Integrable probability:
Beyond the Gaussian universality class

Ivan Corwin

(Columbia University, Clay Mathematics Institute, Institute Henri Poincare)

Integrable probability

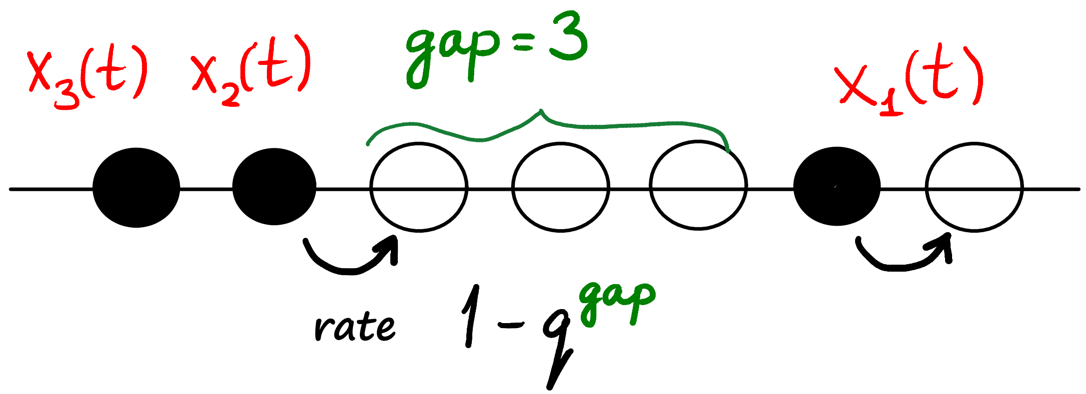
An **integrable probabilistic system** has two main properties:

1. It is possible to compute concise formulas for averages of a rich class of observables.
2. Taking limits of the system, observables and formulas, it is possible to access detailed descriptions of universal classes.

I will focus on a few examples in the Kardar-Parisi-Zhang class and describe how their integrability springs from connections to representation theory and quantum integrable systems.

Basic reason that these models turn out to be analyzable is the existence of a large family of observables whose averages are explicit.

Example 1: q-TASEP [Borodin-C, 2011]

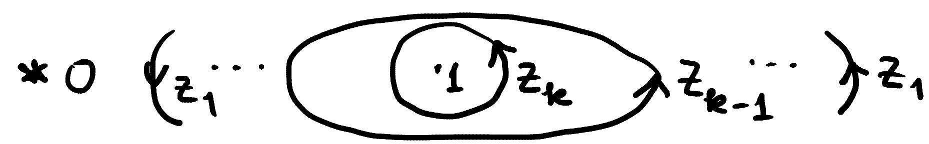


$0 < q < 1$

Theorem [B-C'11], [B-C-Sasamoto'12] For step initial data $\{x_n(0) = -n\}_{n\geq 1}$

$$\mathbb{E}\left[q^{(x_{N_1}(t)+N_1)+\dots+(x_{N_k}(t)+N_k)}\right] = \frac{(-1)^k q^{\frac{k(k-1)}{2}}}{(2\pi i)^k}\oint\cdots\oint \prod_{A<B}\frac{z_A-z_B}{z_A-qz_B}\prod_{j=1}^{k}\frac{e^{(q-1)tz_j}}{(1-z_j)^{N_j}}\frac{dz_j}{z_j}$$

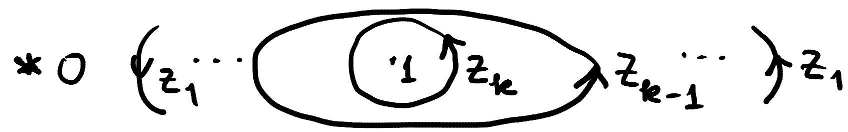
$(N_1 \geq N_2 \geq \dots \geq N_k)$



Let us briefly explain why such formulas are useful for asymptotics.

For q-TASEP with step initial data, one specializes to q-moments

$$\mathbb{E}\left(q^{x_N(t)+N}\right)^k = \frac{(-1)^k q^{\frac{k(k-1)}{2}}}{(2\pi i)^k} \oint \cdots \oint \prod_{A<B} \frac{z_A - z_B}{z_A - q z_B} \prod_{j=1}^{k} \frac{e^{(q-1)t z_j}}{(1-z_j)^N} \frac{dz_j}{z_j}$$



and takes their generating function (q-Laplace transform)

$$\mathbb{E} \frac{1}{\prod_{m\geq 0}\left(1-\zeta q^m q^{x_N+N}\right)} = \mathbb{E} \sum_{k=0}^{\infty} \frac{\left(q^{x_N+N}\right)^k \zeta^k}{(1-q)\cdots(1-q^k)} = \det(\mathbb{1}+K)_{L^2(\mathbb{N}\times \circlearrowleft)}$$

with

$$K(n_1, w_1; n_2, w_2) = \frac{f(w_1)\cdots f(q^{n_1-1}w_1)\,\zeta^{n_1}}{q^{n_1}w_1 - w_2}, \qquad f(w) = \frac{e^{(q-1)\tau w}}{(1-w)^N}$$

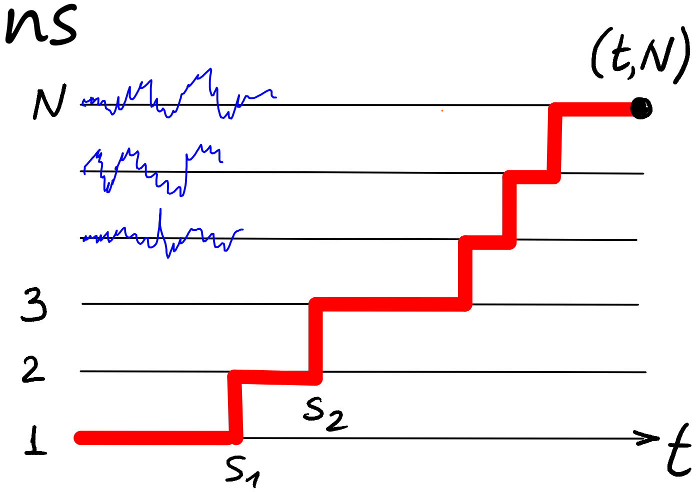
The result is suitable for taking various limits.

Example 2: semi-discrete Brownian polymer [O'Connell-Yor, 2001]

Taking a suitable scaling limit of the q-TASEP as $q \to 1$,

one arrives at the following partition functions

$$Z^N(t) = \int_{0<s_1<\dots<s_{N-1}<t} e^{B_1(0,s_1)+B_2(s_1,s_2)+\dots+B_N(s_{N-1},t)}\, ds_1 \dots ds_{N-1}$$



Brownian increments $B_k(\alpha,\beta) := B_k(\beta) - B_k(\alpha)$

Theorem [B-C, B-C-Ferrari, 2011-12] Set $F^N(t) = \log Z^N(t)$

$$\lim_{N\to\infty} \mathbb{P}\left\{ \frac{F^N(\varkappa N) - N f_\varkappa}{N^{1/3}} \le r \right\} = F_{GUE}\left(\left(\frac{g_\varkappa}{2}\right)^{-1/3} r \right)$$

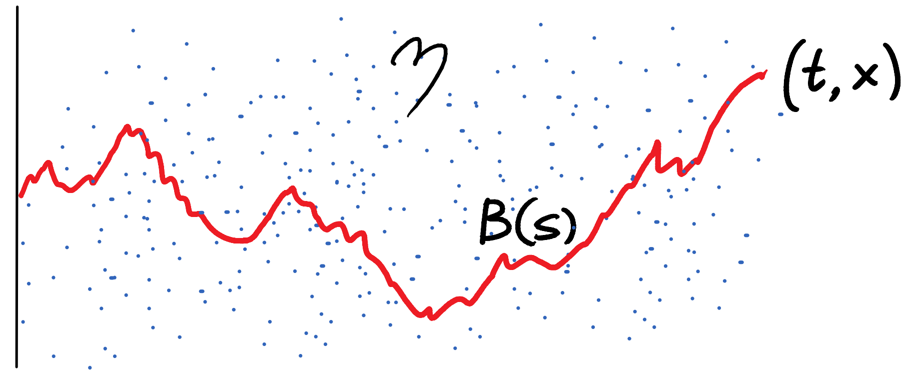
GUE Tracy-Widom distribution describes the largest eigenvalue of a large Gaussian Hermitian matrix

Here one cannot use the moment expansion of the Laplace transform because it is divergent!
Leads to the (non-rigorous) replica trick. Using q-TASEP provides a rigorous replica trick.

Example 3: SHE/KPZ equation/continuum polymer

A weak-noise limit yields the continuum polymer partition function

$$\mathcal{Z}(t,x) = \int_{\text{Brownian paths } B(0)=0,\ B(t)=x} e^{\int_0^t \eta(s, B(s))\,ds}$$



Equivalently, the stochastic heat equation: $\mathcal{Z}_t = \frac{1}{2}\Delta \mathcal{Z} + \eta \cdot \mathcal{Z}$

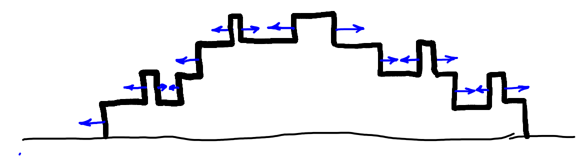
or setting $h := \log(\mathcal{Z})$, the KPZ equation: $h_t = \frac{1}{2}\Delta h + \frac{1}{2}(\nabla h)^2 + \eta$

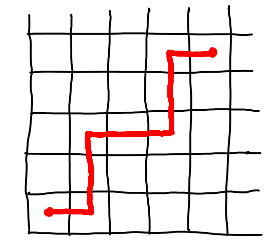
$t^{1/3}$ fluctuations, KPZ class statistics shown in works of

rigorous: [Amir-C-Quastel, '10], [C-Quastel '11], [B-C-Ferrari, '12], [B-C-Ferrari-Veto '14]

non-rigorous: [Sasamoto-Spohn, '10], [Dotsenko, '10+], [Calabrese-Le Doussal-Rosso, '10+], [Sasamoto-Imamura, '11]

KPZ class integrable probabilistic systems





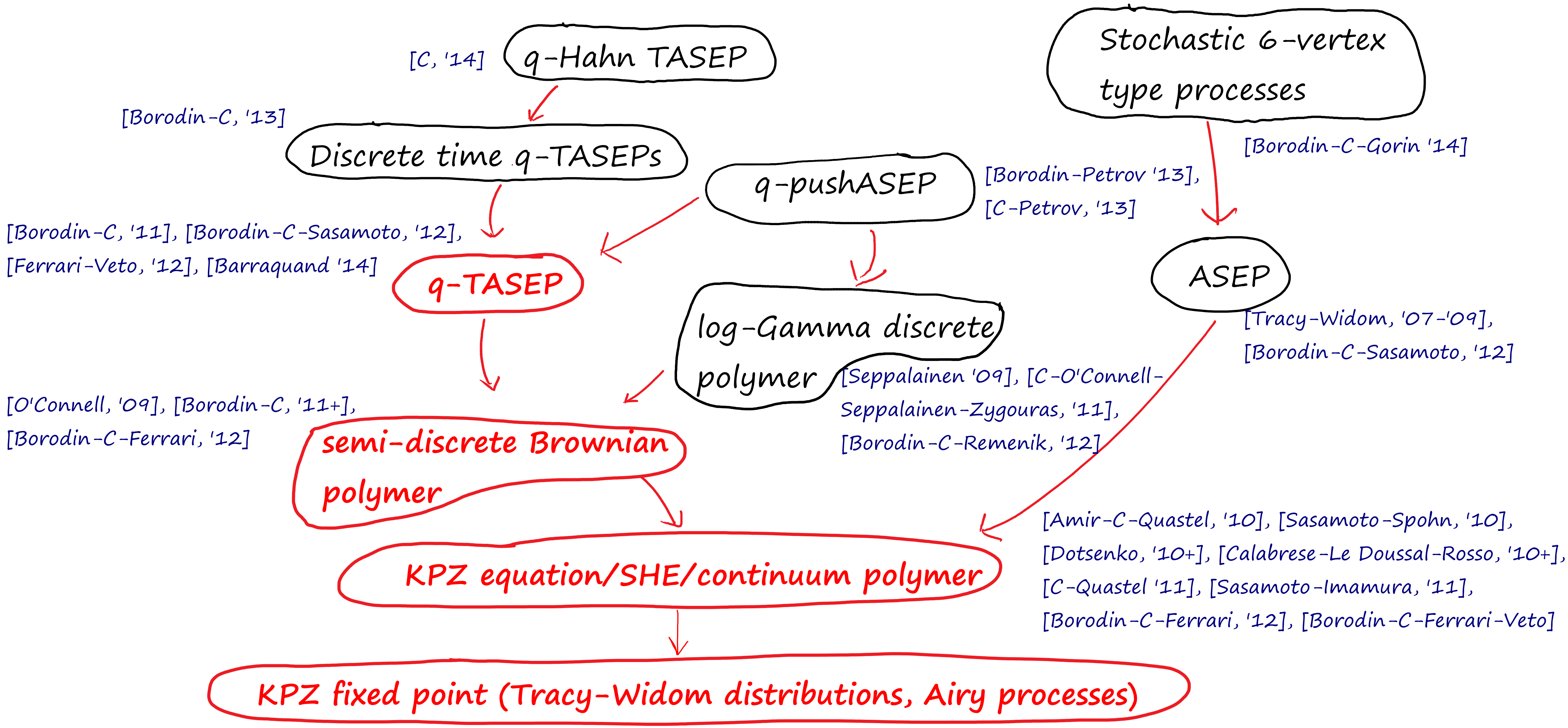
Late '90s: Asymptotic analysis of some totally asymmetric / zero temperature models in KPZ class (PNG/LIS, TASEP/LPP) [Baik-Deift-Johansson '99] [Johansson '99].

Early '00s: Determinantal point process framework developed (Schur processes, free Fermions, non-intersecting paths).

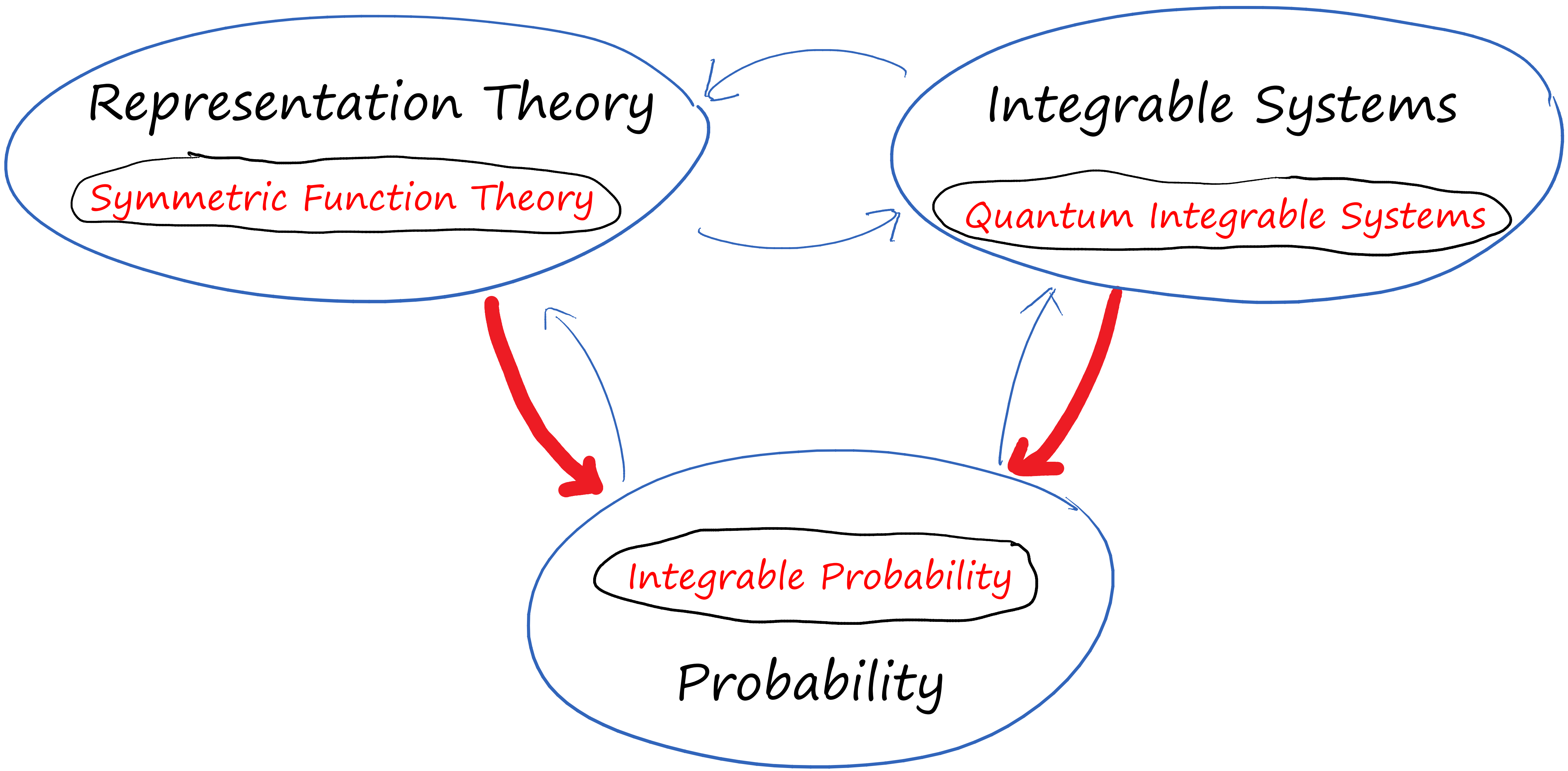
Late '00s-present: Beginning with [Tracy-Widom '07-'09] some non-determinantal, partially asymmetric / positive temperature models have been discovered and studied asymptotically.

Developed in parallel: Probabilistic means to study KPZ exponents.

Some non-determinantal models which have been analyzed



Structures behind KPZ class integrable probabilistic systems



Macdonald polynomials $P_\lambda(x_1,\dots,x_N)\in\mathbb{Q}(q,t)[x_1,\dots,x_N]^{S(N)}$ with partitions $\lambda=(\lambda_1\geq\lambda_2\geq\dots\geq\lambda_N\geq 0)$ form a basis in symmetric polynomials in N variables over $\mathbb{Q}(q,t)$. They diagonalize

$$(\mathcal{D}_1 f)(x_1,\dots,x_N)=\sum_{i=1}^{n}\prod_{j\neq i}\frac{t\,x_i-x_j}{x_i-x_j}\,f(x_1,\dots,qx_i,\dots,x_N)$$

with (generically) pairwise different eigenvalues

$$\mathcal{D}_1 P_\lambda=(q^{\lambda_1}t^{N-1}+q^{\lambda_2}t^{N-2}+\dots+q^{\lambda_N})P_\lambda\,.$$

They have many remarkable properties that include orthogonality (dual basis Q_λ), simple reproducing kernel (Cauchy type identity), Pieri and branching rules, index/variable duality, explicit generators of the algebra of (Macdonald) operators commuting with $\mathcal{D}_1$, etc.

(Ascending) Macdonald processes are probability measures on interlacing triangular arrays (Gelfand-Tsetlin patterns)

$$
\begin{array}{ccccccccc}
\lambda^{(N)}_N & & \lambda^{(N)}_{N-1} & & \cdots & & \lambda^{(N)}_2 & & \lambda^{(N)}_1 \\
& \lambda^{(N-1)}_{N-1} & & \lambda^{(N-1)}_{N-2} & & \cdots & & \lambda^{(N-1)}_1 & \\
& & \ddots & & \cdots & & \iddots & & \\
& & & \lambda^{(2)}_2 & & \lambda^{(2)}_1 & & & \\
& & & & \lambda^{(1)}_1 & & & &
\end{array}
$$

$$\lambda^{(m)}_j \in \mathbb{Z}_{\geq 0}$$

Macdonald polynomials

$$\mathbb{P}\left(\lambda^{(k)}\right) = \frac{P_{\lambda^{(k)}}(a_1, \ldots, a_k)\, Q_{\lambda^{(k)}}(b_1, \ldots, b_M)}{\Pi(\underbrace{a_1, \ldots, a_k}; \underbrace{b_1, \ldots, b_M})}$$

normalization constant

two groups of parameters

Macdonald processes

Integrable structure of Macdonald polynomials translates into probabilistic content enabling us to:

1. Describe nice (2+1) dimensional Markov chains which preserve the class of measures and have interesting marginals

2. Compute formulas for averages of a rich class of observables

Initiated in [Borodin-C '11+]; many other developments involving Bufetov, Ferrari, Gorin, O'Connell, Pei, Petrov, Remenik, Seppalainen, Shakirov, Shkolnikov, Veto, Zygouras...

Macdonald processes $q, t \in [0,1)$

Ruijsenaars–Macdonald system

Representations of Double Affine Hecke Algebras

q-Whittaker processes $t=0$

q-TASEP, 2d dynamics

q-deformed quantum Toda lattice

Representations of $\hat{\mathfrak{gl}}_N$, $U_q(\mathfrak{gl}_N)$

Hall–Littlewood processes $q=0$

Random matrices over finite fields

Spherical functions for p-adic groups

General β RMT $t = q^{\beta/2} \to 1$

Random matrices over $\mathbb{R}, \mathbb{C}, \mathbb{H}$

Calogero–Sutherland, Jack polynomials

Spherical functions for Riem. Symm. Sp.

Whittaker processes $t=0$, $q \to 1$

Directed polymers and their hierarchies

Quantum Toda lattice, repr. of $GL(n,\mathbb{R})$

Kingman partition structures $q=0$, $t=1$

Cycles of random permutations

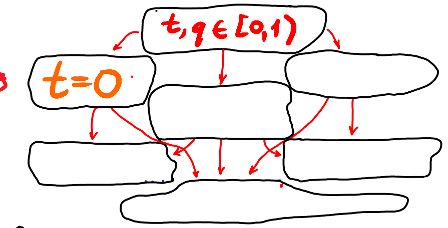
Poisson–Dirichlet distributions

Schur processes $q=t$

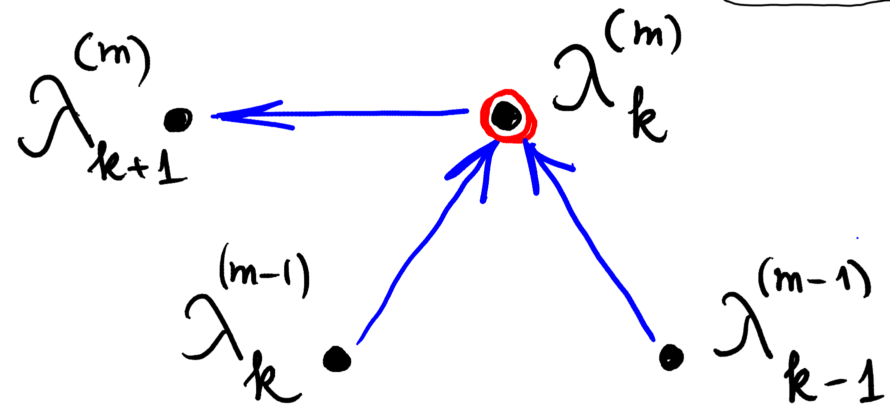
Plane partitions, tilings/shuffling, TASEP, PNG, last passage percolation, GUE

Characters of symmetric, unitary groups

Markov process preserving $t=0$ Macdonald process



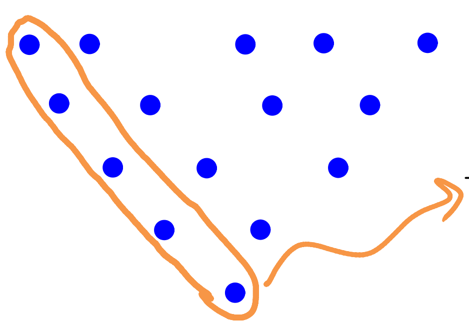
Each coordinate of the triangular array jumps by 1 to the right independently of the others with

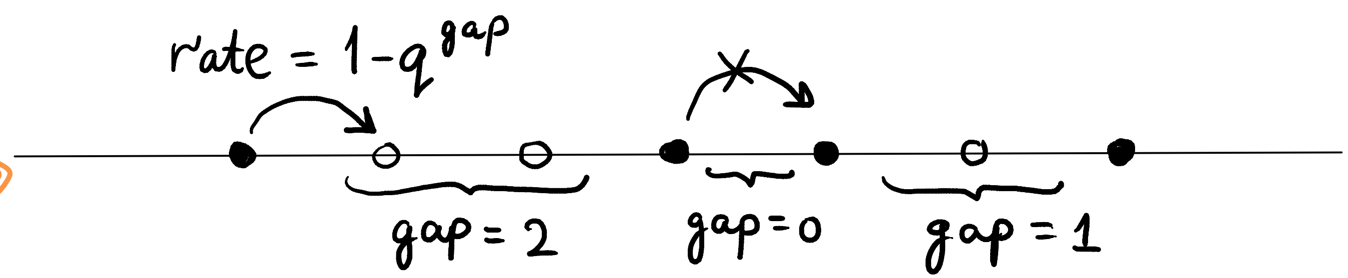


$$\mathrm{rate}\left(\lambda^{(m)}_k\right)=\frac{\left(1-q^{\lambda^{(m-1)}_{k-1}-\lambda^{(m)}_k}\right)\left(1-q^{\lambda^{(m)}_k-\lambda^{(m)}_{k+1}+1}\right)}{\left(1-q^{\lambda^{(m)}_k-\lambda^{(m-1)}_k}\right)}$$

Other dynamics preserve Macdonald process
[O'Connell-Pei '12], [Borodin-Petrov '13]

The set of coordinates $\left\{\lambda^{(m)}_m-m\right\}_{m\geq 1}$ forms q-TASEP





Evaluation of Macdonald process averages

Take an operator diagonal in Macdonald poly's: $\mathcal{D} P_\lambda = d_\lambda P_\lambda$.

$$(\mathcal{D}f)(x_1,\dots,x_N) = \sum_{i=1}^{n} \prod_{j\neq i} \frac{t x_i - x_j}{x_i - x_j} f(x_1,\dots,q x_i,\dots,x_N), \quad \mathcal{D} P_\lambda = (q^{\lambda_1} t^{N-1} + q^{\lambda_2} t^{N-2} + \dots + q^{\lambda_N}) P_\lambda.$$

Apply it to the Cauchy type identity $\sum_\lambda P_\lambda(a) Q_\lambda(b) = \Pi(a;b)$

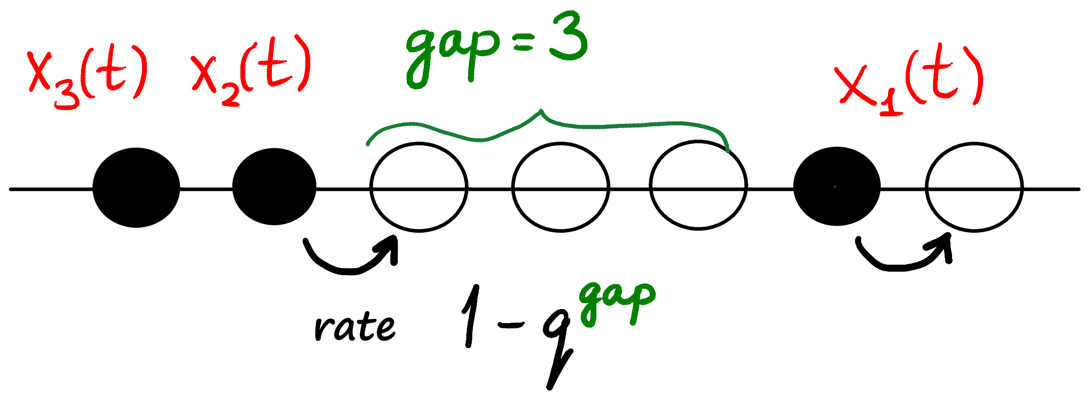
to obtain
$$\mathbb{E}[d_\lambda] = \frac{\mathcal{D}^{(a)} \Pi(a;b)}{\Pi(a;b)}.$$

Since all ingredients are explicit, we obtain meaningful probabilistic information without explicit formulas for Macdonald polynomials.

When $t=0$, $d_\lambda = q^{\lambda_N}$ and this yields the q-TASEP moment formulas!

Basic reason that these models turned out to be analyzable is the existence of a large family of observables whose averages are explicit.

Example 1: q-TASEP [Borodin-C, 2011]

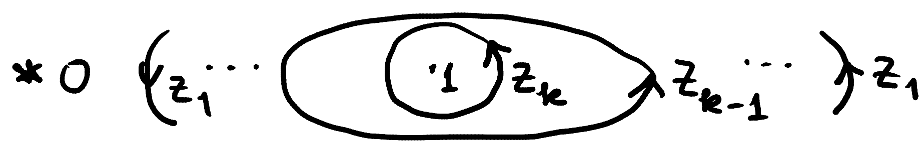


$0 < q < 1$

Theorem [B-C'11], [B-C-Sasamoto'12] *For step initial data* $\{x_n(0)=-n\}_{n\geq 1}$

$$\mathbb{E}\left[q^{(x_{N_1}(t)+N_1)+\dots+(x_{N_k}(t)+N_k)}\right] = \frac{(-1)^k q^{\frac{k(k-1)}{2}}}{(2\pi i)^k}\oint\cdots\oint \prod_{A<B}\frac{z_A-z_B}{z_A-qz_B}\prod_{j=1}^{k}\frac{e^{(q-1)tz_j}}{(1-z_j)^{N_j}}\frac{dz_j}{z_j}$$

$(N_1 \geq N_2 \geq \dots \geq N_k)$



Quantum integrable systems approach

It is not hard to check that the q-TASEP moments

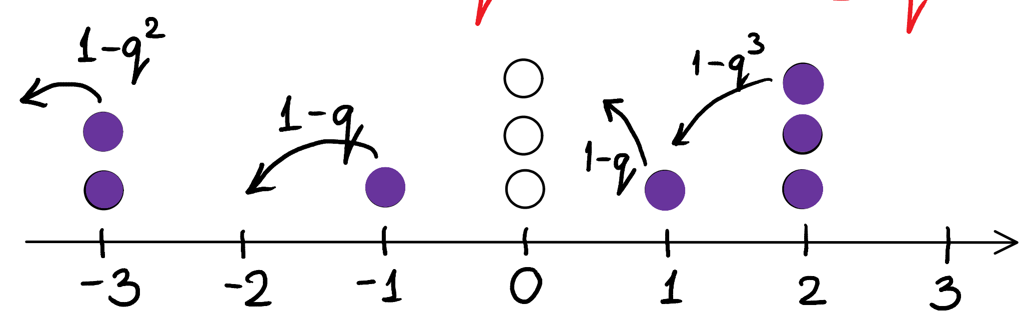
$$V(t;N_1,\dots,N_k)=\mathbb{E}\left[q^{(x_{N_1}(t)+N_1)+\dots+(x_{N_k}(t)+N_k)}\right]$$

satisfy [B-C-Sasamoto '12] the q-Boson system [Sasamoto-Wadati '97]

$$\partial_t V(t;\vec{N}) = \sum_{\text{clusters } i}(1-q^{c_i})\left[V(t;\vec{N}^{\,-}_{c_1+\dots+c_i}) - V(t;\vec{N})\right]$$

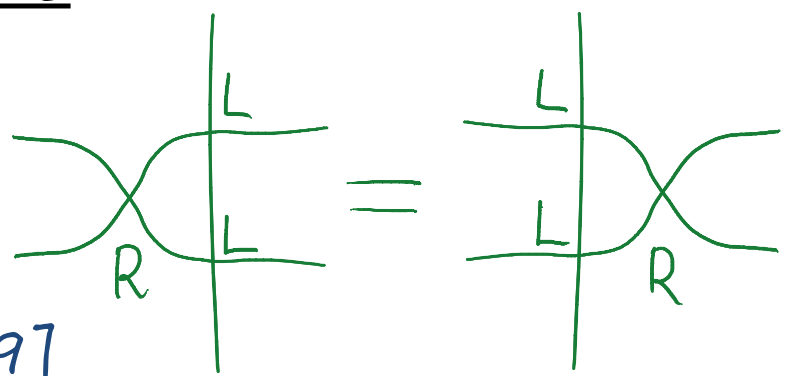
It is easy to verify that the contour integral formulas satisfy this (closed) system with desired initial data, thus proving them in an elementary way.

rate $= 1-q^{\text{site occupancy}} = 1-q^3$



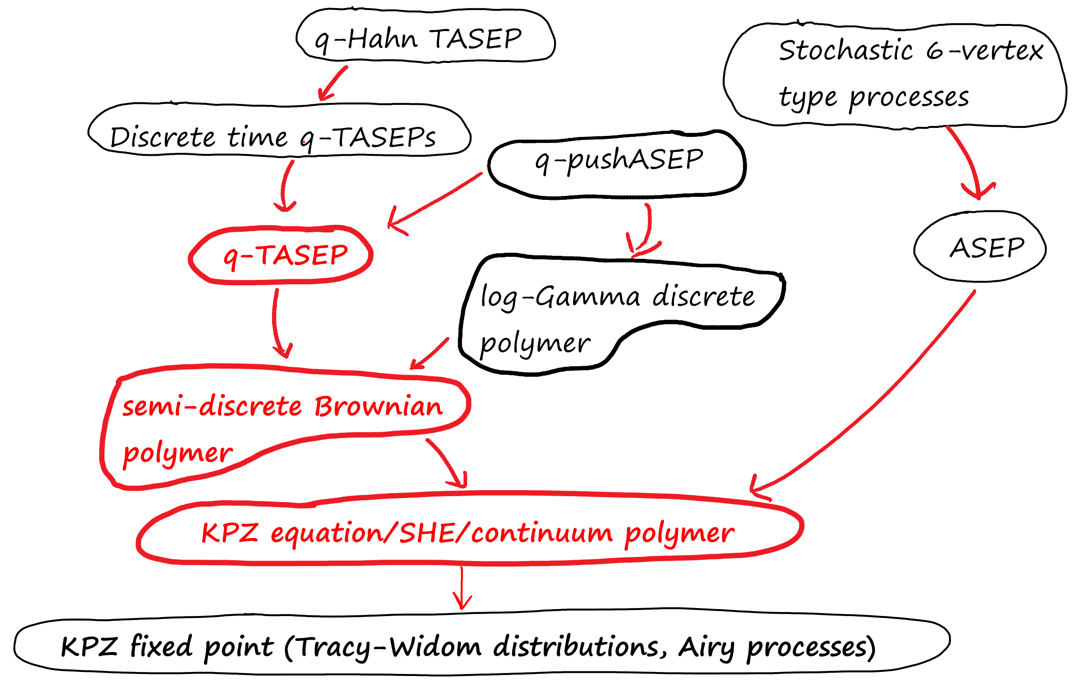
(Local) stochastic quantum integrable systems

This approach can be made less ad hoc:



- QISM / algebraic Bethe ansatz [Faddeev '79] is a rich source of exactly solvable 2d lattice/vertex models and (1+1)-dimensional quantum spin chains [Baxter '82].
- For 6-vertex/XXZ type systems there exist commuting stochastic transfer matrices which produce interacting particle systems (directly or via Markov dualities) [Borodin-C-Gorin '14].
- Completeness: Plancherel theory for Bethe ansatz yields moment formulas for general initial data [Borodin-C-Petrov-Sasamoto '13].

What structures degenerate?



Under the limit to the SHE,

$$Z_t = \frac{1}{2}\Delta Z + \eta \cdot Z \quad , Z(0,x) = \delta(x),$$

$$\bar{Z}(t; x_1, \dots, x_k) = \mathbb{E}\left[Z(t,x_1) \cdots Z(t,x_k) \right]$$

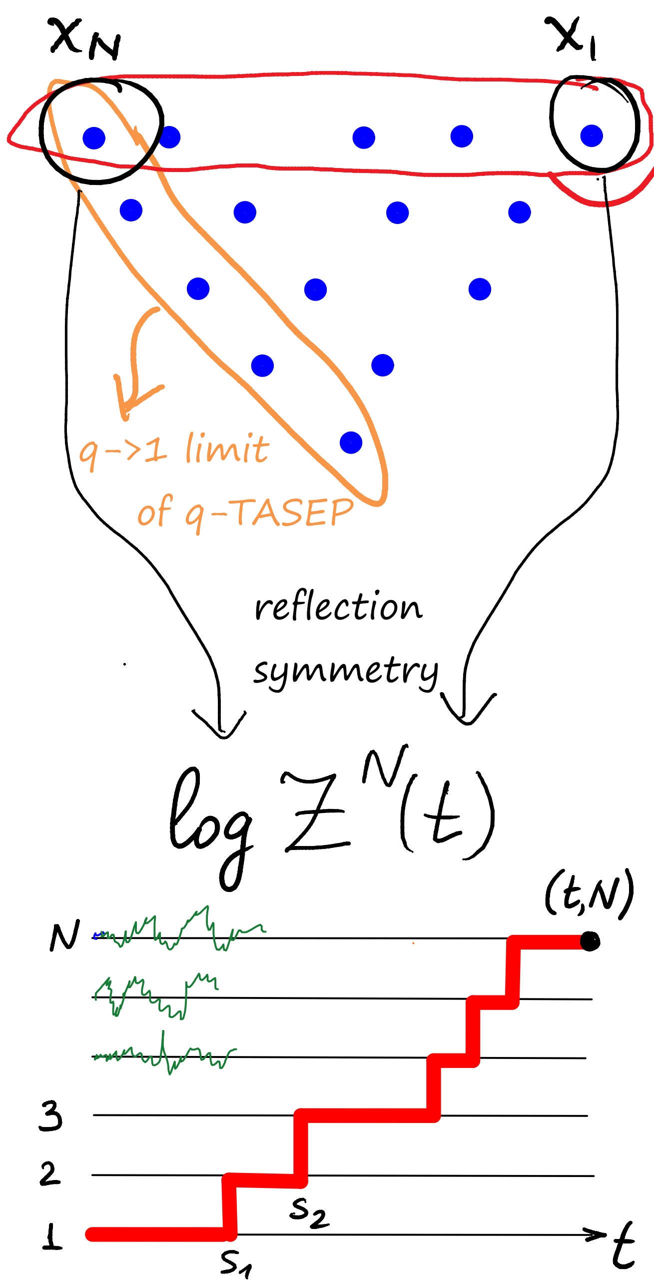
satisfies the delta-Bose gas evolution [Molchanov '86], [Kardar '87]

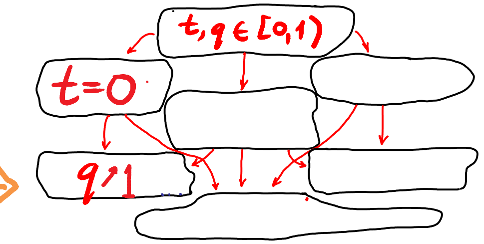
$$\partial_t \bar{Z} = \frac{1}{2}\left(\Delta + \sum_{i \neq j} \delta(x_i - x_j) \right) \bar{Z}$$

which is solvable via Bethe ansatz [Lieb-Liniger '63, Bethe '31]

(completeness proved in [Oxford '79, Heckman-Opdam '97]).

(Non-local) stochastic quantum integrable system





Top row evolves via **quantum Toda** diffusion, with generator [O'Connell '09]

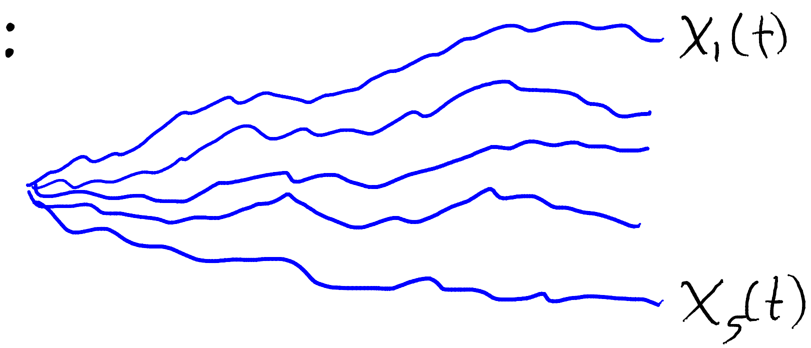
$$\psi_0^{-1}\left(\frac{1}{2}\Delta+\sum_{i=1}^{N-1} e^{x_{i+1}-x_i}\right)\psi_0$$

Class one GL(N) Whittaker function

which is a soft / positive temperature version of Dyson's Brownian motion:

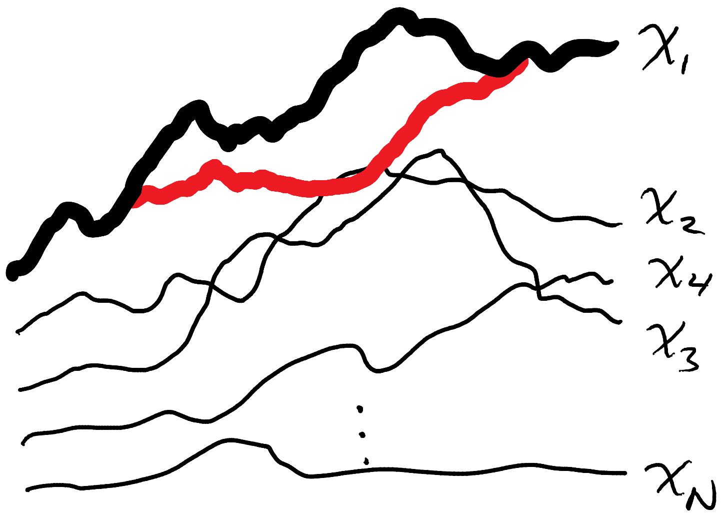
$$V^{-1}\Delta_{\text{Dirichlet}}V$$

Vandermonde determinant



Geometric lifting of RSK correspondence gives path interpretation for all $x_i(t)$ [O'Connell '09], [C-O'Connell-Seppalainen-Zygouras '13]

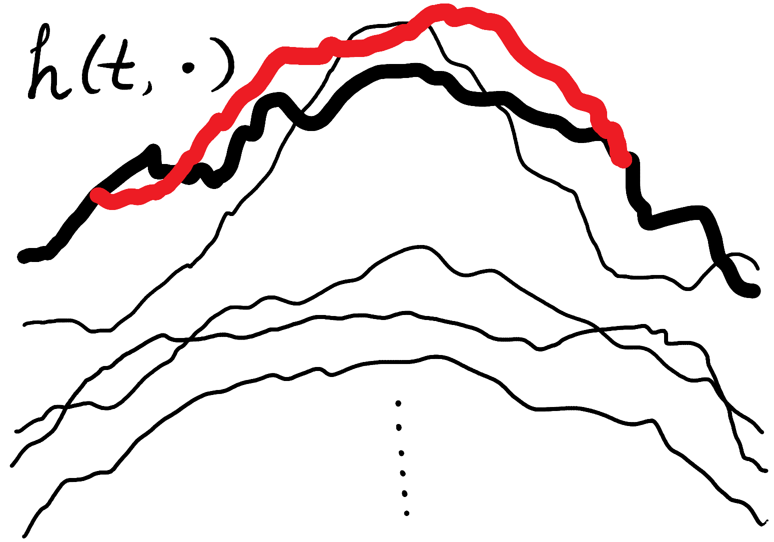
Limiting structure: KPZ/Airy line ensembles [C-Hammond '11,'13]



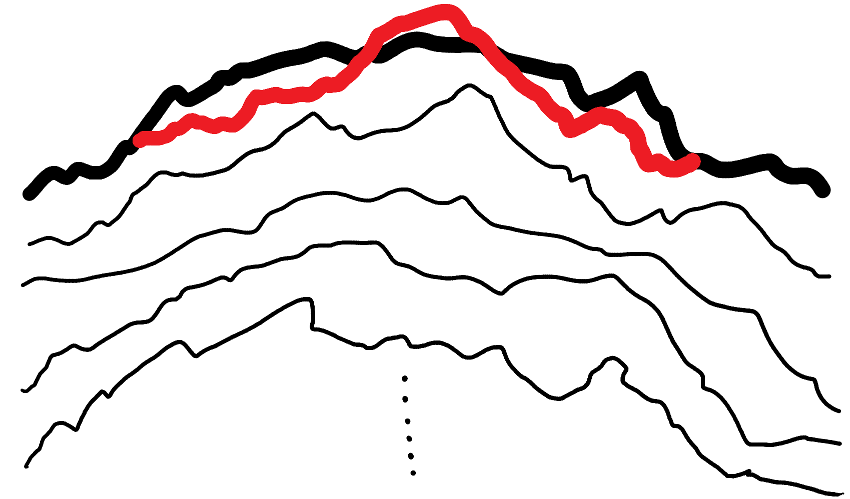
Sample paths of

$$\Psi_0^{-1}\left(\tfrac{1}{2}\Delta+\sum_{i=1}^{N-1}e^{x_{i+1}-x_i}\right)\Psi_0$$

are **invariant** in law under *soft* Brownian bridge exponential energy resampling



Careful limit yields **KPZ line ensemble**: Invariance survives giving regularity of $h(t,\cdot)$, even as $t\to\infty$ under KPZ scaling.



Conj: As $t\to\infty$ limit is **Airy line ensemble** (ALE) with *hard* Gibbs prop.

Conj: ALE is the unique trans. invariant ergodic [C-Sun '14] Gibbs meas.

Summary

- Integrable examples provide detailed information about universality classes, expand their scope, refine their properties.
- They originate from algebraic origins, two of which are representation theory (sym. functs) and int. systems (QISM).
- Building bridges from these areas to probability gives us tools to discover and analyze many new systems.
- The examples considered have revealed much about KPZ class.
- There remain many challenges and further directions to explore.

We are at the De Moivre/Laplace stage and not yet Lyapunov.