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
Customer 1 Enters:
  - sits alone $\rightarrow$ record $\{1\}$

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

... 

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

- $\Pi_n$: the $n^{th}$ record
  - random set partition of $\{1, 2, \ldots, n\}$

- CRP: $(\Pi_n)_{n \geq 1}$

- An interesting property:
  $$\Pi_n \xrightarrow{\text{CRP}} \Pi_{n+1} \xrightarrow{\text{uniform deletion}} \Pi'_n \overset{d}{=} \Pi_n$$

- CRP + uniform deletion
  - Markov chain on partitions of $\{1, \ldots, n\}$
  - law of $\Pi_n$ is stationary

- Theme:
  - growth + decay models $\leftrightarrow$ well-known stationary distributions

- Growth + decay model = up-down chain
Petrov’s Chains

- Sizes: \( \{A_1, \ldots, A_N\} \mapsto (|A_1|, \ldots, |A_N|)_{\text{dec}} \)
- Up-step: \( \text{Sizes}(\Pi_n) \mapsto \text{Sizes}(\Pi_{n+1}) \)
- Down-step: uniformly delete a box
- Results:
  - there is a scaling limit
  - the limit generalizes the IMNA\(^*\) model
  - \( \text{PD}(\alpha, \theta) \)\(^\dagger\) is stationary

---

\(^*\)Infinitely-Many-Neutral-Alleles
\(^\dagger\)Poisson-Dirichlet distribution
Ordered Analogues

The ordered CRP universe:

▶ oCRP
▶ ordered PD($\alpha, \theta$)
▶ *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 $n$**
  - $n \geq 1$: $\sigma = (\sigma_1, \ldots, \sigma_k)$
    - $\sigma_i$ are positive integers
    - $\sum \sigma_i = n$
  - $n = 0$: $\emptyset = ()$

- $C_n = \{\text{compositions of } n\}$.

- $C = \bigcup C_n$

- Correspond to diagrams

\[\sigma = (2, 3, 1, 1, 2)\]
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** \((U_n)\):

\[
\begin{align*}
\text{Up-step: } U_n &= \theta - \alpha - \alpha - \alpha - \alpha \\
\text{Down-step: } D_n &= \text{uniform deletion}
\end{align*}
\]
The Ambient Space: $\mathcal{U}$

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

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

\[
\iota \left( X_n \left( \lfloor n^2 t \rfloor \right) \right) \rightarrow X(t)
\]

whenever the initial distributions converge.

(convergence of paths in distribution on Skorokhod space)
The Approach

Path-Counting Identity on $\mathcal{C}$ → Express identity with QS functions → Description of $\mathcal{D}_n$

- combinatorics
- simple
- challenging

Description of $\mathcal{C}$ chains

- $T_n = U_n D_n$
- Gnedin (1997)

Description of $\mathcal{U}$ chains

- simple!

Convergence of generators

- Ethier, Kurtz (1986)
- Borodin, Olshanski (2009)

Convergence of processes
The Path-Counting Identity

\[ g(\sigma, \tau) = \# \text{ paths from } \sigma \text{ to } \tau \]

\[ \text{path } \leftrightarrow \text{ deconstruction} \]

\[ \text{Decomposition of a path:} \]

1. choose which boxes to remove

\[ b = (1, 2, 1) \]

2. choose when to remove each box

\[ \pi = (\{3\}, \{1, 4\}, \{2\}) \]

\[ \text{Result:} \]

\[ g(\sigma, \tau) = \sum_b \frac{(|\tau| - |\sigma|)!}{\prod b_r!} \]
The Algebra of Quasisymmetric Functions: $\Lambda$

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

- **For example,**

$$y_1^4y_2 + y_1^4y_3 + y_1^4y_4 + \ldots = \sum_{i_1 < i_2} y_{i_1}^4 y_{i_2}$$

$$+ y_2^4y_3 + y_2^4y_4 + \ldots + y_7^4y_9 + \ldots$$

- **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

\[ \Downarrow \]

\[ m^*_\sigma(\tau) = g(\sigma, \tau) \frac{|\tau| \downarrow |\sigma|}{g(\emptyset, \tau)} \]

where

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

\[ (\ell = \# \text{ components in } \tau) \]
The Down Operators

QS Path-Counting Identity

\[ D_n m^*_\sigma = \text{constant} \times m^*_\sigma \]

\[ \in C(C_n) \]

\[ \in C(C_{n+1}) \]
explicit formulas for $D_n, U_n$

\[ T_n = U_n D_n \]

\[ T_n m^*_\sigma = a_\sigma m^*_\sigma + \sum_{\mu \to \sigma} b_{\mu,\sigma} m^*_\mu \]

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

Transition operator: $\iota(T_n)$

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

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

\[ m^*_\sigma = m_\sigma + \text{lower order terms} \]

\[ (m^*_\sigma)^o = m^o_\sigma + o(1) \]
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 ($m_\sigma^*$)
  - $\nu(T_n)$ is also triangular ($(m_\sigma^*)^o$)
  - $m_\sigma^* = m_\sigma + \text{lower order terms}$
Thank you!

Questions?