d$ with side length $L > 0$ $(d \geq 3)$
- n particles in the cube Λ
- A permutation σ of particles with the same state

We now define a measure on $\mathbb{P}[.]$ on $S_n \times \Lambda^n$ with

$$\mathbb{P}[\sigma, dx] = \frac{1}{Y_n n!} e^{-H(\sigma,x)} dx$$

with dx the Lebesgue measure, Y_n a normalisation constant and

We now define a measure on $\mathbb{P}[.]$ on $S_n \times \Lambda^n$ with

$$\mathbb{P}[\sigma, dx] = \frac{1}{Y_n n!} e^{-H(\sigma,x)} dx$$

with dx the Lebesgue measure, Y_n a normalisation constant and

$$H(\sigma, x) = \sum_{k=1}^{n} \alpha_k C_k + \sum_{j=1}^{n} \|x_j - x_{\sigma(j)}\|$$

where $\alpha_k \in \mathbb{R}$, $x = (x_1, \ldots, x_n) \in \Lambda^n \subset \mathbb{R}^{nd}$ are the coordinates of the particles.

Let us take a look at

$$H(\sigma, x) = \sum_{k=1}^{n} \alpha_k C_k + \sum_{j=1}^{n} \|x_j - x_{\sigma(j)}\|$$

Let us take a look at

$$H(\sigma, x) = \sum_{k=1}^{n} \alpha_k C_k + \sum_{j=1}^{n} \|x_j - x_{\sigma(j)}\|$$

The α_k model the particle interaction.

Let us take a look at

$$H(\sigma, x) = \sum_{k=1}^{n} \alpha_k C_k + \sum_{j=1}^{n} \|x_j - x_{\sigma(j)}\|$$

The α_k model the particle interaction.
A reasonable choice is $\alpha_k \to \alpha$.

Let us take a look at

$$H(\sigma, x) = \sum_{k=1}^{n} \alpha_k C_k + \sum_{j=1}^{n} \|x_j - x_{\sigma(j)}\|$$

The α_k model the particle interaction.
A reasonable choice is $\alpha_k \to \alpha$.

The norm $\| \cdot \|$ forces particles of the same state to stay together.

Let us take a look at

$$H(\sigma, x) = \sum_{k=1}^{n} \alpha_k C_k + \sum_{j=1}^{n} \|x_j - x_{\sigma(j)}\|$$

The α_k model the particle interaction.
A reasonable choice is $\alpha_k \to \alpha$.

The norm $\| \cdot \|$ forces particles of the same state to stay together.

Particles of different states do not interact.

We study the behavior of this system with respect to the thermodynamic limit.

We study the behavior of this system with respect to the thermodynamic limit.

Thermodynamic limit $= n \to \infty$ while keeping $\rho := \frac{n}{L^d} = \frac{n}{|\Lambda|}$ fixed

We study the behavior of this system with respect to the thermodynamic limit.

Thermodynamic limit = $n \to \infty$ while keeping $\rho := \frac{n}{L^d} = \frac{n}{|\Lambda|}$ fixed

Condensation = Only one state occurs in the limit infinitely often.

We study the behavior of this system with respect to the thermodynamic limit.

Thermodynamic limit $=$ $n \to \infty$ while keeping $\rho := \frac{n}{L^d} = \frac{n}{|\Lambda|}$ fixed

Condensation $=$ Only one state occurs in the limit infinitely often.

The first step is study the existence or non-existence of infinite sets of particles of the same state (in the limit).

We study the behavior of this system with respect to the thermodynamic limit.

Thermodynamic limit $= n \to \infty$ while keeping $\rho := \frac{n}{L^d} = \frac{n}{|\Lambda|}$ fixed

Condensation = Only one state occurs in the limit infinitely often.

The first step is study the existence or non-existence of infinite sets of particles of the same state (in the limit).

In this setting large sets of particles of the same state correspond to long cycles in the cycle decomposition of $\sigma \in S_n$.

How can one measure the existence of infinite cycles in the limit?

How can one measure the existence of infinite cycles in the limit?
For finite n, all cycles are finite.

How can one measure the existence of infinite cycles in the limit?
For finite n, all cycles are finite.
Let us consider for $p \in \mathbb{N}$

$$\frac{1}{n} \sum_{k \geq p} k C_k$$

How can one measure the existence of infinite cycles in the limit?
For finite n, all cycles are finite.
Let us consider for $p \in \mathbb{N}$

$$\frac{1}{n} \sum_{k \geq p} k C_k$$

This has the interpretation as the fraction of particles in cycles of length at least p.

We now define

$$\nu_p := \liminf_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\sum_{k\geq p} kC_k\right]$$

We now define

$$\nu_p := \liminf_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\sum_{k\geq p} kC_k\right]$$

and

$$\nu := \lim_{p\to\infty} \nu_p.$$

We now define

$$\nu_p := \liminf_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\sum_{k\geq p} kC_k\right]$$

and

$$\nu := \lim_{p\to\infty} \nu_p.$$

For the Ewens measure, we have $\nu = 1$.

It was shown by Betz and Ueltschi for
'$\alpha_k \to \alpha$' and $\alpha_k = \gamma \log(k)$ for $\gamma > 0$

$$\nu = \max\left\{0, 1 - \frac{\rho_c}{\rho}\right\}$$

It was shown by Betz and Ueltschi for
'$\alpha_k \to \alpha$' and $\alpha_k = \gamma \log(k)$ for $\gamma > 0$

$$\nu = \max\left\{0, 1 - \frac{\rho_c}{\rho}\right\}$$

with

$$\rho_c := \sum_{k=1}^{\infty} e^{-\alpha_k} \int_{\mathbb{R}^d} e^{-k\|x\|}\, dx.$$

Betz and Ueltschi could also compute the behaviour of the large cycles for $\rho > \rho_c$.

Betz and Ueltschi could also compute the behaviour of the large cycles for $\rho > \rho_c$.
In the case '$\alpha_k \to \alpha$' we have

$$\left(\frac{\lambda_1}{\nu n}, \frac{\lambda_2}{\nu n}, \dots \right) \xrightarrow{d} \mathcal{PD}(e^{-\alpha})$$

and in the case $\alpha_k = \gamma \log(k)$

$$\frac{\lambda_1}{\nu n} \xrightarrow{d} 1$$

Outline

Remember we had

$$\mathbb{P}[\sigma, dx] = \frac{1}{Y_n n!} e^{-H(\sigma,x)} dx$$

with

$$H(\sigma, x) = \sum_{k=1}^{n} \alpha_k C_k + \sum_{j=1}^{n} \|x_j - x_{\sigma(j)}\|.$$

Remember we had

$$\mathbb{P}\left[\sigma, dx\right] = \frac{1}{Y_n n!} e^{-H(\sigma,x)} dx$$

with

$$H(\sigma, x) = \sum_{k=1}^{n} \alpha_k C_k + \sum_{j=1}^{n} \|x_j - x_{\sigma(j)}\|.$$

"Periodizing" the boundary conditions gives and integrating out the x gives

$$\mathbb{P}\left[\sigma\right] := \frac{1}{Y_n n!} \prod_{k=1}^{n} \left(e^{-\alpha_k} \sum_{m \in \mathbb{Z}^d} e^{-k\|m/L\|} \right)^{C_k},$$

We now have for fixed k and using $\rho = \frac{n}{L^d}$

$$\frac{1}{L^d} \sum_{m \in \mathbb{Z}^d} e^{-k\|m/L\|} \approx \int_{\mathbb{R}^d} e^{-k\|x\|}\, dx$$

We now have for fixed k and using $\rho = \frac{n}{L^d}$

$$\sum_{m\in\mathbb{Z}^d} e^{-k\|m/L\|} \approx \frac{n}{\rho}\int_{\mathbb{R}^d} e^{-k\|x\|}\,dx$$

We now have for fixed k and using $\rho = \frac{n}{L^d}$

$$\sum_{m \in \mathbb{Z}^d} e^{-k\|m/L\|} \approx \frac{n}{\rho} \int_{\mathbb{R}^d} e^{-k\|x\|}\, dx$$

and thus

$$\begin{aligned}
\mathbb{P}\left[\sigma\right] &= \frac{1}{Y_n n!} \prod_{k=1}^{n} \left(e^{-\alpha_k} \sum_{m \in \mathbb{Z}^d} e^{-k\|m/L\|} \right)^{C_k} \\
&\approx \frac{1}{Y_n n!} \prod_{k=1}^{n} \left(n \cdot \frac{e^{-\alpha_k}}{\rho} \int_{\mathbb{R}^d} e^{-k\|x\|}\, dx \right)^{C_k}.
\end{aligned}$$

We make the following Ansatz

Definition

Let $\Theta = (\theta_k)_{k\geq 1}$ and $\Upsilon = (\tau_k)_{k\geq 1}$ be given, with $\theta_k, \tau_k \geq 0$. We then define the *surrogate spatial* probability measure on permutations as

$$\mathbb{P}_n^{(sur)}[\sigma] := \frac{1}{H_n n!} \prod_{k=1}^{n} (n \cdot \tau_k + \theta_k)^{C_k},$$

with H_n some constant.

Are the error-terms θ_k important?

Are the error-terms θ_k important?
If one computes the τ_k arising from spatial permutations, one gets

'$\alpha_k \to \alpha$'	$\alpha_k = \gamma \log(k)$
$\tau_k = \frac{1}{\rho} e^{-\alpha} k^{-d/2}$	$\tau_k = \frac{1}{\rho} k^{-d/2-\gamma}$

Are the error-terms θ_k important?
If one computes the τ_k arising from spatial permutations, one gets

'$\alpha_k \to \alpha$'	$\alpha_k = \gamma \log(k)$
$\tau_k = \frac{1}{\rho} e^{-\alpha} k^{-d/2}$	$\tau_k = \frac{1}{\rho} k^{-d/2-\gamma}$
$\left(\frac{\lambda_1}{\nu n}, \frac{\lambda_2}{\nu n}, \dots \right) \xrightarrow{d} \mathcal{PD}(e^{-\alpha})$	$\frac{\lambda_1}{\nu n} \xrightarrow{d} 1$

Strategy to analyse large$-n$ asymptotics of surrogate-spatial permutations:

- Use generating functions to give ‘nice’ expressions
- Apply Cauchy’s integral formula

Strategy to analyse large$-n$ asymptotics of surrogate-spatial permutations:

- Use generating functions to give ‘nice’ expressions
- Apply Cauchy’s integral formula

We illustrate this with the normalisation constant H_n in the surrogate-spatial measure.

We work here only with class functions on S_n and it is well known that the conjugacy class of S_n can be parametrised with partitions.

We work here only with class functions on S_n and it is well known that the conjugacy class of S_n can be parametrised with partitions.

Lemma

For any class function $u : S_n \to \mathbb{C}$, there is the identity

$$\frac{1}{n!}\sum_{\sigma\in S_n} u(\sigma) = \sum_{\lambda\in\mathcal{P}_n} \frac{1}{z_\lambda} u(\mathcal{C}_\lambda)$$

where $\mathcal{C}_\lambda$ is the conjugacy class corresponding to partition λ and $z_\lambda := \prod_{k=1}^{n} k^{C_k} C_k!$.

The main tool in this talk to write down generating functions

Lemma (Polya)

Let $(a_m)_{m\in\mathbb{N}}$ be a sequence of complex numbers. Then

$$\sum_{\lambda\in\mathcal{P}} \frac{1}{z_\lambda} \left(\prod_{m=1}^{\ell(\lambda)} a_{\lambda_m} \right) t^{|\lambda|} = \exp\left(\sum_{m=1}^{\infty} \frac{1}{m} a_m t^m \right)$$

with z_λ as above.
If one of the sums is absolutely convergent then so is the other one.

It follows form the definition of $\mathbb{P}_n^{(sur)}[\cdot]$ that

$$H_n = \sum_{\lambda \in \mathcal{P}_n} \frac{1}{z_\lambda} \prod_{k=1}^{\infty} (n \cdot \tau_k + \theta_k)^{C_k}$$

It follows form the definition of $\mathbb{P}_n^{(sur)}[\cdot]$ that

$$H_n = \sum_{\lambda \in \mathcal{P}_n} \frac{1}{z_\lambda} \prod_{k=1}^{\infty} (n \cdot \tau_k + \theta_k)^{C_k}$$

Unfortunately we can not directly compute

$$\sum_{n=0}^{\infty} H_n t^n \quad \text{or} \quad \sum_{n=0}^{\infty} H_n \frac{t^n}{n!} \quad \text{or} \quad \ldots$$

It follows form the definition of $\mathbb{P}_n^{(sur)}[\cdot]$ that

$$H_n = \sum_{\lambda \in \mathcal{P}_n} \frac{1}{z_\lambda} \prod_{k=1}^{\infty} (n \cdot \tau_k + \theta_k)^{C_k}$$

Unfortunately we can not directly compute

$$\sum_{n=0}^{\infty} H_n t^n \quad \text{or} \quad \sum_{n=0}^{\infty} H_n \frac{t^n}{n!} \quad \text{or} \quad \dots$$

We thus introduce for $v \in \mathbb{N}$

$$h_n(v) = \sum_{\lambda \in \mathcal{P}_n} \frac{1}{z_\lambda} \prod_{k=1}^{\infty} (v \cdot \tau_k + \theta_k)^{C_k}$$

We get for each $v \in \mathbb{N}$

$$\sum_{n=0}^{\infty} h_n(v)\, t^n =$$

We get for each $v \in \mathbb{N}$

$$\sum_{n=0}^{\infty} h_n(v)\, t^n = \exp\left(g_\Theta(t) + v \cdot p_\Upsilon(t)\right)$$

with

$$g_\Theta(t) := \sum_{k=1}^{\infty} \frac{\theta_k}{k}\, t^k, \qquad p_\Upsilon(t) := \sum_{k=1}^{\infty} \frac{\tau_k}{k}\, t^k.$$

We get for each $v \in \mathbb{N}$

$$\sum_{n=0}^{\infty} h_n(v)\, t^n = \exp\left(g_\Theta(t) + v \cdot p_\Upsilon(t)\right)$$

with

$$g_\Theta(t) := \sum_{k=1}^{\infty} \frac{\theta_k}{k}\, t^k, \qquad p_\Upsilon(t) := \sum_{k=1}^{\infty} \frac{\tau_k}{k}\, t^k.$$

and thus

$$H_n = [t^n]\left[\exp\left(g_\Theta(t) + n \cdot p_\Upsilon(t)\right)\right]$$

Assume that $g_\Theta(t)$ and $p_\Upsilon(t)$ have radius of convergence $R > 0$.

Assume that $g_\Theta(t)$ and $p_\Upsilon(t)$ have radius of convergence $R > 0$. It turns out that one has to distinguish the cases

$$a(R) := Rp'_\Upsilon(R) > 1 \quad \text{and} \quad a(R) < 1 \quad \text{and} \quad a(R) = 1$$

Assume that $g_\Theta(t)$ and $p_\Upsilon(t)$ have radius of convergence $R > 0$. It turns out that one has to distinguish the cases

$$\underbrace{a(R) := Rp'_\Upsilon(R) > 1}_{\text{sub-critical}} \quad \text{and} \quad \underbrace{a(R) < 1}_{\text{super-critical}} \quad \text{and} \quad \underbrace{a(R) = 1}_{\text{critical}}$$

Sub-critical case

If we apply Cauchy's integral formula, we get

$$H_n = \frac{1}{2\pi i} \int_{\gamma} \frac{\exp\left(g_\Theta(z) + n \cdot p_\Upsilon(z)\right)}{z^{n+1}} dz$$

Sub-critical case

If we apply Cauchy's integral formula, we get

$$H_n = \frac{1}{2\pi i} \int_\gamma \frac{\exp\left(g_\Theta(z) + n \cdot p_\Upsilon(z)\right)}{z^{n+1}} dz \approx \int \Big(f(x)\Big)^n dx$$

Sub-critical case

If we apply Cauchy's integral formula, we get

$$H_n = \frac{1}{2\pi i}\int_\gamma \frac{\exp\left(g_\Theta(z) + n \cdot p_\Upsilon(z)\right)}{z^{n+1}}dz \approx \int \Big(f(x)\Big)^n dx$$

We choose $\gamma(\varphi) = re^{i\varphi}$ with $\varphi \in [-\pi, \pi]$.

Sub-critical case

If we apply Cauchy's integral formula, we get

$$H_n = \frac{1}{2\pi i} \int_\gamma \frac{\exp\left(g_\Theta(z) + n \cdot p_\Upsilon(z)\right)}{z^{n+1}} dz \approx \int \Big(f(x)\Big)^n dx$$

We choose $\gamma(\varphi) = re^{i\varphi}$ with $\varphi \in [-\pi, \pi]$. It's clear that $\mathrm{Re}(p_\Upsilon(re^{i\varphi}))$ has a maximum for $\varphi = 0$.

We expand

$$p_\Upsilon(re^{i\varphi}) = p_\Upsilon(r) + i\varphi a(r) - \frac{\varphi^2}{2} b(r) + O(\varphi^3),$$

with $a(r) = rp'_\Upsilon(r)$ and $b(r) = rp'_\Upsilon(r) + r^2 p''_\Upsilon(r)$.

We expand

$$p_\Upsilon(re^{i\varphi}) = p_\Upsilon(r) + i\varphi a(r) - \frac{\varphi^2}{2} b(r) + O(\varphi^3),$$

with $a(r) = rp'_\Upsilon(r)$ and $b(r) = rp'_\Upsilon(r) + r^2 p''_\Upsilon(r)$.
Inserting this expansion into the integral and considering only a small neighbourhood $[-\kappa_n, \kappa_n]$ of $\varphi = 0$ gives (setting $g_\Theta \equiv 0$)

We expand

$$p_\Upsilon(re^{i\varphi}) = p_\Upsilon(r) + i\varphi a(r) - \frac{\varphi^2}{2} b(r) + O(\varphi^3),$$

with $a(r) = rp'_\Upsilon(r)$ and $b(r) = rp'_\Upsilon(r) + r^2 p''_\Upsilon(r)$.
Inserting this expansion into the integral and considering only a small neighbourhood $[-\kappa_n, \kappa_n]$ of $\varphi = 0$ gives (setting $g_\Theta \equiv 0$)

$$\frac{\exp\left(n \cdot p_\Upsilon(r)\right)}{2\pi r^n \sqrt{n}} \int_{-\kappa_n\sqrt{n}}^{\kappa_n\sqrt{n}} e^{i\sqrt{n}\varphi(a(r)-1)} e^{-b(r)\frac{x^2}{2}} (1 + o(1)) dx$$

We expand

$$p_\Upsilon(re^{i\varphi}) = p_\Upsilon(r) + i\varphi a(r) - \frac{\varphi^2}{2} b(r) + O(\varphi^3),$$

with $a(r) = rp'_\Upsilon(r)$ and $b(r) = rp'_\Upsilon(r) + r^2 p''_\Upsilon(r)$.
Inserting this expansion into the integral and considering only a small neighbourhood $[-\kappa_n, \kappa_n]$ of $\varphi = 0$ gives (setting $g_\Theta \equiv 0$)

$$\frac{\exp\left(n \cdot p_\Upsilon(r)\right)}{2\pi r^n \sqrt{n}} \int_{-\kappa_n\sqrt{n}}^{\kappa_n\sqrt{n}} e^{i\sqrt{n}\varphi(a(r)-1)} e^{-b(r)\frac{x^2}{2}} (1 + o(1)) dx$$

We now choose r to be the solution of $a(r) = 1$ (if it's possible).

Theorem

Assume that $g_\Theta(t)$ and $p_\Upsilon(t)$ have radius of convergence $R > 0$. Suppose that

$$a(R) > 1 \text{ with } a(R) := \sup_{0<r<R} a(r) \leq \infty$$

and let r_1 be the unique solution of $a(r_1) = 1$. Suppose further that $p_\Upsilon(t) \neq f(t^k)$ with $k > 1$ and f holomorphic.
We then have as $n \to \infty$

$$H_n \sim \frac{\exp\big(g_\Theta(r_1) + np_\Upsilon(r_1)\big)}{r_1^{n-d}\sqrt{2\pi nb(r_1)}}$$

Super-critical case

What can we do if $a(R) < 1$?

Super-critical case

What can we do if $a(R) < 1$?
We have in this case $p_\Upsilon(R) < \infty$ since $a(t) = tp'_\Upsilon(t)$ and remember

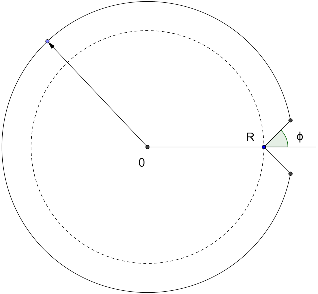
$$H_n = [t^n] \left[\exp\left(g_\Theta(t) + n \cdot p_\Upsilon(t)\right)\right]$$

Super-critical case

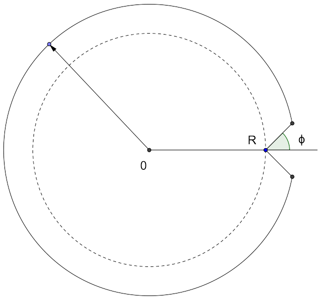
What can we do if $a(R) < 1$?
We have in this case $p_\Upsilon(R) < \infty$ since $a(t) = tp'_\Upsilon(t)$ and remember

$$H_n = [t^n]\left[\exp\left(g_\Theta(t) + n \cdot p_\Upsilon(t)\right)\right]$$

If $g_\Theta(t)$ is diverging at R, we can hope that H_n is determined by the behaviour of p_Υ and g_Θ near R.



$$g_\Theta(t) = -\vartheta \log(1 - t/R) + O(1 - t/R),$$
$$p_\Upsilon(t) = p_\Upsilon(R) + a(R)(t/R - 1) + O((1 - t/R)^2).$$



$$g_\Theta(t) = -\vartheta \log(1 - t/R) + O(1 - t/R),$$
$$p_\Upsilon(t) = p_\Upsilon(R) + a(R)(t/R - 1) + O((1 - t/R)^2).$$

Theorem

For $a(R) < 1$ and under the above assumptions we have

$$H_n \sim \frac{n^{\vartheta-1} \exp(n p_\Upsilon(R))}{R^n} \frac{(1 - a(R))^{\vartheta-1}}{\Gamma(\vartheta)}$$

Let us define the quantity r_* as

$$r_* := \begin{cases} r_1, & a(R) \geq 1, \\ R, & a(R) \leq 1, \end{cases} \qquad (1)$$

where r_1 is the (unique) solution of the equation

$$a(r_1) = 1$$

Cycle counts

Theorem

We have

$$\mathbb{E}_n^{(sur)}\left[(C_k)^m\right] \sim \left(n \cdot \frac{\tau_k r_*^k}{k}\right)^m$$

and $\frac{C_k}{n}$ *converges in law to the constant* $\frac{\tau_k r_*^k}{k}$.

Using the previous result, we obtain for each $p \in \mathbb{N}$

$$\nu_p = \liminf_{n\to\infty}\left(1 - \frac{1}{n}\mathbb{E}\left[\sum_{k=1}^{p} kC_k\right]\right) = 1 - \sum_{k=1}^{p} \tau_k r_*^k$$

Using the previous result, we obtain for each $p \in \mathbb{N}$

$$\nu_p = \liminf_{n\to\infty} \left(1 - \frac{1}{n}\mathbb{E}\left[\sum_{k=1}^{p} kC_k\right]\right) = 1 - \sum_{k=1}^{p} \tau_k r_*^k$$

This then gives

$$\nu = \lim_{p\to\infty} \left(1 - \sum_{k=1}^{p} \tau_k r_1^m\right) = 1 - a(r_*)$$

Using the previous result, we obtain for each $p \in \mathbb{N}$

$$\nu_p = \liminf_{n\to\infty} \left(1 - \frac{1}{n}\mathbb{E}\left[\sum_{k=1}^{p} kC_k\right]\right) = 1 - \sum_{k=1}^{p} \tau_k r_*^k$$

This then gives

$$\nu = \lim_{p\to\infty} \left(1 - \sum_{k=1}^{p} \tau_k r_1^m\right) = 1 - a(r_*)$$
$$= \begin{cases} 0 & \text{if } a(R) \geq 1 \\ 1 - a(R) & \text{if } a(R) < 1 \end{cases}$$

Total number of cycles

Theorem

(a) *Let* $a(R) > 1$. *Then,*

$$\frac{T_n - np_\Upsilon(r_1)}{\sqrt{n\left[p_\Upsilon(r_1) - 1/b(r_1)\right]}} \xrightarrow{d} \mathcal{N}(0,1), \qquad n \to \infty, \tag{2}$$

(b) *If* $a(R) < 1$ *then,*

$$\frac{T_n - np_\Upsilon(R)}{\sqrt{np_\Upsilon(R)}} \xrightarrow{d} \mathcal{N}(0,1), \qquad n \to \infty. \tag{3}$$

First comparison of models

If we take $\tau_k = \frac{k^{-\alpha}}{\rho}$ with $\alpha > 1$, we get the polylogarithm

$$p_\Upsilon(t) = \frac{\mathrm{Li}_{\alpha+1}(t)}{\rho} = \sum_{k=1}^{\infty} \frac{t^k}{\rho k^{\alpha+1}} \text{ and } a(t) = \frac{\mathrm{Li}_\alpha(t)}{\rho}.$$

First comparison of models

If we take $\tau_k = \frac{k^{-\alpha}}{\rho}$ with $\alpha > 1$, we get the polylogarithm

$$p_\Upsilon(t) = \frac{\mathrm{Li}_{\alpha+1}(t)}{\rho} = \sum_{k=1}^{\infty} \frac{t^k}{\rho k^{\alpha+1}} \text{ and } a(t) = \frac{\mathrm{Li}_\alpha(t)}{\rho}.$$

We have in this case $R = 1$

$$a(1) = \frac{\mathrm{Li}_\alpha(1)}{\rho} < \infty \quad \text{and} \quad \nu = \max\left\{0, 1 - \frac{\mathrm{Li}_\alpha(1)}{\rho}\right\}$$

The study of the long cycles requires further assumptions on the derivative.

The study of the long cycles requires further assumptions on the derivative. We assume that there exist an $\alpha > 2, \alpha \notin \mathbb{N}$ such that for all $d \in \mathbb{N}$

$$\left(\frac{\partial}{\partial t}\right)^d p_\Upsilon(t) = \left(\left(\frac{\partial}{\partial t}\right)^d Li_\alpha(t)\right)(1 + o(1))$$

$$\left(\frac{\partial}{\partial t}\right)^d g_\Theta(t) = \left(\left(\frac{\partial}{\partial t}\right)^d \vartheta \log\left(\frac{1}{1-t}\right)\right)(1 + o(1))$$

We find here a new phenomenon:
the stick–breaking process with an unbreakable part.

We find here a new phenomenon:
the stick–breaking process with an unbreakable part.
Let a stick of length 1 be given, splitted into a breakable part of length ν and an unbreakable part $1-\nu$.

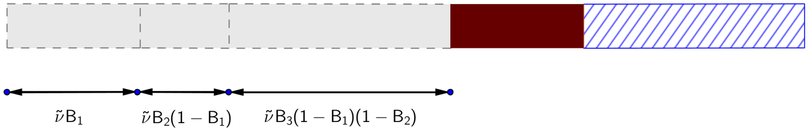
We find here a new phenomenon:
the stick–breaking process with an unbreakable part.
Let a stick of length 1 be given, splitted into a breakable part of length ν and an unbreakable part $1-\nu$.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1,\vartheta)$.

We find here a new phenomenon:
the stick–breaking process with an unbreakable part.
Let a stick of length 1 be given, splitted into a breakable part of length ν and an unbreakable part $1-\nu$.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1,\vartheta)$.

We find here a new phenomenon:
the stick–breaking process with an unbreakable part.
Let a stick of length 1 be given, splitted into a breakable part of length ν and an unbreakable part $1-\nu$.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1,\vartheta)$.

We find here a new phenomenon:
the stick–breaking process with an unbreakable part.
Let a stick of length 1 be given, splitted into a breakable part of length ν and an unbreakable part $1-\nu$.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1,\vartheta)$.

We find here a new phenomenon:
the stick–breaking process with an unbreakable part.
Let a stick of length 1 be given, splitted into a breakable part of length ν and an unbreakable part $1-\nu$.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1,\vartheta)$.



Theorem

Suppose that $\theta_k \to \vartheta > 0$ *and* $\nu > 0$. *Then*

$$\left(\frac{\lambda_1}{\nu n}, \frac{\lambda_2}{\nu n}, \dots\right) \xrightarrow{d} \mathcal{PD}(\vartheta)$$

with $\mathcal{PD}(\vartheta)$ *the Poisson–Dirichlet distribution with parameter* ϑ.

Theorem

Suppose that $\theta_k \to \vartheta > 0$ *and* $\nu > 0$. *Then*

$$\left(\frac{\lambda_1}{\nu n}, \frac{\lambda_2}{\nu n}, \dots\right) \xrightarrow{d} \mathcal{PD}(\vartheta)$$

with $\mathcal{PD}(\vartheta)$ *the Poisson–Dirichlet distribution with parameter* ϑ.

Theorem

Suppose that $\theta_k \equiv 0$ *and* $\nu > 0$. *Then*

$$\frac{\lambda_1}{\nu n} \xrightarrow{d} 1 \tag{4}$$

Thank you