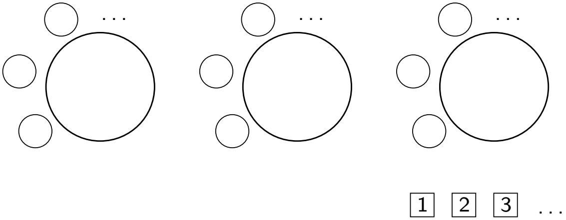
Up-down chains arising from the ordered Chinese Restaurant Process

Kelvin Rivera-Lopez
Joint with Douglas Rizzolo

Random Permutations Meet Random Matrices

April 8, 2022

The Chinese Restaurant Process



- Customer 1 Enters:
  - sits alone $\longrightarrow$ record $\{\{1\}\}$
- Customer 2 Enters:
  - $1 - \alpha \longrightarrow$ joins 1 $\longrightarrow$ record $\{\{1, 2\}\}$
  - $\alpha + \theta \longrightarrow$ sits alone $\longrightarrow$ record $\{\{1\}, \{2\}\}$

...

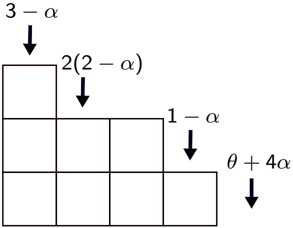
- Customer k Enters:
  - $m - \alpha \longrightarrow$ joins table with m customers $\longrightarrow$ add k to a set
  - $\alpha N + \theta \longrightarrow$ sits alone $\longrightarrow$ add the singleton $\{k\}$

The Chinese Restaurant Process

- Π_n: the n^{th} record
  - random set partition of $\{1, 2, \ldots, n\}$
- CRP: $(\Pi_n)_{n \geq 1}$
- *An interesting property:*
  - $\Pi_n \xrightarrow[\text{CRP}]{} \Pi_{n+1} \xrightarrow[\substack{\text{uniform}\\ \text{deletion}}]{} \Pi'_n \overset{d}{=} \Pi_n$
- CRP + uniform deletion
  - Markov chain on partitions of $\{1, \ldots, n\}$
  - law of Π_n is stationary
- Theme:
  - growth + decay models $\longleftrightarrow$ well-known stationary distributions
- Growth + decay model = *up-down chain*

Petrov's Chains

- Sizes: $\{A_1, \ldots, A_N\} \longmapsto (|A_1|, \ldots, |A_N|)_{\text{dec}}$
- Up-step: Sizes(Π_n) $\longrightarrow$ Sizes(Π_{n+1})
- Down-step: uniformly delete a box
- Results:
  - there is a scaling limit
  - the limit generalizes the IMNA* model
  - PD(α, θ)† is stationary



*Infinitely-Many-Neutral-Alleles
†Poisson-Dirichlet distribution

Ordered Analogues

The ordered CRP universe:

- oCRP
- ordered PD(α, θ)
- *Missing:* ordered variant of Petrov diffusions

Some work:

- Shi, Winkel
  - *Stationarity:* 2020
- Forman, Pal, Rizzolo, Winkel
  - *Stationarity:* 2017, 19, 19, 20
  - *Ordering:* in progress
  - *Convergence:* conjectured by Rogers, Winkel

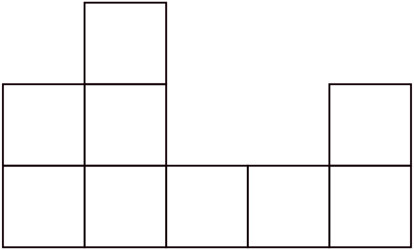
Our work:

- Rivera-Lopez, Rizzolo
  - *Stationarity, Ordering, Convergence:* 2020

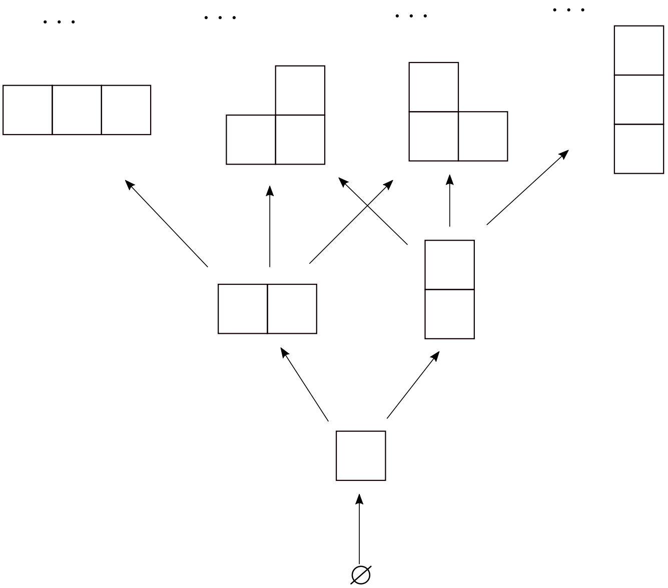
Integer Compositions

- *Composition* of n
  - $n \geq 1$: $\sigma = (\sigma_1, ..., \sigma_k)$
    - σ_i are positive integers
    - $\sum \sigma_i = n$
  - $n = 0$: $\varnothing = ()$
- $\mathcal{C}_n = \{\text{compositions of } n\}$.
- $\mathcal{C} = \cup \mathcal{C}_n$
- Correspond to diagrams

$$\sigma = (2, 3, 1, 1, 2) \quad \longmapsto$$

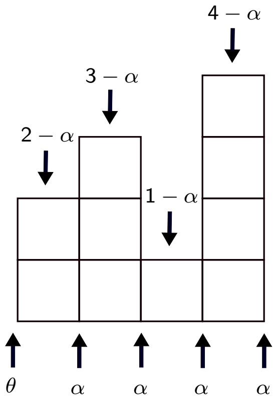


The Graph of Compositions: $\mathcal{C}$



The Up-Down Chains: $\{X_n\}_{n\geq 1}$

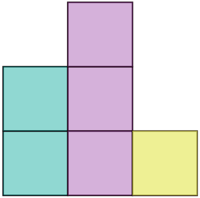
- State Space: $\mathcal{C}_n$
- Transition Matrix: $T_n = U_n D_n$
  - Up-step (U_n):



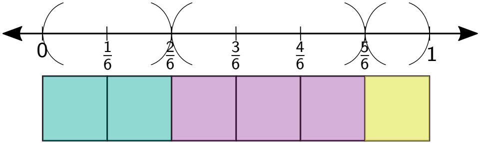
  - Down-step (D_n): uniform deletion

The Ambient Space: $\mathcal{U}$

- $\mathcal{U} = \{\text{open subsets of } (0,1)\}$
- Metric:
  - $d(U,V) = \text{Hausdorff Distance}(U^c, V^c)$
- Inclusion:
  - $\iota : \mathcal{C} \to \mathcal{U}$



$\mapsto$



The Main Result

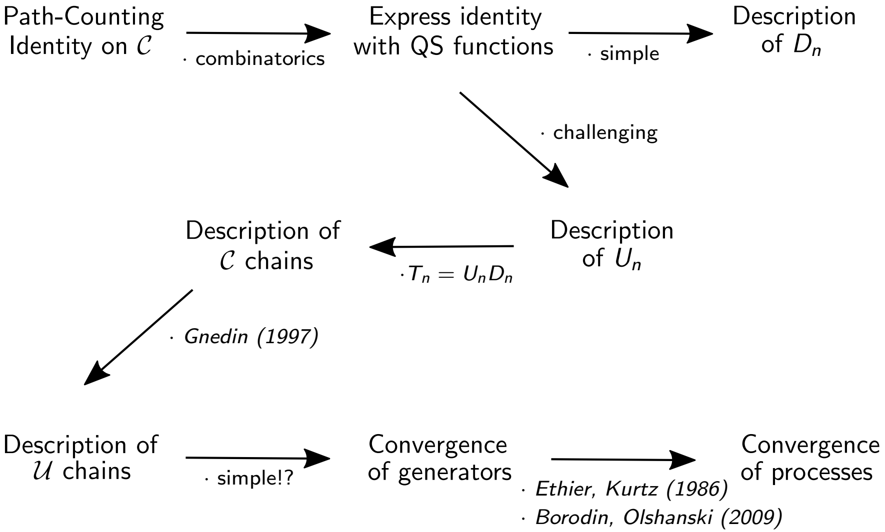
There exists a Feller diffusion $(X(t))_{t\geq 0}$ on $\mathcal{U}$ such that

$$\iota\Big(X_n\big(\lfloor n^2 t\rfloor\big)\Big) \longrightarrow X(t)$$

whenever the initial distributions converge.

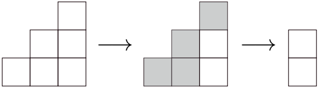
(convergence of paths in distribution on Skorokhod space)

The Approach

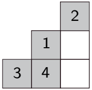


The Path-Counting Identity

- $g(\sigma, \tau) = \#$ paths from σ to τ
- path $\longleftrightarrow$ deconstruction
- Decomposition of a path:
  1. choose which boxes to remove

     - $b = (1, 2, 1)$
  2. choose when to remove each box

     - $\pi = (\{3\}, \{1, 4\}, \{2\})$
- Result:

$$g(\sigma, \tau) = \sum_{b} \frac{(|\tau| - |\sigma|)!}{\prod b_r!}.$$

The Algebra of Quasisymmetric Functions: Λ

- Quasisymmetric function:
  - Formal power series in $y_1, y_2, \ldots$
  - Bounded degree
  - Coefficient of $y_1^{a_1} \ldots y_k^{a_k}$ = Coefficient of $y_{i_1}^{a_1} \ldots y_{i_k}^{a_k}$ for i monotone
- For example,

$$\begin{aligned} y_1^4 y_2 + y_1^4 y_3 + y_1^4 y_4 + \ldots \quad &= \sum_{i_1 < i_2} y_{i_1}^4 y_{i_2} \\ + y_2^4 y_3 + y_2^4 y_4 + \ldots + y_7^4 y_9 + \ldots \quad & \end{aligned}$$

- Basis of *monomials*:

$$m_\sigma = \sum_{i_1 < \ldots < i_k} \prod_{r=1}^{k} y_{i_r}^{\sigma_r}, \quad \sigma = (\sigma_1, \ldots, \sigma_k) \in \mathcal{C}$$

- Basis of *monomial-variants*:

$$m_\sigma^* = \sum_{i_1 < \ldots < i_k} \prod_{r=1}^{k} y_{i_r}^{\downarrow \sigma_r}, \quad \sigma \in \mathcal{C}$$

where $a^{\downarrow b} = a(a-1) \cdot \ldots \cdot (a - (b-1))$

The QS Path-Counting Identity

Path-Counting Identity

$+$

some combinatorics

$$\Downarrow$$

$$m_\sigma^*(\tau) = g(\sigma, \tau) \frac{|\tau|^{\downarrow|\sigma|}}{g(\varnothing, \tau)}$$

where

$$m_\sigma^*(\tau) = \sum_{i_1 < \ldots < i_k \leq \ell} \prod_{r=1}^{k} \tau_{i_r}^{\downarrow \sigma_r}$$

$(\ell = \#$ components in $\tau)$

The Down Operators

QS Path-Counting Identity

$+$

some more combinatorics

$$\Downarrow$$

$$\underset{\in C(\mathcal{C}_n)}{\underset{\uparrow}{D_n m^*_\sigma}} = \text{constant} \times \underset{\in C(\mathcal{C}_{n+1})}{\underset{\uparrow}{m^*_\sigma}}$$

The Transition Operators (in $\mathcal{C}$)

explicit formulas for D_n, U_n

$$+$$

$$T_n = U_n D_n$$

$$\Downarrow$$

$$T_n\, m^*_\sigma = a_\sigma m^*_\sigma + \sum_{\mu \nearrow \sigma} b_{\mu,\sigma} m^*_\mu$$

- triangular

The Transition Operators (in $\mathcal{U}$)

Transition operator: $\iota(T_n)$

- $\iota(T_n)$: $C(\mathcal{U}) \to C(\mathcal{U})$

Gnedin (1997)

- Each $q \in \Lambda$ has a natural identification q^o in $C(\mathcal{U})$

Result:

$$T_n\, m^*_\sigma = a_\sigma m^*_\sigma + \sum_{\mu < \sigma} b_{\mu,\sigma} m^*_\mu$$

$$\Downarrow$$

$$\iota(T_n)\,(m^*_\sigma)^o = a_\sigma (m^*_\sigma)^o + \sum_{\mu < \sigma} b_{\mu,\sigma} (m^*_\mu)^o$$

The Convergence Argument

$$m_\sigma^* = m_\sigma + \text{lower order terms}$$

$$\Downarrow$$

$$(m_\sigma^*)^o = m_\sigma^o + o(1)$$

A summary

- CRP
  - up-down chains are natural
- Petrov + ordered CRP universe
  - left a gap
    - there has been partial progress
    - we complete the picture
- Approach
  - algebraic + combinatorial
    - inspired by Petrov and others
  - $g(\sigma,\tau) \approx m^*_\sigma$
  - T_n are triangular (m^*_σ)
  - $\iota(T_n)$ is also triangular ($(m^*_\sigma)^o$)
  - $m^*_\sigma = m_\sigma +$ lower order terms

Thank you!

Questions?