COURSE NOTES ON THE KPZ FIXED POINT

DANIEL REMENIK

Abstract. These notes are a supplement to the lectures at the minicourse on the KPZ fixed point given at the CIRM Research School "Random Structures in Statistical Mechanics and Mathematical Physics" in March 2017. Much of the material is taken directly from the recent paper [MQR17b] by Matetski, Quastel and the author; the aim here is to simplify and shorten the presentation of the main results, stripping the arguments of many technical details and focusing mostly on the key ideas leading to this development.

1. The KPZ fixed point

The aim of these notes is to present the recent development [MQR17b], where the KPZ fixed point, a scaling invariant Markov process taking values in real valued functions which look locally like Brownian motion, was constructed as limit of the totally asymmetric exclusion process. The construction leads to an explicit Fredholm determinant formula for its transition probabilities and to proofs of its main properties. We begin with some brief motivation.

1.1. **The KPZ equation and universality class.** The *Kardar-Parisi-Zhang (KPZ) universality class* is a broad collection of one-dimensional, asymmetric, randomly forced systems, including stochastic interface growth on a one-dimensional substrate, directed polymer chains in a random potential, driven lattice gas models, reaction-diffusion models in two-dimensional random media and randomly forced Hamilton-Jacobi equations. It can be loosely characterized by having local dynamics, a smoothing mechanism, slope dependent growth rate (lateral growth) and space-time random forcing with rapid decay of correlations.

Models in this class present an unusual size and distribution of fluctuations. Thinking in terms of the growth of a random one-dimensional interface, KPZ models exhibit interfaces moving at a non-zero velocity proportional to time t, with fluctuations of size $t^{1/3}$ and decorrelating at a spatial scale of $t^{2/3}$. The distribution of the fluctuations depends on the geometry of the specific problem being studied, and it has been found that in many interesting cases they are related to objects coming from random matrix theory (RMT). For background on the different aspects of the study of this class and an account of some of the main developments in the subject in the last fifteen years see the reviews [Cor12; Qua11; QR14; QS15].

The model which gives its name to this class is the *Kardar-Parisi-Zhang equation* [KPZ86],

$$\partial_t H = \lambda(\partial_x H)^2 + \nu\partial_x^2 H + \sqrt{D}\xi, \tag{1.1}$$

a canonical continuum equation for the evolution of a randomly growing one-dimensional interface. It was predicted in [HH85; KPZ86] that the 1:2:3 rescaled solution

$$H_\varepsilon(t,x) = \varepsilon^{1/2} H(\varepsilon^{-3/2}t, \varepsilon^{-1}x) \tag{1.2}$$

should have a non-trivial limiting behavior. The (conjectural) limit under this scaling is the so-called *KPZ fixed point*, which is believed to be the universal limit for models on the KPZ universality class under the KPZ (1:2:3) scaling, and should contain all the fluctuation behaviour seen in this class. As a (conjectural) object, it was introduced non-rigorously in [CQR15], to which we point (together with [MQR17b]) for further background. One can understand many (but by no means all) of the main advances in the field as attempts to understand some aspects of this limiting object (mostly restricted to a very special family of particular initial conditions).

Revision: March 3, 2017.

1.2. **TASEP and the KPZ fixed point.**

1.2.1. *TASEP.* Our starting point in the construction of the KPZ fixed point will be another basic model in the KPZ universality class, the *totally asymmetric exclusion process (TASEP)*. We will introduce this model in more detail in Section 3, so for now describe it slightly informally.

Definition 1.1. (TASEP height function) The *height function* associated to TASEP is a continuous time Markov process $\big(h_t(x)\big)_{x\in\mathbb{Z}}$ taking values in the space of simple random walk paths, by which we mean continuous curves obtained by linearly interpolating the graph of a function from $\mathbb{Z}$ to $\mathbb{Z}$ which moves up or down by one in each step. We think of this curve as an interface growing (downwards) in time, according to the following dynamics: each local maximum ($\wedge$) turns into a local minimum ($\vee$) at rate one, that is if $h_t(z) = h_t(z \pm 1) + 1$ then $h_t(z) \mapsto h_t(z) - 2$ at rate 1 (see Figure 1).

1.2.2. *The KPZ fixed point.* For each $\varepsilon > 0$ the 1:2:3 rescaled TASEP height function is[1]

$$\mathfrak{h}^\varepsilon(\mathbf{t},\mathbf{x}) = \varepsilon^{1/2}\Big[h_{2\varepsilon^{-3/2}\mathbf{t}}(2\varepsilon^{-1}\mathbf{x}) + \varepsilon^{-3/2}\mathbf{t}\Big]. \tag{1.3}$$

The initial data corresponds to just rescaling h_0 diffusively, $\mathfrak{h}^\varepsilon(0,\mathbf{x}) = \varepsilon^{1/2}h_0(2\varepsilon^{-1}\mathbf{x})$ (the extra 2 is just a choice of normalization), and we allow $h_0 = h_0^{(\varepsilon)}$ itself to depend on ε and assume that

$$\mathfrak{h}_0 = \lim_{\varepsilon\to 0}\mathfrak{h}^\varepsilon(0,\cdot) \tag{1.4}$$

(in a sense to be specified later).

Definition 1.2. (The KPZ fixed point) We *define* the *KPZ fixed point* as the limit

$$\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{h}_0) = \lim_{\varepsilon\to 0}\mathfrak{h}^\varepsilon(\mathbf{t},\mathbf{x}). \tag{1.5}$$

We will often omit $\mathfrak{h}_0$ from the notation when it is clear from the context.

(Of course, one of the main points which we need to settle is that the limit (1.5) exists).

As we already mentioned, the KPZ fixed point $\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{h}_0)$ should be the 1:2:3 scaling limit of all models in the KPZ universality class. This is known as the *strong KPZ universality conjecture*. In these lectures (and in [MQR17b]) we only deal with TASEP, but our method works for several variants of TASEP (which also have a representation through biorthogonal ensembles, such as PushASEP as well as discrete time TASEPs and PNGs); this is the content of [MQR17a].

1.2.3. *Earlier work.* TASEP has been one of the most heavily studied models in the KPZ class, and much effort has been devoted to the study of the distributional limit (1.5) for a few, very special, choices of initial data h_0.

Example 1.3. (Step) The most basic case is the *step* initial data $h_0(x) = -|x|$ (the name step comes from the derivative of h_0), in which case it is known that $\mathfrak{h}^\varepsilon(2,\mathbf{x})$ converges to the Airy$_2$ process $\mathcal{A}_2(\mathbf{x})$ [PS02; Joh03], which has one-point marginals given by the Tracy-Widom GUE distribution [TW94]. In this case we have that $\mathfrak{h}(0,\cdot)$ goes to $\mathfrak{d}_0$, where $\mathfrak{d}_\mathbf{u}$ is the *narrow wedge at* $\mathbf{u}$, which is the function $\mathfrak{d}_\mathbf{u}(\mathbf{u}) = 0$, $\mathfrak{d}_\mathbf{u}(\mathbf{x}) = -\infty$ for $\mathbf{x} \neq \mathbf{u}$. So we have

$$\mathfrak{h}(1,\mathbf{x};\mathfrak{d}_0) + \mathbf{x}^2 \;=\; \mathcal{A}_2(\mathbf{x}). \tag{1.6}$$

[1] There is a difference between the scalings chosen in (1.3) and in the current arXiv version of [MQR17b], where the analogous definition was done with $\mathbf{t}$ replaced by $\mathbf{t}/2$ on the right hand side. The present choice is better, because it makes the usual Airy processes arise at time $\mathbf{t} = 1$ at the fixed point level (instead of $\mathbf{t} = 2$), and because it matches a natural choice of parameters for (1.1); it will be the scaling used in the upcoming revised version of that paper. This difference means that, when referencing [MQR17b] in these notes, some constants will be off by a factor of 2 (for example, in that paper we defined $\mathbf{S}_{\mathbf{t},\mathbf{x}} = e^{\mathbf{x}\partial^2 - \frac{\mathbf{t}}{6}\partial^3}$), but this does not affect the arguments in any other way.

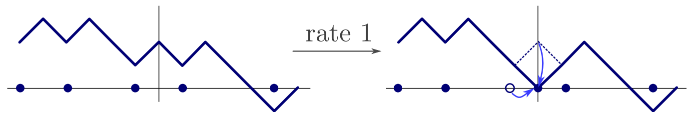

Figure 1. The height function associated to TASEP. A particle moving to the right corresponds to a local maximum of the height function moving down to turn into a local minimum (see also Definition 3.1).

Example 1.4. (Flat) Another case is that of *flat* initial data $h_0(x) = 1$ for even x, $h_0(x) = 0$ for odd x, interpolated lineraly in between. Here we have convergence to the Airy$_1$ process $\mathcal{A}_1(\mathbf{x})$ [Sas05; BFPS07], which has one-point marginals given by the Tracy-Widom GOE distribution [TW96], and so we have

$$\mathfrak{h}(1, \mathbf{x}; 0) \;=\; \mathcal{A}_1(\mathbf{x}). \tag{1.7}$$

Example 1.5. (Stationary) The other basic case corresponds to starting with h_0 coming from sampling a double-sided simple symmetric random walk path, which yields $\mathfrak{h}(1, \mathbf{x}; \mathbf{B})$ with $\mathbf{B}$ a double-sided Brownian motion with diffusion coefficient 2; this is known as the Airy$_{\mathrm{stat}}$ process [BFP10] (the name stationary comes from the fact that TASEP is invariant under the product measure initial initial condition associated to this choice of h_0).

What has made these cases tractable is the fact that exact formulas were available for TASEP, which were derived using the very special properties of these initial conditions. The same methods also yield formulas for the three mixed cases corresponding to putting one of these initial conditions on one side of the origin and another on the other.

Our main result will show that the limit (1.5) exists, in a suitable sense, for all suitable $\mathfrak{h}_0$.

1.3. **State space and topology.** The state space in which we will show that (1.5) holds and where (1.4) will be assumed to hold (in distribution) will be the following:

Definition 1.6. (UC functions) We define UC as the space of upper semicontinuous functions $\mathfrak{h} \colon \mathbb{R} \to [-\infty, \infty)$ with $\mathfrak{h}(\mathbf{x}) \leq C(1 + |\mathbf{x}|)$ for some $C < \infty$.

We will endow this space with the topology of local UC convergence. This is the natural topology for lateral growth, and will allow us to compute $\mathfrak{h}(\mathbf{t}, \mathbf{x}; \mathfrak{h}_0)$ in all cases of interest[2]. In order to define this topology, recall that $\mathfrak{h}$ is upper semicontinuous (UC) if and only if its *hypograph*

$$\mathrm{hypo}(\mathfrak{h}) = \{(\mathbf{x}, \mathbf{y}) : \mathbf{y} \leq \mathfrak{h}(\mathbf{x})\}$$

is closed in $[-\infty, \infty) \times \mathbb{R}$. Slightly informally, local UC convergence can be defined as follows:

Definition 1.7. (Local UC convergence) We say that $\big(\mathfrak{h}_\varepsilon\big)_\varepsilon \subseteq \mathrm{UC}$ *converges locally in* UC to $\mathfrak{h} \in \mathrm{UC}$ if there is a $C > 0$ such that $\mathfrak{h}_\varepsilon(\mathbf{x}) \leq C(1 + |\mathbf{x}|)$ for all $\varepsilon > 0$ and for every $M \geq 1$ there is a $\delta = \delta(\varepsilon, M) > 0$ going to 0 as $\varepsilon \to 0$ such that the hypographs $\mathfrak{H}_{\varepsilon,M}$ and $\mathfrak{H}_M$ of $\mathfrak{h}_\varepsilon$ and $\mathfrak{h}$ restricted to $[-M, M]$ are δ-close in the sense that

$$\cup_{(\mathbf{t},\mathbf{x}) \in \mathfrak{H}_{\varepsilon,M}} B_\delta((\mathbf{t}, \mathbf{x})) \subseteq \mathfrak{H}_M \qquad \text{and} \qquad \cup_{(\mathbf{t},\mathbf{x}) \in \mathfrak{H}_M} B_\delta((\mathbf{t}, \mathbf{x})) \subseteq \mathfrak{H}_{\varepsilon,M}.$$

See [MQR17b, Sec. 3.1] for more details. We will also use an analogous space LC, made of lower semicontinuous functions:

Definition 1.8. (LC functions and local convergence) We let

$$\mathrm{LC} = \big\{\mathfrak{g} \colon -\mathfrak{g} \in \mathrm{UC}\big\}.$$

[2] Actually the bound $\mathfrak{h}(\mathbf{x}) \leq C(1 + |\mathbf{x}|)$ which we are imposing here and in [MQR17b] on UC functions is not as general as possible, but makes the arguments a bit simpler and it suffices for most cases of interest (see also [MQR17b, Foot. 9]).

We endow this space with the topology of *local* LC *convergence* which is defined analogously to local UC convergence, now in terms of *epigraphs*,

$$\mathrm{epi}(\mathfrak{g}) = \{(\mathbf{x}, \mathbf{y}) : \mathbf{y} \geq \mathfrak{g}(\mathbf{x})\}.$$

Explicitly, $\mathfrak{g}_\varepsilon$ converges locally in LC to $\mathfrak{g}$ if and only if $-\mathfrak{g}_\varepsilon \longrightarrow -\mathfrak{g}$ locally in UC.

1.4. **Operators.** In order to state our main result we need to introduce several operators, which will appear in the explicit Fredholm determinant formula for the fixed point.

Definition 1.9. (Projections) For a fixed vector $a \in \mathbb{R}^M$ and indices $n_1 < \ldots < n_M$ we introduce the functions

$$\chi_a(n_j, x) = \mathbf{1}_{x > a_j}, \qquad \bar{\chi}_a(n_j, x) = \mathbf{1}_{x \leq a_j},$$

which we also regard as multiplication operators acting on the spaces $L^2(\{t_1, \ldots, t_M\} \times \mathbb{R})$ and $\ell^2(\{n_1, \ldots, n_M\} \times \mathbb{Z})$. We will use the same notation if a is a scalar, writing

$$\chi_a(x) = 1 - \bar{\chi}_a(x) = \mathbf{1}_{x > a}.$$

We think of χ_a as a projection from $L^2(\mathbb{R})$ to $L^2((a, \infty))$ (and analogously in the vector case).

Our basic building block is the following (almost) group of operators:

Definition 1.10. We define

$$\mathbf{S}_{\mathbf{t},\mathbf{x}} = \exp\{\mathbf{x}\partial^2 - \tfrac{\mathbf{t}}{3}\partial^3\}, \qquad \mathbf{x}, \mathbf{t} \in \mathbb{R}^2 \setminus \{\mathbf{x} < 0, \mathbf{t} = 0\},$$

which satisfy

$$\mathbf{S}_{\mathbf{s},\mathbf{x}}\mathbf{S}_{\mathbf{t},\mathbf{y}} = \mathbf{S}_{\mathbf{s}+\mathbf{t},\mathbf{x}+\mathbf{y}} \tag{1.8}$$

as long as all subscripts avoid $\{\mathbf{x} < 0, \mathbf{t} = 0\}$.

We can think of these as unbounded operators with domain $\mathscr{C}_0^\infty(\mathbb{R})$.

Remark 1.11. In these notes we will not worry about precise statements (or proofs) justifying the convergence of kernels as needed in each step. So for example, at this particular point we will not worry about making the domains and analytical properties of the $\mathbf{S}_{\mathbf{t},\mathbf{x}}$ precise).

But it is not even clear whether the operators make any sense for $\mathbf{x} < 0$, $\mathbf{t} \neq 0$ (notice that for $\mathbf{t} = 0$ they clearly don't). The fact that they do can be checked using the following explicit representation for the operators: $\mathbf{S}_{\mathbf{t},\mathbf{x}}$ acts by convolution

$$\mathbf{S}_{\mathbf{t},\mathbf{x}} f(z) = \int_{-\infty}^{\infty} dy\, \mathbf{S}_{\mathbf{t},\mathbf{x}}(z, y) f(y) = \int_{-\infty}^{\infty} dy\, \mathbf{S}_{\mathbf{t},\mathbf{x}}(z - y) f(y),$$

with convolution kernel $\mathbf{S}_{\mathbf{t},\mathbf{x}}(z, y) = \mathbf{S}_{\mathbf{t},\mathbf{x}}(z - y)$ given by

$$\mathbf{S}_{\mathbf{t},\mathbf{x}}(z) = \frac{1}{2\pi\mathrm{i}} \int_{\langle} dw\, e^{\frac{\mathbf{t}}{3}w^3 + \mathbf{x}w^2 - zw} = \mathbf{t}^{-1/3} e^{\frac{2\mathbf{x}^3}{3\mathbf{t}^2} + \frac{z\mathbf{x}}{\mathbf{t}}} \mathrm{Ai}(\mathbf{t}^{-1/3} z + \mathbf{t}^{-4/3}\mathbf{x}^2) \tag{1.9}$$

for $\mathbf{t} > 0$ together with

$$\mathbf{S}_{-\mathbf{t},\mathbf{x}} = (\mathbf{S}_{\mathbf{t},\mathbf{x}})^* \quad \text{or} \quad \mathbf{S}_{-\mathbf{t},\mathbf{x}}(z, y) = \mathbf{S}_{-\mathbf{t},\mathbf{x}}(z - y) = \mathbf{S}_{\mathbf{t},\mathbf{x}}(y - z) \tag{1.10}$$

for $\mathbf{t} < 0$, where $\langle$ is the positively oriented contour going in straight lines from $e^{-\mathrm{i}\pi/3}\infty$ to $e^{\mathrm{i}\pi/3}\infty$ through 0 and Ai is the Airy function $\mathrm{Ai}(z) = \frac{1}{2\pi\mathrm{i}} \int_{\langle} dw\, e^{\frac{1}{3}w^3 - zw}$.

Remark 1.12. The fact that these operators make sense is just a generalization of the fact that $e^{-\mathbf{x}\partial^2}\mathrm{Ai}$ is well defined for all $\mathbf{x} \in \mathbb{R}$, which has been used extensively in the field (see e.g. [BFPS07; QR13; QR16]), and corresponds simply to taking $\mathbf{x} = 1$ in the above definition.

From (1.10) we get directly the identity

$$(\mathbf{S}_{\mathbf{t},\mathbf{x}})^*\mathbf{S}_{\mathbf{t},-\mathbf{x}} = \mathbf{I}.$$

Definition 1.13. (Hit operators) For $\mathfrak{g} \in \mathrm{LC}$ we define the operator

$$\bar{\mathbf{S}}^{\mathrm{epi}(\mathfrak{g})}_{\mathbf{t},\mathbf{x}}(v,u) = \mathbb{E}_{\mathbf{B}(0)=v}\big[\mathbf{S}_{\mathbf{t},\mathbf{x}-\boldsymbol{\tau}}(\mathbf{B}(\boldsymbol{\tau}),u)\mathbf{1}_{\boldsymbol{\tau}<\infty}\big] = \int_0^\infty \mathbb{P}_{\mathbf{B}(0)=v}(\boldsymbol{\tau}\in d\mathbf{s})\mathbf{S}_{\mathbf{t},\mathbf{x}-\mathbf{s}}(\mathbf{B}(\boldsymbol{\tau}),u)$$

where $\mathbf{B}(x)$ is a Brownian motion with diffusion coefficient 2 and $\boldsymbol{\tau}$ is the hitting time of the epigraph of $\mathfrak{g}|_{[0,\infty)}$. Note that, trivially,

$$\bar{\mathbf{S}}^{\mathrm{epi}(\mathfrak{g})}_{\mathbf{t},\mathbf{x}}(v,u) = \mathbf{S}_{\mathbf{t},\mathbf{x}}(v,u) \quad \text{for } v \geq \mathfrak{g}(0). \tag{1.11}$$

If $\mathfrak{h} \in \mathrm{UC}$, there is a similar operator $\bar{\mathbf{S}}^{\mathrm{hypo}(\mathfrak{h})}_{\mathbf{t},\mathbf{x}}$ with the same definition, except that now $\boldsymbol{\tau}$ is the hitting time of the hypograph of $\mathfrak{h}$ and $\bar{\mathbf{S}}^{\mathrm{hypo}(\mathfrak{h})}_{\mathbf{t},\mathbf{x}}(v,u) = \mathbf{S}_{\mathbf{t},\mathbf{x}}(v,u)$ for $v \leq \mathfrak{h}(0)$.

One way to think of $\bar{\mathbf{S}}^{\mathrm{epi}(\mathfrak{g})}_{\mathbf{t},\mathbf{x}}(v,u)$ is as a sort of asymptotic tranformed transition "probability" for the Brownian motion $\mathbf{B}$ to go from v to u hitting the epigraph of $\mathfrak{g}$ (note that $\mathfrak{g}$ is not necessarily continuous, so hitting $\mathfrak{g}$ is not the same as hitting $\mathrm{epi}(\mathfrak{g})$; in particular, $\mathbf{B}(\boldsymbol{\tau}) \geq \mathfrak{g}(\boldsymbol{\tau})$ and in general the equality need not hold). To see what we mean, write

$$\bar{\mathbf{S}}^{\mathrm{epi}(\mathfrak{g})}_{\mathbf{t},\mathbf{x}} = \lim_{\mathbf{T}\to\infty}\mathbf{S}_{\mathbf{t},\mathbf{x}-\mathbf{T}}\bar{\mathbf{S}}^{\mathrm{epi}(\mathfrak{g}),\mathbf{T}} \qquad \text{with} \qquad \bar{\mathbf{S}}^{\mathrm{epi}(\mathfrak{g}),\mathbf{T}}(v,u) = \mathbb{E}_{\mathbf{B}(0)=v}\big[\mathbf{S}_{0,\mathbf{T}-\boldsymbol{\tau}}(\mathbf{B}(\boldsymbol{\tau}),u)\mathbf{1}_{\boldsymbol{\tau}\leq\mathbf{T}}\big] \tag{1.12}$$

and note that $\bar{\mathbf{S}}^{\mathrm{epi}(\mathfrak{g}),\mathbf{T}}(v,u)$ is nothing but the transition probability for $\mathbf{B}$ to go from v at time 0 to u at time $\mathbf{T}$ hitting $\mathrm{epi}(\mathfrak{g})$ in $[0,\mathbf{T}]$.

Definition 1.14. (Epi/hypo operators) For $\mathfrak{g} \in \mathrm{LC}$ and $\mathbf{x} \in \mathbb{R}$ we define

$$\mathbf{K}^{\mathrm{epi}(\mathfrak{g})}_{\mathbf{t}} = \mathbf{I} - (\mathbf{S}_{\mathbf{t},\mathbf{x}} - \bar{\mathbf{S}}^{\mathrm{epi}(\mathfrak{g}^-_{\mathbf{x}})}_{\mathbf{t},\mathbf{x}})^*\bar{\chi}_{\mathfrak{g}(\mathbf{x})}(\mathbf{S}_{\mathbf{t},-\mathbf{x}} - \bar{\mathbf{S}}^{\mathrm{epi}(\mathfrak{g}^+_{\mathbf{x}})}_{\mathbf{t},-\mathbf{x}}), \tag{1.13}$$

where

$$\mathfrak{g}^+_{\mathbf{x}}(\mathbf{y}) = \mathfrak{g}(\mathbf{x}+\mathbf{y}), \qquad \mathfrak{g}^-_{\mathbf{x}}(\mathbf{y}) = \mathfrak{g}(\mathbf{x}-\mathbf{y}).$$

Note that the projection $\bar{\chi}_{\mathfrak{g}(\mathbf{x})}$ can be removed from (1.13), thanks to (1.11).

There is another operator which uses $\mathfrak{h} \in \mathrm{UC}$, and hits "from above",

$$\mathbf{K}^{\mathrm{hypo}(\mathfrak{h})}_{\mathbf{t}} = \mathbf{I} - (\mathbf{S}_{\mathbf{t},\mathbf{x}} - \bar{\mathbf{S}}^{\mathrm{hypo}(\mathfrak{h}^-_{\mathbf{x}})}_{\mathbf{t},\mathbf{x}})^*\chi_{\mathfrak{h}(\mathbf{x})}(\mathbf{S}_{\mathbf{t},-\mathbf{x}} - \bar{\mathbf{S}}^{\mathrm{hypo}(\mathfrak{h}^+_{\mathbf{x}})}_{\mathbf{t},-\mathbf{x}}), \tag{1.14}$$

As above, $\mathbf{S}_{\mathbf{t},\mathbf{x}} - \bar{\mathbf{S}}^{\mathrm{epi}(\mathfrak{g}_{\mathbf{x}})}_{\mathbf{t},\mathbf{x}}$ may be thought of as a sort of asymptotic transformed transition probability for a Brownian motion $\mathbf{B}$, started at time 0, not to hit $\mathrm{epi}(\mathfrak{g}|_{[0,\infty)})$. Therefore $\mathbf{K}^{\mathrm{epi}(\mathfrak{g})}_{\mathbf{t}}$ may be thought of as the same sort of asymptotic transformed transition probability for $\mathbf{B}$, in this case hitting $\mathrm{epi}(\mathfrak{g})$, which is built out of the product of left and right "no hit" operators.

Remark 1.15. Note that in this description the point $\mathbf{x}$ at which $\mathfrak{g}$ is split in (1.13) plays absolutely no role; the operator $\mathbf{K}^{\mathrm{epi}(\mathfrak{g})}_{\mathbf{t}}$ actually does not depend on the choice of $\mathbf{x}$, and this fact will be used at a crucial step in the derivation. This was proved in [QR16] (see the next remark), essentially by using the representation given in (1.12).

Remark 1.16. For $\mathbf{t} = 1$, $\mathbf{I} - \mathbf{K}^{\mathrm{epi}(\mathfrak{g})}_{\mathbf{t}}$ is the *Brownian scattering operator* introduced in [QR16]. In that paper we derived, under an assumption which is widely believe to hold but which currently escapes rigorous treatment (namely that the *partially* asymmetric exclusion process with step initial data converges to the Airy_2 process) explicit formulas for the one-point marginals of the limit of the rescaled solution of the KPZ equation, (1.2). These formulas coincide with those which follow from the KPZ fixed point formulas to be given below, which provides further (non-rigorous) confirmation of the strong KPZ universality conjecture.

1.5. **Existence and formulas for the KPZ fixed point.** We are ready to state the main result of [MQR17b] (Theorem 4.1 in that paper), which proves the existence of the KPZ fixed point $\mathfrak{h}(\mathbf{t},\mathbf{x})$ and characterizes its distribution for each $\mathbf{t}\geq 0$.

Theorem 1.17. (KPZ fixed point formula) *Fix $\mathfrak{h}_0\in\mathrm{UC}$. Let $h_0^{(\varepsilon)}$ be initial data for the TASEP height function such that the corresponding rescaled height functions $\mathfrak{h}_0^{\varepsilon}$ (1.3) converge to $\mathfrak{h}_0$ locally in* UC *as $\varepsilon\to 0$. Then the limit (1.5) for $\mathfrak{h}(\mathbf{t},\mathbf{x})$ exists (in distribution) locally in* UC *and is given as follows: for any $\mathfrak{g}\in\mathrm{LC}$,*

$$\mathbb{P}(\mathfrak{h}(\mathbf{t},\mathbf{x})\leq\mathfrak{g}(\mathbf{x}),\,\mathbf{x}\in\mathbb{R})=\det\Big(\mathbf{I}-\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t}/2}\mathbf{K}^{\mathrm{epi}(\mathfrak{g})}_{-\mathbf{t}/2}\Big)_{L^2(\mathbb{R})}. \tag{1.15}$$

The determinant on the right hand side of (1.15) is the Fredholm determinant on the Hilbert space $L^2(\mathbb{R})$, which may be defined as follows: for K an operator with integral kernel $K(x,y)$,

$$\det(I-K)=\sum_{n\geq 0}\frac{(-1)^n}{n!}\int_{\mathbb{R}^n}dx_1\cdots dx_n\,\det[K(x_i,x_j)]_{i,j=1}^n.$$

For background on Fredholm determinants we refer to [Sim05] or [QR14, Sec. 2].

Remark 1.18. The fact that the Fredholm determinant in the formula is finite is a consequence of the fact that there is a (multiplication) operator M such that $M\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t}/2}\mathbf{K}^{\mathrm{epi}(\mathfrak{g})}_{-\mathbf{t}/2}M^{-1}$ is trace class (this fact is proved in [MQR17b, Sec. B.1]). However, as we already mentioned, in these lectures we will omit all these issues.

Note that (1.15) gives a lot more information about the distribution of the fixed point than finite-dimensional distributions. This type of "continuum statistics" formulas have been previously derived for Airy and related processes in [CQR13; QR13; BCR15; NR15; NR16]. In our case it is crucial that we have such a general formula, for it will allow us to prove that the fixed point is a Markov process. However, the formulas for the finite dimensional distributions from which Theorem 1.17 will follow are interesting and useful in themselves, so we state them here:

Theorem 1.19. (KPZ fixed point finite dimensional distributions) *Let $\mathfrak{h}_0\in\mathrm{UC}$ and choose $\mathbf{x}_1<\mathbf{x}_2<\cdots<\mathbf{x}_M$. Then*

$$\begin{aligned}&\mathbb{P}_{\mathfrak{h}_0}(\mathfrak{h}(\mathbf{t},\mathbf{x}_1)\leq\mathbf{a}_1,\dots,\mathfrak{h}(\mathbf{t},\mathbf{x}_M)\leq\mathbf{a}_M)\\&\quad=\det\Big(\mathbf{I}-\chi_{\mathbf{a}}\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t},\mathrm{ext}}\chi_{\mathbf{a}}\Big)_{L^2(\{\mathbf{x}_1,\dots,\mathbf{x}_M\}\times\mathbb{R})} &(1.16)\\&\quad=\det\Big(\mathbf{I}-\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t},\mathbf{x}_M}+\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t},\mathbf{x}_M}e^{(\mathbf{x}_1-\mathbf{x}_M)\partial^2}\bar\chi_{\mathbf{a}_1}e^{(\mathbf{x}_2-\mathbf{x}_1)\partial^2}\bar\chi_{\mathbf{a}_2}\cdots e^{(\mathbf{x}_M-\mathbf{x}_{M-1})\partial^2}\bar\chi_{\mathbf{a}_M}\Big)_{L^2(\mathbb{R})} &(1.17)\end{aligned}$$

where

$$\mathbf{K}^{\mathrm{hypo}(\mathbf{h}_0)}_{\mathbf{t},\mathrm{ext}}(\mathbf{x}_i,\cdot;\mathbf{x}_j,\cdot)=-e^{(\mathbf{x}_j-\mathbf{x}_i)\partial^2}\mathbf{1}_{\mathbf{x}_i<\mathbf{x}_j}+e^{-\mathbf{x}_i\partial^2}\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t}}e^{\mathbf{x}_j\partial^2} \tag{1.18}$$

and

$$\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t},\mathbf{x}}=e^{-\mathbf{x}\partial^2}\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t}}e^{\mathbf{x}\partial^2} \tag{1.19}$$

with $\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t}}$ the kernel defined in (1.14).

The kernel in (1.16) is usually referred to as an *extended kernel* (note that the Fredholm determinant is being computed on the "extended L^2 space" $L^2(\{\mathbf{x}_1,\dots,\mathbf{x}_M\}\times\mathbb{R})$). The kernel appearing after the second hypo operator in (1.17) is sometimes referred to as a *path integral kernel* [BCR15], and should be thought of as a discrete, pre-asymptotic version of the epi operators (on a finite interval).

Remark 1.20. The fact that $e^{-\mathbf{x}\partial^2}\mathbf{K}_{\mathbf{t}}^{\mathrm{hypo}(\mathfrak{h}_0)}e^{\mathbf{x}\partial^2}$ makes sense is not entirely obvious, but follows from the fact that $\mathbf{K}_{\mathbf{t}}^{\mathrm{hypo}(\mathfrak{h}_0)}$ equals

$$(\mathbf{S}_{\mathbf{t},\mathbf{x}})^*\chi_{\mathfrak{g}(\mathbf{x})}\mathbf{S}_{\mathbf{t},-\mathbf{x}}+(\bar{\mathbf{S}}_{\mathbf{t},\mathbf{x}}^{\mathrm{epi}(\mathfrak{g}_{\mathbf{x}}^-)})^*\bar{\chi}_{\mathfrak{g}(\mathbf{x})}\mathbf{S}_{\mathbf{t},-\mathbf{x}}+(\mathbf{S}_{\mathbf{t},\mathbf{x}})^*\bar{\chi}_{\mathfrak{g}(\mathbf{x})}\bar{\mathbf{S}}_{\mathbf{t},-\mathbf{x}}^{\mathrm{epi}(\mathfrak{g}_{\mathbf{x}}^+)}-(\bar{\mathbf{S}}_{\mathbf{t},\mathbf{x}}^{\mathrm{epi}(\mathfrak{g}_{\mathbf{x}}^-)})^*\bar{\chi}_{\mathfrak{g}(\mathbf{x})}\bar{\mathbf{S}}_{\mathbf{t},-\mathbf{x}}^{\mathrm{epi}(\mathfrak{g}_{\mathbf{x}}^+)},$$

together with the group property (1.8) and the definition of the hit operators.

2. Properties of the KPZ fixed point

In this section we present the main properties of the KPZ fixed point, as well as a sketch of the proof of some of them.

2.1. **Markov property.** The sets $A_{\mathfrak{g}}=\{\mathfrak{h}\in\mathrm{UC}:\mathfrak{h}(\mathbf{x})\leq\mathfrak{g}(\mathbf{x}),\ \mathbf{x}\in\mathbb{R}\}$, $\mathfrak{g}\in\mathrm{LC}$, form a generating family for the Borel sets $\mathcal{B}(\mathrm{UC})$. Hence from (1.15) we can define the fixed point transition probabilities $p_{\mathfrak{h}_0}(\mathbf{t},A_{\mathfrak{g}})$:

Lemma 2.1. *For fixed $\mathfrak{h}_0\in\mathrm{UC}$ and $\mathbf{t}>0$, the measure $p_{\mathfrak{h}_0}(\mathbf{t},\cdot)$ defined above is a probability measure on* UC*.*

Sketch of the proof. It is clear from the construction that $\mathbb{P}_{\mathfrak{h}_0}(\mathfrak{h}(\mathbf{t},\mathbf{x}_i)\leq\mathbf{a}_i, i=1,\dots,n)$ is non-decreasing in each $\mathbf{a}_i$ and is in $[0,1]$. We need to show then that this quantity goes to 1 as all $\mathbf{a}_i$'s go to infinity, and to 0 if any $\mathbf{a}_i$ goes to 0. The first one is standard, and relies on the inequality $|\det(\mathbf{I}-\mathbf{K})-1|\leq\|\mathbf{K}\|_1e^{\|\mathbf{K}\|_1+1}$ (with $\|\cdot\|_1$ denoting trace norm). The second limit is in general very hard to show for a formula given in terms of a Fredholm determinant, but it turns out to be rather easy in our case, because the multipoint probability is trivially bounded by $\mathbb{P}_{\mathfrak{h}_0}(\mathfrak{h}(\mathbf{t},\mathbf{x}_i)\leq\mathbf{a}_i)$, and then one can use the skew time reversal symmetry and affine invariance of the fixed point (Prop. 2.8(ii,v)) to compare this to the one-point marginals of the Airy$_2$ process. See [MQR17b, Sec. 4.2] for the details. $\square$

Theorem 2.2. *The KPZ fixed point $\big(\mathfrak{h}(\mathbf{t},\cdot)\big)_{\mathbf{t}>0}$ is a (Feller) Markov process taking values in* UC*.*

The proof is based on the fact that $\mathfrak{h}(\mathbf{t},\mathbf{x})$ is the limit of $\mathfrak{h}^{\varepsilon}(\mathbf{t},\mathbf{x})$, which is Markovian. To derive from this the Markov property of the limit requires some compactness, which in our case is provided by Theorem 2.4 below (see [MQR17b, Sec. 4.2] for more details).

2.2. **Regularity and local Brownian behavior.** Up to this point we only know that the fixed point is in UC, but by the smoothing mechanism inherent to models in the KPZ class one should expect $\mathfrak{h}(\mathbf{t},\cdot)$ to at least be continuous for each fixed $\mathbf{t}>0$. The next result shows that $\mathfrak{h}(\mathbf{t},\cdot)$ is actually locally Hölder β for any $\beta<1/2$.

Definition 2.3. (Local Hölder spaces) Let $\mathscr{C}=\{\mathfrak{h}\colon\mathbb{R}\to[-\infty,\infty)$ continuous with $\mathfrak{h}(\mathbf{x})\leq C(1+|\mathbf{x}|)$ for some $C<\infty\}$. We define the local Hölder norm

$$\|\mathfrak{h}\|_{\beta,[-M,M]}=\sup_{\mathbf{x}_1\neq\mathbf{x}_2\in[-M,M]}\frac{|\mathfrak{h}(\mathbf{x}_2)-\mathfrak{h}(\mathbf{x}_1)|}{|\mathbf{x}_2-\mathbf{x}_1|^{\beta}}$$

and the local Hölder spaces

$$\mathscr{C}^{\beta}=\{\mathfrak{h}\in\mathscr{C}\text{ with }\|\mathfrak{h}\|_{\beta,[-M,M]}<\infty\text{ for each }M=1,2,\dots\}.$$

The topology on UC, when restricted to $\mathscr{C}$, is the topology of uniform convergence on compact sets. UC is a Polish space and the spaces $\mathscr{C}^{\beta}$ are compact in UC.

Theorem 2.4. (Space regularity) *Fix $\mathbf{t}>0$, $\mathfrak{h}_0\in\mathrm{UC}$ and initial data $h_0^{(\varepsilon)}$ for the TASEP height function such that $\mathfrak{h}_0^{\varepsilon}\longrightarrow\mathfrak{h}_0$ locally in* UC *as $\varepsilon\to0$. Then for each $\beta\in(0,1/2)$ and $M<\infty$,*

$$\lim_{A\to\infty}\limsup_{\varepsilon\to0}\mathbb{P}(\|\mathfrak{h}^{\varepsilon}(\mathbf{t})\|_{\beta,[-M,M]}\geq A)=\lim_{A\to\infty}\mathbb{P}(\|\mathfrak{h}\|_{\beta,[-M,M]}\geq A)=0.$$

The proof proceeds through an application of the Kolmogorov continuity theorem, which reduces regularity to two-point functions, and is based on the arguments introduced in [QR13] for the Airy_1 and Airy_2 processes, and depends heavily on the representation (1.17) for the two-point function in terms of path integral kernels. We skip the details.

Theorem 2.5. (Local Brownian behavior) *For any initial condition $\mathfrak{h}_0 \in \mathrm{UC}$ $\mathfrak{h}(\mathbf{t},\mathbf{x})$ is locally Brownian in $\mathbf{x}$ in the sense that for each $\mathbf{y} \in \mathbb{R}$, the finite dimensional distributions of*

$$\mathfrak{b}_\varepsilon^+(\mathbf{x}) = \varepsilon^{-1/2}(\mathfrak{h}(\mathbf{t},\mathbf{y}+\varepsilon\mathbf{x}) - \mathfrak{h}(\mathbf{t},\mathbf{y})) \qquad \text{and} \qquad \mathfrak{b}_\varepsilon^-(\mathbf{x}) = \varepsilon^{-1/2}(\mathfrak{h}(\mathbf{t},\mathbf{y}-\varepsilon\mathbf{x}) - \mathfrak{h}(\mathbf{t},\mathbf{y}))$$

converge, as $\varepsilon \searrow 0$, to Brownian motions with diffusion coefficient 2.

Very brief sketch of the proof. The proof is based again on the arguments of [QR13]. One uses (1.17) and Brownian scale invariance to show that

$$\mathbb{P}\big(\mathfrak{h}(\mathbf{t},\varepsilon\mathbf{x}_1) \le \mathbf{u} + \sqrt{\varepsilon}\mathbf{a}_1, \dots, \mathfrak{h}(\mathbf{t},\varepsilon\mathbf{x}_n) \le \mathbf{u} + \sqrt{\varepsilon}\mathbf{a}_n \,\big|\, \mathfrak{h}(\mathbf{t},0) = \mathbf{u}\big) \\ = \mathbb{E}\big(\mathbf{1}_{\mathbf{B}(\mathbf{x}_i)\le \mathbf{a}_i, i=1,\dots,n}\, \phi^\varepsilon_{\mathbf{x},\mathbf{a}}(\mathbf{u},\mathbf{B}(\mathbf{x}_n))\big)$$

for some explicit function $\phi^\varepsilon_{\mathbf{x},\mathbf{a}}(\mathbf{u},\mathbf{b})$. The Brownian motion appears from the product of heat kernels in (1.17), while $\phi^\varepsilon_{\mathbf{x},\mathbf{a}}$ contains the dependence on everything else in the formula (the Fredholm determinant structure and $\mathfrak{h}_0$ through the hypo operator $\mathbf{K}_{\mathbf{t}}^{\mathrm{hypo}(\mathfrak{h}_0)}$). Then one shows that $\phi^\varepsilon_{\mathbf{x},\mathbf{a}}(\mathbf{u},\mathbf{b})$ goes to 1 in a suitable sense as $\varepsilon \to 0$. □

Proposition 2.6. (Time regularity) *Fix $\mathbf{x}_0 \in \mathbb{R}$ and initial data $\mathfrak{h}_0 \in \mathrm{UC}$. For $\mathbf{t} > 0$, $\mathfrak{h}(\mathbf{t},\mathbf{x}_0)$ is locally Hölder α in $\mathbf{t}$ for any $\alpha < 1/3$.*

The proof uses the variational formula for the fixed point, we sketch it in Section 2.5.

2.3. **Symmetries and invariance.** The KPZ fixed point satisfies a number of properties which follow directly from the construction. The first one is the 1:2:3 scaling invariance, which we state separately since it lies at the heart of our interest in this object (and justifies the name of the process, which is fixed, in distribution, under 1:2:3 rescaling).

Proposition 2.7. (1:2:3 scaling invariance)

$$\alpha\mathfrak{h}(\alpha^{-3}\mathbf{t},\alpha^{-2}\mathbf{x};\alpha\mathfrak{h}_0(\alpha^{-2}\mathbf{x})) \overset{\mathrm{dist}}{=} \mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{h}_0), \quad \alpha > 0.$$

This property follows straightforwardly from Definition 1.2 (and Theorem 1.17).

Proposition 2.8 (Symmetries of $\mathfrak{h}$).

(i) *(Skew time reversal)* $\mathbb{P}\big(\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{g}) \le -\mathfrak{f}(\mathbf{x})\big) = \mathbb{P}\big(\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{f}) \le -\mathfrak{g}(\mathbf{x})\big)$, $\quad \mathfrak{f},\mathfrak{g} \in \mathrm{UC}$;

(ii) *(Shift invariance)* $\mathfrak{h}(\mathbf{t},\mathbf{x}+\mathbf{u};\mathfrak{h}_0(\mathbf{x}+\mathbf{u})) \overset{\mathrm{dist}}{=} \mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{h}_0)$;

(iii) *(Reflection invariance)* $\mathfrak{h}(\mathbf{t},-\mathbf{x};\mathfrak{h}_0(-\mathbf{x})) \overset{\mathrm{dist}}{=} \mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{h}_0)$;

(iv) *(Affine invariance)* $\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{f}(\mathbf{x})+\mathbf{a}+c\mathbf{x}) \overset{\mathrm{dist}}{=} \mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{f}(\mathbf{x}+\tfrac{1}{2}c\mathbf{t})) + \mathbf{a} + c\mathbf{x} + \tfrac{1}{4}c^2\mathbf{t}$;

(v) *(Preservation of max)* $\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{f}_1 \vee \mathfrak{f}_2) = \mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{f}_1) \vee \mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{f}_2)$.

Proof. (ii) and (iii) follow directly from the fact that the TASEP height function satisfies exactly the same properties.

(i) is slightly more delicate, but it also follows easily from the analog property for TASEP, which is perhaps most easily seen by using the graphical representation for TASEP, see e.g. [Lig85], to couple TASEP starting with height profile g at time 0 and running to time t with another copy of TASEP, started now with height profile $-f$ at time t and running backwards to time 0.

(v) also follows from the exact same property for TASEP, which can be proved by showing that if $h_1(t,x)$ and $h_2(t,x)$ are two copies of TASEP coupled again using the graphical representation and such that $h_1(0,x) \ge h_2(0,x)$ for all x then $h_1(t,x) \ge h_2(t,x)$ for all x and all t.

(iv) can be also be proved from TASEP, but it is easier to use the variational formula provided in Theorem 2.15, so we postpone the proof until Section 2.5. □

2.4. **Airy processes.** We show here how to recover (at time $\mathbf{t}=1$) the known Airy$_1$, Airy$_2$ and Airy$_{2\to1}$ processes from Theorem 1.17. The first two actually follow from the third (which interpolates between the two), so we do the computation only in that case.

Definition 2.9. The UC function $\mathfrak{d}_{\mathbf{u}}(\mathbf{u})=0$, $\mathfrak{d}_{\mathbf{u}}(\mathbf{x})=-\infty$ for $\mathbf{x}\neq\mathbf{u}$, is known as a *narrow wedge at* $\mathbf{u}$ (we already used this function in (1.6)).

Example 2.10. (Airy$_2$ process) Narrow wedge initial data leads to the *Airy$_2$ process* [PS02; Joh03]:

$$\mathfrak{h}(1,\mathbf{x};\mathfrak{d}_{\mathbf{u}})+(\mathbf{x}-\mathbf{u})^2 \;=\; \mathcal{A}_2(\mathbf{x})$$

(which we already stated in (1.6)).

Example 2.11. (Airy$_1$ process) *Flat* initial data $\mathfrak{h}_0\equiv0$ leads to the *Airy$_1$ process* [Sas05; BFPS07]:

$$\mathfrak{h}(1,\mathbf{x};0) \;=\; \mathcal{A}_1(\mathbf{x})$$

(which we already stated in (1.7)).

Example 2.12. (Airy$_{2\to1}$ process) *Wedge* or *half-flat* initial data $\mathfrak{h}_{\text{h-f}}(\mathbf{x})=-\infty$ for $\mathbf{x}<0$, $\mathfrak{h}_{\text{h-f}}(\mathbf{x})=0$ for $\mathbf{x}\geq0$, leads to the *Airy$_{2\to1}$ process* [BFS08b]:

$$\mathfrak{h}(1,\mathbf{x};\mathfrak{h}_{\text{h-f}})+\mathbf{x}^2\mathbf{1}_{\mathbf{x}<0} \;=\; \mathcal{A}_{2\to1}(\mathbf{x}).$$

To get this formula we will show that the finite dimensional distributions match, by computing the kernel on the right hand side of (1.16) with $\mathfrak{h}_0^-\equiv-\infty$ and $\mathfrak{h}_0^+(\mathbf{x})\equiv0$. One can also start directly from Theorem 1.17 (see [MQR17b, Sec. 4.4]).

It is straightforward to check that $\bar{\mathbf{S}}_{\mathbf{t},0}^{\text{hypo}(\mathfrak{h}_0^-)}\equiv0$. On the other hand, as in [QR16, Prop. 3.6] one checks that for $v\geq0$,

$$\bar{\mathbf{S}}_{\mathbf{t},0}^{\text{hypo}(\mathfrak{h}_0^+)}(v,u)=\int_0^\infty\mathbb{P}_v(\tau_0\in d\mathbf{y})\mathbf{S}_{\mathbf{t},-\mathbf{y}}(0,u)=\mathbf{S}_{\mathbf{t},0}(-v,u),$$

which gives, with ϱ the reflection operator $\varrho f(x)=f(-x)$,

$$\mathbf{K}_{\mathbf{t}}^{\text{hypo}(\mathfrak{h}_0)}=\mathbf{I}-(\mathbf{S}_{\mathbf{t},0})^*\chi_0[\mathbf{S}_{\mathbf{t},0}-\varrho\mathbf{S}_{\mathbf{t},0}]=(\mathbf{S}_{\mathbf{t},0})^*(\mathbf{I}+\varrho)\bar{\chi}_0\mathbf{S}_{\mathbf{t},0}.$$

This yields (using (1.18) and (1.19))

$$\begin{aligned}\mathbf{K}_{\mathbf{t},\text{ext}}^{\text{hypo}(\mathfrak{h}_{\text{h-f}})}(\mathbf{x}_i,\cdot;\mathbf{x}_j,\cdot)&=-e^{(\mathbf{x}_j-\mathbf{x}_i)\partial^2}\mathbf{1}_{\mathbf{x}_i<\mathbf{x}_j}+\mathbf{S}_{0,-\mathbf{x}_i}(\mathbf{S}_{\mathbf{t},0})^*(\mathbf{I}+\varrho)\bar{\chi}_0\mathbf{S}_{\mathbf{t},0}\mathbf{S}_{0,\mathbf{x}_i}\\&=-e^{(\mathbf{x}_j-\mathbf{x}_i)\partial^2}\mathbf{1}_{\mathbf{x}_i<\mathbf{x}_j}+(\mathbf{S}_{\mathbf{t},-\mathbf{x}_i})^*(\mathbf{I}+\varrho)\bar{\chi}_0\mathbf{S}_{\mathbf{t}/2,\mathbf{x}_i}.\end{aligned}$$

Choosing $\mathbf{t}=1$ and using (1.9) yields that the second term on the right hand side equals

$$\begin{aligned}\mathbf{K}_{\mathbf{t},\text{ext}}^{\text{hypo}(\mathfrak{h}_{\text{h-f}})}(\mathbf{x}_i,u;\mathbf{x}_j,v)&=\int_{-\infty}^0 d\lambda\, e^{-2\mathbf{x}_i^3/3-\mathbf{x}_i(u-\lambda)}\operatorname{Ai}(u-\lambda+\mathbf{x}_i^2)\,e^{2\mathbf{x}_j^3/3+\mathbf{x}_j(v-\lambda)}\operatorname{Ai}(v-\lambda+\mathbf{x}_j^2)\\&\quad+\int_{-\infty}^0 d\lambda\, e^{-2\mathbf{x}_i^3/3-\mathbf{x}_i(u+\lambda)}\operatorname{Ai}(u+\lambda+\mathbf{x}_i^2)\,e^{2\mathbf{x}_j^3/3+\mathbf{x}_j(v-\lambda)}\operatorname{Ai}(v-\lambda+\mathbf{x}_j^2)\end{aligned}$$

which, after a simple conjugation, gives the kernel for the Airy$_{2\to1}$ process [BFS08b, App. A].

2.5. **Variational formulas and the Airy sheet.**

Definition 2.13. (Airy sheet)

$$\mathcal{A}(\mathbf{x},\mathbf{y}) \;=\; \mathfrak{h}(1,\mathbf{y};\mathfrak{d}_{\mathbf{x}})+(\mathbf{x}-\mathbf{y})^2$$

is called the *Airy sheet* [CQR15]. Fixing either one of the variables, it is an Airy$_2$ process in the other. We also write

$$\hat{\mathcal{A}}(\mathbf{x},\mathbf{y})=\mathcal{A}(\mathbf{x},\mathbf{y})-(\mathbf{x}-\mathbf{y})^2.$$

Remark 2.14. The KPZ fixed point inherits from TASEP a canonical coupling between the processes started with different initial data (using the same "noise"). It is this the property that allows us to define the two-parameter Airy sheet.

It is natural to wonder whether the fixed point formulas at our disposal determine the joint probabilities $\mathbb{P}(\mathcal{A}(\mathbf{x}_i,\mathbf{y}_i)\leq\mathbf{a}_i, i=1,\dots,M)$ for the Airy sheet. Unfortunately, this is not the case. In fact, the most we can compute using our formulas is $\mathbb{P}(\hat{\mathcal{A}}(\mathbf{x},\mathbf{y})\leq\mathfrak{f}(\mathbf{x})+\mathfrak{g}(\mathbf{y}),\ \mathbf{x},\mathbf{y}\in\mathbb{R})=\det\Big(\mathbf{I}-\mathbf{K}_1^{\mathrm{hypo}(-\mathfrak{g})}\mathbf{K}_{-1}^{\mathrm{epi}(\mathfrak{f})}\Big)$. Suppose we want to compute the two-point distribution for the Airy sheet $\mathbb{P}(\hat{\mathcal{A}}(\mathbf{x}_i,\mathbf{y}_i)\leq\mathbf{a}_i,\ i=1,2)$ from this. We would need to choose $\mathfrak{f}$ and $\mathfrak{g}$ taking two non-infinite values, which yields a formula for $\mathbb{P}(\hat{\mathcal{A}}(\mathbf{x}_i,\mathbf{y}_j)\leq\mathfrak{f}(\mathbf{x}_i)+\mathfrak{g}(\mathbf{y}_j),\ i,j=1,2)$, and thus we need to take $\mathfrak{f}(\mathbf{x}_1)+\mathfrak{g}(\mathbf{y}_1)=\mathbf{a}_1$, $\mathfrak{f}(\mathbf{x}_2)+\mathfrak{g}(\mathbf{y}_2)=\mathbf{a}_2$ and $\mathfrak{f}(\mathbf{x}_1)+\mathfrak{g}(\mathbf{y}_2)=\mathfrak{f}(\mathbf{x}_2)+\mathfrak{g}(\mathbf{y}_1)=L$ with $L\to\infty$. But $\{\mathfrak{f}(\mathbf{x}_i)+\mathfrak{g}(\mathbf{y}_j),\ i,j=1,2\}$ only spans a 3-dimensional linear subspace of $\mathbb{R}^4$, so this is not possible.

The preservation of max property allows us to write an important variational formula for the KPZ fixed point in terms of the Airy sheet (analogous to Hopf's formula for certain Hamilton-Jacobi equations), which was conjectured in [CQR15]:

Theorem 2.15. (Airy sheet variational formula)

$$\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{h}_0)=\sup_{\mathbf{y}}\big\{\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{d}_{\mathbf{y}})+\mathfrak{h}_0(\mathbf{y})\big\}\overset{\mathrm{dist}}{=}\sup_{\mathbf{y}}\Big\{\mathbf{t}^{1/3}\hat{\mathcal{A}}(\mathbf{t}^{-2/3}\mathbf{x},\mathbf{t}^{-2/3}\mathbf{y})+\mathfrak{h}_0(\mathbf{y})\Big\}.\tag{2.1}$$

In particular, the Airy sheet satisfies the semi-group property*: If $\hat{\mathcal{A}}^1$ and $\hat{\mathcal{A}}^2$ are independent copies and $\mathbf{t}_1+\mathbf{t}_2=\mathbf{t}$ are all positive, then*

$$\sup_{\mathbf{z}}\Big\{\mathbf{t}_1^{1/3}\hat{\mathcal{A}}^1(\mathbf{t}_1^{-2/3}\mathbf{x},\mathbf{t}_1^{-2/3}\mathbf{z})+\mathbf{t}_2^{1/3}\hat{\mathcal{A}}^2(\mathbf{t}_2^{-2/3}\mathbf{z},\mathbf{t}_2^{-2/3}\mathbf{y})\Big\}\overset{\mathrm{dist}}{=}\mathbf{t}^{1/3}\hat{\mathcal{A}}^1(\mathbf{t}^{-2/3}\mathbf{x},\mathbf{t}^{-2/3}\mathbf{y}).$$

Proof. Let $\mathfrak{h}_0^n$ be a sequence of initial conditions taking finite values $\mathfrak{h}_0^n(\mathbf{y}_i^n)$ at $\mathbf{y}_i^n$, $i=1,\dots,k_n$, and $-\infty$ everywhere else, which converges to $\mathfrak{h}_0$ in UC as $n\to\infty$. By repeated application of Prop. 2.8(v) (and the easy fact that $\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{h}_0+\mathbf{a})=\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{h}_0)+\mathbf{a}$ for $\mathbf{a}\in\mathbb{R}$) we get

$$\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{h}_0^n)=\sup_{i=1,\dots,k_n}\big\{\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{d}_{\mathbf{y}_i^n})+\mathfrak{h}_0^n(\mathbf{y}_i^n)\big\},$$

and taking $n\to\infty$ yields the result (the second equality in (2.1) follows from scaling invariance, Proposition 2.7). □

One of the interests in this variational formula is that it leads to proofs of properties of the fixed point, as we already mentioned in earlier sections.

Proof of affine invariance, Proposition 2.8(iv). The fact that the fixed point is invariant under translations of the initial data is straightforward, so we may assume $\mathbf{a}=0$. By Theorem 2.15 we have

$$\begin{aligned}\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{h}_0+c\mathbf{x})&\overset{\mathrm{dist}}{=}\sup_{\mathbf{y}}\Big\{\mathbf{t}^{1/3}\mathcal{A}(\mathbf{t}^{-2/3}\mathbf{x},\mathbf{t}^{-2/3}\mathbf{y})-\mathbf{t}^{-1}(\mathbf{x}-\mathbf{y})^2+\mathfrak{h}_0(\mathbf{y})+c\mathbf{y}\Big\}\\&=\sup_{\mathbf{y}}\Big\{\mathbf{t}^{1/3}\mathcal{A}(\mathbf{t}^{-2/3}\mathbf{x},\mathbf{t}^{-2/3}(\mathbf{y}+c\mathbf{t}/2))-\mathbf{t}^{-1}(\mathbf{x}-\mathbf{y})^2+\mathfrak{h}_0(\mathbf{y}+c\mathbf{t}/2)+c\mathbf{x}+c^2\mathbf{t}/4\Big\}\\&\overset{\mathrm{dist}}{=}\sup_{\mathbf{y}}\Big\{\mathbf{t}^{1/3}\hat{\mathcal{A}}(\mathbf{t}^{-2/3}\mathbf{x},\mathbf{t}^{-2/3}\mathbf{y})+\mathfrak{h}_0(\mathbf{y}+c\mathbf{t}/2)+c\mathbf{x}+c^2\mathbf{t}/4\Big\}\\&=\mathfrak{h}(\mathbf{t},\mathbf{x};\mathfrak{h}_0(\mathbf{x}+c\mathbf{t}/2))+c\mathbf{x}+c^2\mathbf{t}/4.\end{aligned}$$

□

Sketch of the proof of time regularity, Theorem 2.6. Fix $\alpha<1/3$ and choose $\beta<1/2$ so that $\beta/(2-\beta)=\alpha$. By the Markov property it is enough to assume that $\mathfrak{h}_0\in\mathscr{C}^\beta$ and check the Hölder-α regularity at time 0. By space regularity of the Airy$_2$ process (proved in [QR13], but which also follows from Theorem 2.4) there is an $R < \infty$ a.s. such that $|\mathcal{A}_2(\mathbf{x})| \le R(1 + |\mathbf{x}|^\beta)$, and making R larger if necessary we may also assume $|\mathfrak{h}_0(\mathbf{x}) - \mathfrak{h}_0(\mathbf{x}_0)| \le R(|\mathbf{x} - \mathbf{x}_0|^\beta + |\mathbf{x} - \mathbf{x}_0|)$. From the variational formula (2.1), $|\mathfrak{h}(\mathbf{t}, \mathbf{x}_0) - \mathfrak{h}(0, \mathbf{x}_0)|$ is then bounded by

$$\sup_{\mathbf{x} \in \mathbb{R}} \Big(R(|\mathbf{x} - \mathbf{x}_0|^\beta + |\mathbf{x} - \mathbf{x}_0| + \mathbf{t}^{1/3} + \mathbf{t}^{(1-2\beta)/3} |\mathbf{x}|^\beta) - \tfrac{1}{\mathbf{t}} (\mathbf{x}_0 - \mathbf{x})^2 \Big).$$

The supremum is attained roughly at $x - x_0 = \mathbf{t}^{-\eta}$ with η such that $|\mathbf{x} - \mathbf{x}_0|^\beta \sim \frac{1}{\mathbf{t}}(\mathbf{x}_0 - \mathbf{x})^2$. Then $\eta = 1/(2 - \beta)$ and the supremum is bounded by a constant multiple of $\mathbf{t}^{\beta/(2-\beta)} = \mathbf{t}^\alpha$, as desired. □

3. Full solution for TASEP

Our goal is to obtain a full solution for TASEP (by which we mean an explicit formula for its finite dimensional distributions) which is suitable for asymptotics in the form needed in (1.5). We begin with a more precise description of the process.

3.1. **The model.**

Definition 3.1. (TASEP) The *totally asymmetric simple exclusion process* consists of particles at positions $\cdots < X_t(2) < X_t(1) < X_t(0) < X_t(-1) < X_t(-2) < \cdots$ on $\mathbb{Z} \cup \{-\infty, \infty\}$ performing totally asymmetric nearest neighbour random walks with exclusion: each particle independently attempts jumps to the neighbouring site to the right at rate 1, the jump being allowed only if that site is unoccupied (see [Lig85] for the non-trivial fact that the process with an infinite number of particles makes sense).

We follow the standard practice of ordering particles from the right; for right-finite data the rightmost particle is labelled 1, unless indicated otherwise. Let

$$X_t^{-1}(u) = \min\{k \in \mathbb{Z} : X_t(k) \le u\}$$

denote the label of the rightmost particle which sits to the left of, or at, u at time t. The *TASEP height function* associated to X_t is given for $z \in \mathbb{Z}$ by

$$h_t(z) = -2\big(X_t^{-1}(z-1) - X_0^{-1}(-1)\big) - z, \tag{3.1}$$

which fixes $h_0(0) = 0$. This is the height function which we already introduced in Section 1.2 (one can check directly that the dynamics induced on h_t by the particle dynamics coincide exactly with the dynamics specified there). We will find formulas for the finite dimensional distributions of X_t; (3.1) will allow us to use those formulas to compute the limit (1.5).

3.2. **The master equation and Schütz's formula**[3]**.**

The first step is to solve the master equation for TASEP in order to obtain a formula for its transition probabilities. We will work only in the case where TASEP starts with a finite number $N \ge 2$ of particles.

Define the *Weyl chamber* $\Omega_N = \{(x_1, \dots, x_N) \in \mathbb{Z}^N : x_1 > \cdots > x_N\}$ and for $x, y \in \Omega_N$ consider the transition probabilities

$$P_t^{(N)}(x, y) = \mathbb{P}(X_t = x | X_0 = y).$$

Then $P_t^{(N)}$ satisfies the *master equation* (or *Kolmogorov forward equation*)

$$\frac{d}{dt} P_t^{(N)}(x, y) = \big(\mathcal{L}^{(N)}\big)^* P_t^{(N)}(x, y), \qquad P_0^{(N)}(x, y) = \mathbf{1}_{y=x} \tag{3.2}$$

where the infinitesimal generator of the process $\mathcal{L}^{(N)}$ is given, for $F \colon \Omega_N \longrightarrow \mathbb{R}$, by

$$\mathcal{L}^{(N)} F(x) = \sum_{k=1}^{N} \mathbf{1}_{x_{k-1} - x_k > 1} \nabla_k F(x),$$

[3]The reader may choose to skip directly to Section 3.4; this and the next sections are not strictly needed in the derivation of the fixed point.

with $x_0 = \infty$ and $\nabla_k F(x) = F(x_1 \dots, x_k+1, \dots, x_N) - F(x)$.

The master equation for TASEP was solved by Schütz [Sch97]. He used the Bethe ansatz [Bet31], which in our case means the idea of rewriting (3.2) as an equation without the exclusion constraint (that is, the factor $\mathbf{1}_{x_k > x_{k+1}}$) together with suitable boundary conditions. Explicitly, if for fixed $y \in \Omega_N$ the function $u_t^{(N)} : \mathbb{Z}^N \longrightarrow \mathbb{R}$ solves

$$\frac{d}{dt} u_t^{(N)} = \sum_{k=1}^{N} \nabla_k^* u_t^{(N)}, \qquad u_0^{(N)}(x) = \mathbf{1}_{y=x}, \tag{3.3}$$

(where $\nabla_k^* F(x) = F(x_1 \dots, x_k - 1, \dots, x_N) - F(x)$) with the *boundary conditions*

$$\nabla_k^* u_t^{(N)}(x) = 0 \quad \text{when} \quad x_k = x_{k+1} + 1, \tag{3.4}$$

then when restricted to the Weyl chamber, $u_t^{(N)}$ coincides with $P_t^{(N)}$, that is

$$P_t^{(N)}(x, y) = u_t^{(N)}(x) \quad \forall\, x \in \Omega_N.$$

Theorem 3.2. (Schütz's formula [Sch97]**)**

$$P_t^{(N)}(x, y) = \det(F_{i-j}(t, x_{N+1-i} - y_{N+1-j}))_{1 \le i,j \le N} \tag{3.5}$$

with

$$F_n(t, x) = \frac{(-1)^n}{2\pi \mathrm{i}} \oint_{\Gamma_{0,1}} dw\, \frac{(1-w)^{-n}}{w^{x-n+1}} e^{t(w-1)}, \tag{3.6}$$

where $\Gamma_{0,1}$ is any simple loop oriented anticlockwise which includes $w = 0$ and $w = 1$.

The argument of [Sch97] shows how this remarkable solution can be derived, and the method turns out to work for other similar processes (see for instance [BFP07; BF08; BFS08a]). However, once one has the explicit solution it is not hard to check that it satisfies (3.3)-(3.4), so this is the proof we present[4].

Proof of Theorem 3.2. The proof is based only on the following identities, which follow directly from (3.6):

$$\partial_t F_n(t, x) = \nabla^* F_n(t, x), \qquad F_n(t, x) = -\nabla F_{n+1}(t, x), \tag{3.7}$$

where $\nabla F(x) = F(x+1) - F(x)$, $\nabla^* F(x) = F(x-1) - F(x)$.

Define the column vectors

$$H_i(t, x) = \big(F_{i-1}(t, x - y_N), \cdots, F_{i-N}(t, x - y_1)\big)^{\mathsf{T}}.$$

Then, denoting by $u_t^{(N)}(x)$ the right-hand side of (3.5), we can write

$$\begin{aligned}\partial_t u_t^{(N)}(x) &= \sum_{k=1}^{N} \det\big[\dots, \partial_t H_k(t, x_{N+1-k}), \dots\big] = \sum_{k=1}^{N} \det\big[\dots, \nabla^* H_k(t, x_{N+1-k}, t), \dots\big] \\ &= \sum_{k=1}^{N} \nabla_k^* \det\big[F_{i-j}(t, x_{N+1-i} - y_{N+1-j})\big],\end{aligned}$$

where we used the multilinearity of the determinant. This gives the first equation in (3.3).

To get the boundary conditions (3.4), take $x_k = x_{k+1} + 1$ and use again the multilinearity of the determinant to get

$$\nabla_k^* \det\big[H_1(t, x_N), \dots, H_N(t, x_1)\big] = \det\big[\dots, H_{N-k}(t, x_{k+1}), \nabla^* H_{N+1-k}(t, x_k), \dots\big] = 0$$

because $\nabla^* H_{N+1-k}(t, x_k) = -\nabla H_{N+1-k}(t, x_k - 1) = -\nabla H_{N+1-k}(t, x_k) = H_{N-k}(t, x_k)$.

[4]We thank K. Matetski for this short proof, which to the best of our knowledge is not available in this exact form in the literature.

To get the initial condition we write first, $n \geq 0$,

$$F_{-n}(0,x) = \frac{(-1)^n}{2\pi\mathrm{i}} \oint_{\Gamma_0} dw\, \frac{(1-w)^n}{w^{x+n+1}},$$

which in particular implies that $F_{-n}(0,x) = 0$ for $x < -n$ and $x > 0$, and $F_0(0,x) = \mathbf{1}_{x=0}$. In the case $x_N < y_N$, since $y \in \Omega_N$ we have $x_N < y_k$ for all $k = 1,\dots,N-1$ and $x_N - y_{N+1-j} < 1-j$. This yields $F_{1-j}(x_N - y_{N+1-j}, 0) = 0$ and thus the determinant in (3.5) vanishes, because the matrix contains a row of zeros. If $x_N \geq y_N$ then we have $x_k > y_N$ for all $k = 1,\dots,N-1$, and all entries of the first column in the matrix from (3.5) vanish, except the first entry which equals $\mathbf{1}_{x_N=y_N}$. Repeating this argument for x_{N-1}, x_{N-2} and so on, we see that the matrix is upper-triangular with indicator functions $\mathbf{1}_{X_{N-i+1}=y_{N-i+1}}$ in the diagonal, which gives us the claim. □

3.3. **The non-intersecting line ensemble representation.** Schütz's formula itself is not well-suited for asymptotics, but the remarkable fact that it is given by a determinant opens up an avenue for its treatment. In a breakthrough by Sasamoto [Sas05], which was pursued and extended rigorously in [BFPS07], he realized that the finite dimensional distributions for TASEP can be expressed in terms of a "signed" non-intersecting line ensemble, as follows. Let GT_N be the set of *Gelfand-Tsetlin patterns* $\bar{x}$ given as

$$\mathrm{GT}_N = \left\{\bar{x}_i^j \in \mathbb{Z},\, 1 \leq i \leq j \leq N \colon \bar{x}_i^{j+1} < \bar{x}_i^j \leq \bar{x}_{i+1}^{j+1}\right\}$$

(see Figure 2). For $y \in \Omega_N$ define the (signed) weight

$$W_N(\bar{x}; y) = \Big(\prod_{n=2}^{N} \det\big(\phi(\bar{x}_i^{n-1}, \bar{x}_j^n)\big)_{1\leq i,j\leq n}\Big) \det\big(F_{-j}(t, \bar{x}_{i+1}^N - y_{N-j})\big)_{0\leq i,j<N} \tag{3.8}$$

with $\phi(x_1,x_2) = \mathbf{1}_{x_1>x_2}$. W_N defines a signed measure on GT_N. We think of it as describing, after normalization, the evolution of $\big((\bar{x}_i^j)_{i=1,\dots,j}\big)_{j=1,\dots,N}$ (of course this makes no sense since the measure is not positive; it is meant only as intuition). We are thinking now of j as time, which has nothing to do with the "real" TASEP time, which is just the parameter t in the formula. The statement can then be expressed as follows:

Theorem 3.3 ([BFPS07])**.** *For $x, y \in \Omega_N$,*

$$P_t^{(N)}(x,y) \;=\; \sum_{\bar{x}\in GT_N\colon x_1^i = x_i,\, i=1,\dots,N} W_N(\bar{x}; y).$$

In other words, TASEP at time t can be identified with the evolution of the first particle in the Gelfand-Tsetlin pattern $(\bar{x}_1^1,\dots,\bar{x}_1^N)$ (that is, the leftmost diagonal in Figure 2), and $P_t^{(N)}$ can be obtained as a (signed) marginal of W_N. The proof of this result is based on applying the second equality of (3.7) and the multilinearity of the determinant repeatedly on Schütz's formula and then employing a clever symmetrization argument, see [BFPS07] for the details.

3.4. **Biorthogonal ensembles.** The key point is that from the form of (3.8), [Sas05; BFPS07] could recognize that the correlation functions associated to the "random" Gelfand-Tsetlin pattern $\bar{x}$ should be determinantal, and that one could then apply a suitable version of the Eynard-Mehta theorem [EM98] (see [BFPS07, Lem. 3.4]) to obtain a Fredholm determinant formula for the finite dimensional distributions of TASEP (this step is crucial, and far from trivial, but we refer to [BFPS07] for the details). We describe the result next.

Definition 3.4. Consider a decreasing sequence of integers $\big(X_0(k)\big)_{k\geq 1}$ (the TASEP initial condition). Given a fixed $n \in \mathbb{Z}_{>0}$, we define two families of functions on the integers $\big(\Psi_k^n\big)_{k\leq n-1}$ and $\big(\Phi_k^n\big)_{k=0,\dots,n-1}$ as follows:

$$\Psi_k^n(x) = \frac{1}{2\pi\mathrm{i}} \oint_{\Gamma_0} dw\, \frac{(1-w)^k}{2^{x-X_0(n-k)} w^{x+k+1-X_0(n-k)}} e^{t(w-1)} \tag{3.9}$$

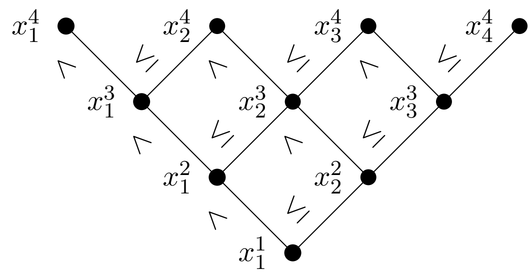

FIGURE 2. Visualization of a Gelfand-Tsetlin pattern in GT_4.

where Γ_0 is any simple loop, anticlockwise oriented and contained in D, which includes $w = 0$ but does not include $w = 1$, while the functions $\Phi_k^n(x)$, $k = 0, \dots, n-1$, are defined implicitly by

(1) The *biorthogonality relation* $\sum_{x\in\mathbb{Z}} \Psi_k^n(x)\Phi_\ell^n(x) = \mathbf{1}_{k=\ell}$;
(2) $2^{-x}\Phi_k^n(x)$ is a polynomial of degree at most $n-1$ in x.

Additionally, we define the following stochastic matrix:

Definition 3.5. Let

$$Q(x,y) = \frac{1}{2^{x-y}}\mathbf{1}_{x>y}.$$

We have, for $m \geq 1$,

$$Q^m(x,y) = \frac{1}{2^{x-y}}\binom{x-y-1}{m-1}\mathbf{1}_{x\geq y+m}.$$

Moreover Q and Q^m are invertible:

$$Q^{-1}(x,y) = 2\cdot\mathbf{1}_{x=y-1} - \mathbf{1}_{x=y}, \qquad Q^{-m}(x,y) = (-1)^{y-x+m}2^{y-x}\binom{m}{y-x}. \tag{3.10}$$

A crucial identity is that for all $m, n \in \mathbb{Z}$

$$Q^{n-m}\Psi_{n-k}^n = \Psi_{m-k}^m. \tag{3.11}$$

Theorem 3.6. (Biorthogonal ensemble formula for TASEP [BFPS07]**)** *Suppose that TASEP starts with particles labeled $1, 2, \dots$ (so that, in particular, there is a rightmost particle) and let $1 \leq n_1 < n_2 < \cdots < n_M \leq N$. Then for $t > 0$ we have*

$$\mathbb{P}(X_t(n_j) \geq a_j,\ j = 1, \dots, M) = \det(I - \bar{\chi}_a K_t \bar{\chi}_a)_{\ell^2(\{n_1,\dots,n_M\}\times\mathbb{Z})} \tag{3.12}$$

where

$$K_t(n_i, x_i; n_j, x_j) = -Q^{n_j-n_i}(x_i, x_j) + \sum_{k=1}^{n_j} \Psi_{n_i-k}^{n_i}(x_i)\Phi_{n_j-k}^{n_j}(x_j), \tag{3.13}$$

with the Ψ_k^n's and the Φ_k^n's given as in Definition 3.4.

Remark 3.7.

1. We are assuming in the theorem that $X_0(j) < \infty$ for all $j \geq 1$; particles at $-\infty$ are allowed.
2. The [BFPS07] result is stated only for initial conditions with finitely many particles, but the extension to right-finite (infinite) initial conditions is straightforward because, given fixed indices $n_1 < n_2 < \cdots < n_M$, the distribution of $\big(X_t(n_1), \dots, X_t(n_M)\big)$ does not depend on the initial positions of the particles with indices beyond n_M.
3. In (3.13) we have conjugated the kernel K_t from [BFPS07] by 2^x for convenience. The additional $X_0(n-k)$ in the power of 2 in the Ψ_k^n's is there also for convenience and is allowed because it just means that the Φ_k^n's have to be multiplied by $2^{X_0(n-k)}$.

Our goal is to find the Φ^n_k's explicitly for any initial data. These functions had been computed previously only for very special initial conditions.

3.5. **TASEP path integral kernel.** Formula (3.12) is what we called an extended kernel formula after Theorem 1.19. There is aso a path integral kernel formula, which is useful to explain the intuition behind the next step in our derivation. It is given as follows:

$$\begin{aligned}\mathbb{P}(X_t(n_j) \geq a_j,\ j = 1, \dots, M) \\ = \det\big(I - K_t^{(n_m)}(I - Q^{n_1-n_m}\chi_{a_1}Q^{n_2-n_1}\chi_{a_2}\cdots Q^{n_m-n_{m-1}}\chi_{a_m})\big)_{L^2(\mathbb{R})},\end{aligned} \tag{3.14}$$

where

$$K_t^{(n)} = K_t(n, \cdot; n, \cdot).$$

This formula follows from the framework of [BCR15] (or rather a minor variation of it, see [MQR17b, App. A.2]) applied to (3.12).

TASEP satisfies the skew time reversibility property

$$\mathbb{P}_f(h_t(x) \leq g(x),\ \ x \in \mathbb{Z}) = \mathbb{P}_{-g}(h_t(x) \leq -f(x),\ \ x \in \mathbb{Z}), \tag{3.15}$$

the subscript indicating the initial data. In other words, the height function evolving backwards in time is indistiguishable from minus the height function. Now suppose we have the solution (3.13) for step initial data centered at x_0, which means h_0 is the peak $-|x - x_0|$ (we actually do have this solution, it is the starting point in the computation of the limit (1.6)). The multipoint distribution at time t is given by (3.14), but we can use (3.15) to reinterpret it as the one-point distribution of h_t at x_0, starting from an (arbitrary) series of peaks, and this gives us directly the one-point distribution for TASEP with any initial condition.

At first sight it looks like we are done, since we can compute now the multipoint kernel from the one-point kernel (using (3.11) and (3.13)) through

$$Q^{n_j-n_i}K_t^{(n_j)} = \sum_{k=0}^{n_j-1} \Psi^{n_i}_{n_i-n_j+k} \otimes \Phi^{n_j}_k = K_t(n_i, \cdot; n_j, \cdot) + Q^{n_j-n_i}\mathbf{1}_{n_i<n_j}.$$

But notice that we have used the distributional identity (3.15) to obtain the one-point formula, so we actually don't have yet a formula for $K_t^{(n)}$ (but only an equality of Fredholm determinants). Instead of trying to use these facts directly, it turns out to be easier to use them to guess the formula for the Φ^n_k's.

The key is to recognize the kernel Q as the transition probabilities of a random walk (which is why we conjugated the [BFPS07] kernel by 2^x) and then $\chi_{a_1}Q^{n_2-n_1}\chi_{a_2}\cdots Q^{n_m-n_{m-1}}\chi_{a_m}(x, y)$ as the probability that this walk goes from x to y in $n_m - n_1$ steps, staying above a_1 at time n_1, above a_2 at time n_2, etc. Based on this intuition we obtained in [MQR17b] a formula for Φ^n_k in terms of the solution of certain boundary value problem for a backwards heat equation involving Q. This is the content of the next result.

3.6. **Explicit biorthogonalization.**

Definition 3.8. Let

$$R_t(x, y) = \frac{1}{2\pi \mathrm{i}}\oint_{\Gamma_0} dw\, \frac{e^{t(w-1)}}{2^{x-y}w^{x-y+1}} = e^{-t}\frac{t^{x-y}}{2^{x-y}(x-y)!}\mathbf{1}_{x\geq y}. \tag{3.16}$$

R_t is invertible, with

$$R_t^{-1}(x, y) = \frac{1}{2\pi \mathrm{i}}\oint_{\Gamma_0} dw\, \frac{e^{t(1-w)}}{2^{x-y}w^{x-y+1}} = e^{t}\frac{(-t)^{x-y}}{2^{x-y}(x-y)!}\mathbf{1}_{x\geq y}. \tag{3.17}$$

Note that by Cauchy's residue theorem we have $\Psi^n_0 = R_t\delta_{X_0(n)}$, where $\delta_y(x) = \mathbf{1}_{x=y}$, so from (3.11) we have

$$\Psi^n_k = R_tQ^{-k}\delta_{X_0(n-k)}. \tag{3.18}$$

Theorem 3.9. *Fix $0 \le k < n$ and consider particles at $X_0(1) > X_0(2) > \cdots > X_0(n)$. Let $h^n_k(\ell, z)$ be the unique solution to the initial–boundary value problem for the backwards heat equation*

$$\begin{cases} (Q^*)^{-1} h^n_k(\ell, z) = h^n_k(\ell+1, z) & \ell < k,\ z \in \mathbb{Z}; \quad (3.19a) \\ h^n_k(k, z) = 2^{z - X_0(n-k)} & z \in \mathbb{Z}; \quad (3.19b) \\ h^n_k(\ell, X_0(n-k)) = 0 & \ell < k. \quad (3.19c) \end{cases}$$

Then

$$\Phi^n_k(z) = (R^*_t)^{-1} h^n_k(0, \cdot)(z) = \sum_{y \in \mathbb{Z}} h^n_k(0, y) R_t{-}1(y, z).$$

Here $Q^*(x, y) = Q(y, x)$ is the kernel of the adjoint of Q (and likewise for R^*_t).

Remark 3.10. It is not true that $Q^* h^n_k(\ell+1, z) = h^n_k(\ell, z)$. In fact, in general $Q^* h^n_k(k, z)$ is divergent.

Before the proof we need:

Lemma 3.11. *With the definitions in the above theorem, $2^{-x} h^n_k(0, x)$ is a polynomial of degree at most k.*

Proof. We proceed by induction. By (3.19b), $2^{-x} h^n_k(k, x)$ is a polynomial of degree 0. Assume now that $\tilde{h}^n_k(\ell, x) = 2^{-x} h^n_k(\ell, x)$ is a polynomial of degree at most $k - \ell$ for some $0 < \ell \le k$. By (3.19a) and (3.10) we have

$$\tilde{h}^n_k(\ell, y) = 2^{-y} (Q^*)^{-1} h^n_k(\ell-1, y) = \tilde{h}^n_k(\ell-1, y-1) - \tilde{h}^n_k(\ell-1, y) \tag{3.20}$$

Taking $x \ge X_0(n-\ell+1)$ and summing (3.20) gives $\tilde{h}^n_k(\ell-1, x) = -\sum_{y=X_0(n-\ell+1)+1}^{x} \tilde{h}^n_k(\ell, y)$ thanks to (3.19c), which by the inductive hypothesis is a polynomial of degree at most $k-\ell+1$ in x. Similarly, taking $x < X_0(n-\ell+1)$ we get $\tilde{h}^n_k(\ell-1, x) = \sum_{y=x+1}^{X_0(n-\ell+1)} \tilde{h}^n_k(\ell, y)$, which again is a polynomial of degree at most $k-\ell+1$. The two polynomials are the same, as can be checked for instance from Faulhaber's formula, whence the claim follows. □

Proof of Theorem 3.9. Note first that the dimension of $\ker(Q^*)^{-1}$ is 1, and it consists of the function 2^z. This allows us to march forwards from the initial condition $h^n_k(k, z) = 2^{z - X_0(n-k)}$ uniquely solving the boundary value problem $h^n_k(\ell, X_0(n-k)) = 0$ at each step and thus get the existence and uniqueness. The details are left as an exercise.

We need to show that the proposed Φ^n_k's satisfy conditions (1) and (2) of Definition 3.4. We check the biorthogonality first. Using (3.18) we get

$$\begin{aligned} \sum_{z \in \mathbb{Z}} \Psi^n_\ell(z) \Phi^n_k(z) &= \sum_{z_1, z_2 \in \mathbb{Z}} \sum_{z \in \mathbb{Z}} R_t(z, z_1) Q^{-\ell}(z_1, X_0(n-\ell)) h^n_k(0, z_2) R_t^{-1}(z_2, z) \\ &= \sum_{z \in \mathbb{Z}} Q^{-\ell}(z, X_0(n-\ell)) h^n_k(0, z) = (Q^*)^{-\ell} h^n_k(0, X_0(n-\ell)) \end{aligned}$$

(exercise: justify the application of Fubini here). We want to show that the last expression equals $\mathbf{1}_{k=\ell}$. Consider first $\ell \le k$. Then we may use the boundary condition $h^n_k(\ell, X_0(n-k)) = \mathbf{1}_{k=\ell}$, which is both (3.19b) and (3.19c), to get

$$(Q^*)^{-\ell} h^n_k(0, X_0(n-\ell)) = h^n_k(\ell, X_0(n-\ell)) = \mathbf{1}_{k=\ell}.$$

Next for $\ell > k$ we use (3.19a) and $2^z \in \ker (Q^*)^{-1}$:

$$(Q^*)^{-\ell} h^n_k(0, X_0(n-\ell)) = (Q^*)^{-(\ell-k-1)} (Q^*)^{-1} h^n_k(k, X_0(n-\ell)) = 0.$$

Next we show that $2^{-x} \Phi^n_k(x)$ is a polynomial of degree at most k in x. By (3.16) we have

$$2^{-x} \Phi^n_k(x) = \sum_{y \ge 0} e^{-t} \frac{t^y}{y!} 2^{-(x+y)} h^n_k(0, x+y).$$

This sum is absolutely convergent, so we may compute the $(k+1)$-th derivate in x of the right hand side by taking the derivative inside the sum and use the fact that $2^{-z}h_k^n(0,z)$ is a polynomial of degree at most k to deduce that $\frac{d^{k+1}}{dx^{k+1}}[2^{-x}\Phi_k^n(x)] = 0$ as desired. □

3.7. **Representation of the kernel below the curve as a transition probability with hitting.**

Definition 3.12. (Geometric random walks) We denote by B_m^* a random walk with transition probabilities given by Q^* (that is, B_m^* has Geom$[\frac{1}{2}]$ jumps strictly to the right). We also introduce, for $0 \le \ell \le k \le n-1$, the stopping times

$$\tau^{\ell,n} = \min\{m \in \{\ell,\dots,n-1\} \colon B_m^* > X_0(n-m)\},$$

with the convention that $\min\emptyset = \infty$.

Similarly, we denote by B_m a random walk with transition probabilities given by Q (that is, B_m has Geom$[\frac{1}{2}]$ jumps strictly to the left), and introduce the stopping time

$$\tau = \min\{m \ge 0 : B_m > X_0(m+1)\}. \tag{3.21}$$

Lemma 3.13. *For $z \le X_0(n-\ell)$ we have*

$$h_k^n(\ell,z) = \mathbb{P}_{B_{\ell-1}^*=z}\big(\tau^{\ell,n} = k\big).$$

Proof. It is easy to see that $h_k^n(\ell,z)$ satisfies (3.19b) and (3.19c). On the other hand, for $z \le X_0(n-\ell-1)$ it also satisfies (3.19a) and it is given by 2^z times a polynomial in z of degree at most $n-1$. Exercise: conclude the proof by using Lemma 3.11. □

Definition 3.14. We write

$$G_{0,n}(z_1,z_2) = \sum_{k=0}^{n-1} Q^{n-k}(z_1, X_0(n-k))h_k^n(0,z_2),$$

with h_k^n the solution of (3.19), so that

$$K_t^{(n)} = R_t Q^{-n} G_{0,n} R_t^{-1}.$$

Lemma 3.15. *For $z_2 \le X_0(n)$,*

$$G_{0,n}(z_1,z_2) = \mathbb{P}_{B_{-1}^*=z_2}\big(\tau^{0,n} < n,\, B_{n-1}^* = z_1\big) \tag{3.22}$$

(which is the probability for the walk starting at z_2 at time -1 to end up at z_1 after n steps, having hit the curve $\big(X_0(n-m)\big)_{m=0,\dots,n-1}$ in between).

Proof. From the memoryless property of the geometric distribution we have for all $z \le X_0(n-k)$ that

$$\mathbb{P}_{B_{-1}^*=z}\big(\tau^{0,n} = k,\, B_k^* = y\big) = 2^{X_0(n-k)-y}\mathbb{P}_{B_{-1}^*=z}\big(\tau^{0,n} = k\big), \tag{3.23}$$

and as a consequence we get, for $z_2 \le X_0(n)$,

$$\begin{aligned}
G_{0,n}(z_1,z_2) &= \sum_{k=0}^{n-1} \mathbb{P}_{B_{-1}^*=z_2}\big(\tau^{0,n} = k\big)(Q^*)^{n-k}(X_0(n-k), z_1) \\
&= \sum_{k=0}^{n-1} \sum_{z > X_0(n-k)} \mathbb{P}_{B_{-1}^*=z_2}\big(\tau^{0,n} = k,\, B_k^* = z\big)(Q^*)^{n-k-1}(z, z_1) \\
&= \mathbb{P}_{B_{-1}^*=z_2}\big(\tau^{0,n} < n,\, B_{n-1}^* = z_1\big).
\end{aligned}$$

□

3.8. **Polynomial extension above the curve.** To finish our derivation of our TASEP solution we need to extend the result of Lemma 3.15 to all values of z_2. We will do this by extending our formula polynomally.

Definition 3.16. (Polynomial extension of Q^n) Let

$$\mathcal{Q}^{(n)}(y_1,y_2)=\frac{1}{2\pi\mathrm{i}}\oint_{\Gamma_0}dw\,\frac{(1+w)^{y_1-y_2-1}}{2^{y_1-Ty_2}w^n}=\frac{(y_1-y_2-1)_{n-1}}{2^{y_1-y_2}(n-1)!},\tag{3.24}$$

where $(x)_k=x(x-1)\cdots(x-k+1)$ for $k>0$ and $(x)_0=1$ is the *Pochhammer symbol.* For each fixed y_1, $2^{-y_2}Q^n(y_1,y_2)$ extends in y_2 to a polynomial $2^{-y_2}\mathcal{Q}^{(n)}(y_1,y_2)$ of degree $n-1$. We have

$$\mathcal{Q}^{(n)}(y_1,y_2)=Q^n(y_1,y_2)\qquad\text{for}\quad y_1-y_2\geq 1\tag{3.25}$$

and

$$Q^{-1}\mathcal{Q}^{(n)}=\mathcal{Q}^{(n)}Q^{-1}=\mathcal{Q}^{(n-1)}\quad\text{for}\quad n>1,\qquad\text{but}\quad Q^{-1}\mathcal{Q}^{(1)}=\mathcal{Q}^{(1)}Q^{-1}=0.$$

Remark 3.17. $\mathcal{Q}^{(n)}\mathcal{Q}^{(m)}$ is divergent, so the $\mathcal{Q}^{(n)}$ are no longer a group like the Q^n.

Lemma 3.18. *For all $z_1,z_2\in\mathbb{Z}$ we have, for τ as in Definition 3.12,*

$$G_{0,n}(z_1,z_2)=\mathbb{E}_{B_0=z_1}\Big[\mathcal{Q}^{(n-\tau)}(B_\tau,z_2)\mathbf{1}_{\tau<n}\Big].$$

Proof. From Lemma 3.11 and its definition, $2^{-z_2}G_{0,n}(z_1,z_2)$ is a polynomial in z_2 for every fixed z_1, and thus it is enough to check that the right hand side of (3.18) is a polynomial in z_2 and coincides with (3.22) for all $z_2\leq X_0(n)$.

From (3.22) we have, for $z_2\leq X_0(n)$,

$$\begin{aligned}G_{0,n}(z_1,z_2)&=\mathbb{P}_{B^*_{-1}=z_2}\big(\tau^{0,n}\leq n-1,\,B^*_{n-1}=z_1\big)=\mathbb{P}_{B_0=z_1}\big(\tau\leq n-1,B_n=z_2\big)\\&=\sum_{k=0}^{n-1}\sum_{z>X_0(k+1)}\mathbb{P}_{B_0=z_1}\big(\tau=k,\,B_k=z\big)Q^{n-k}(z,z_2)=\mathbb{E}_{B_0=z_1}\big[Q^{n-\tau}\big(B_\tau,z_2\big)\mathbf{1}_{\tau<n}\big].\end{aligned}\tag{3.26}$$

Note that we reversed the direction of the walk in this formula. Crucially, on the right hand side z_2 appears as an argument inside the expectation, and not as the inital condition. So we may replace the $Q^{n-\tau}$ in the expectation by $\mathcal{Q}^{(n-\tau)}$ to obtain a formula given as a polynomial in z_2. To finish the proof we need to check that the resulting (extended) formula coincides with the right hand side of (3.26) below the curve, which can be checked easily using the last equality in (3.26): all we need to check is that $\chi_{X_0(k+1)}\mathcal{Q}^{(n-k)}\bar{\chi}_{X_0(n)}=\chi_{X_0(k+1)}Q^{n-k}\bar{\chi}_{X_0(n)}$, which follows from $X_0(k+1)-X_0(n)>n-k-1$ and (3.25) (see [MQR17b] for the details). ☐

3.9. **Main formula for TASEP.** We need to introduce the discrete version of the operators in Definitions 1.10 and 1.13. In the first case we need two discrete versions:

Definition 3.19. We let

$$\mathtt{S}_{t,-n}(z_1,z_2)=(e^{t/2}R_tQ^{-n})^*(z_1,z_2)=\frac{1}{2\pi\mathrm{i}}\oint_{\Gamma_0}dw\,\frac{(1-w)^n}{2^{z_2-z_1}w^{n+1+z_2-z_1}}e^{t(w-1/2)},\tag{3.27}$$

$$\mathcal{S}_{t,n}(z_1,z_2)=e^{-t/2}\mathcal{Q}^{(n)}R_t^{-1}(z_1,z_2)=\frac{1}{2\pi\mathrm{i}}\oint_{\Gamma_0}dw\,\frac{(1-w)^{z_2-z_1+n-1}}{2^{z_1-z_2}w^n}e^{t(1/2-w)}\tag{3.28}$$

(where we used (3.9) and (3.18) for the first one, and (3.17) and (3.24) together with a residue computation for the second one).

Definition 3.20. (Discrete hit operator)

$$\bar{\mathcal{S}}^{\mathrm{\mathring{e}pi}(X_0)}_{t,n}(z_1,z_2)=\mathbb{E}_{B_0=z_1}[\mathcal{S}_{t,n-\tau}(B_\tau,z_2)\mathbf{1}_{\tau<n}].\tag{3.29}$$

Note that τ (defined in (3.21)) is the hitting time of the strict epigraph of the curve $\big(X_0(k+1)\big)_{k=0,\dots,n-1}$ by the random walk B_k; the *strict epigraph* of $\big(g(m)\big)_{m\geq 0}$ is the set

$$\mathring{\mathrm{epi}}(g) = \big\{(m,y)\colon m\geq 0,\ y>g(m)\big\}.$$

The following formula now follows directly from the above arguments and the last two definitions:

Theorem 3.21. (TASEP formula for right-finite initial data) *Assume that initially we have $X_0(j)=\infty$ for all $j\leq 0$, $X_0(1)<\infty$. Then for $1\leq n_1<n_2<\dots<n_M$ and $t\geq 0$,*

$$\mathbb{P}(X_t(n_j)>a_j,\ j=1,\dots,M) = \det(I-\bar{\chi}_a K_t \bar{\chi}_a)_{\ell^2(\{n_1,\dots,n_M\}\times\mathbb{Z})}, \tag{3.30}$$

where

$$K_t(n_i,\cdot;n_j,\cdot) = -Q^{n_j-n_i}\mathbf{1}_{n_i<n_j} + (\mathtt{S}_{t,-n_i})^*\bar{\mathcal{S}}^{\mathring{\mathrm{epi}}(X_0)}_{t,n_j}. \tag{3.31}$$

The path integral version (3.14) also holds.

Remark 3.22. Note that, by definition, $\bar{\mathcal{S}}^{\mathring{\mathrm{epi}}(X_0)}_{t,n_j} = \mathcal{S}_{t,n_j}(y,z)$ for $y>X_0(1)$, so (3.31) can also be written as

$$K_t(n_i,\cdot;n_j,\cdot) = -Q^{n_j-n_i}\mathbf{1}_{n_i<n_j} + (\mathtt{S}_{t,-n_i})^*\chi_{X_0(1)}\mathcal{S}_{t,n_j} + (\mathtt{S}_{t,-n_i})^*\bar{\chi}_{X_0(1)}\bar{\mathcal{S}}^{\mathring{\mathrm{epi}}(X_0)}_{t,n_j}. \tag{3.32}$$

Example 3.23. (Step initial data) Consider TASEP with step initial data, $X_0(i)=-i$ for $i\geq 1$. If we start the random walk in (3.29) from $B_0=z_1$ below the curve, i.e. $z_1\leq -1$, then the random walk clearly never hits the epigraph. Hence, $\bar{\chi}_{X_0(1)}\bar{\mathcal{S}}^{\mathrm{epi}(X_0)}_{t,n}\equiv 0$ and the last term in (3.32) vanishes. For the second term in (3.32) we have, from (3.27) and (3.28),

$$(\mathtt{S}_{t,-n_i})^*\chi_{X_0(1)}\mathcal{S}_{t,n_j}(z_1,z_2) = \frac{1}{(2\pi\mathrm{i})^2}\oint_{\Gamma_0}dw\oint_{\Gamma_0}dv\,\frac{(1-w)^{n_i}(1-v)^{n_j+z_2}}{2^{z_1-z_2}w^{n_i+z_1+1}v^{n_j}}\frac{e^{t(w+v-1)}}{1-v-w},$$

which is exactly the formula previously derived in the literature (see e.g. [Fer15, Eq. 82]).

Example 3.24. (Periodic initial data) Consider now TASEP with periodic initial data $X_0(i)=2i$, $i\in\mathbb{Z}$. One can obtain a formula for the kernel in this case by approximation, considering first the finite periodic initial data $X_0(i)=2(N-i)$ for $i=1,\dots,2N$. The computation is much more involved than in the previous example, but can be carried out explicitly (see [MQR17b]) and leads to

$$K_t^{(n)}(z_1,z_2) = -\frac{1}{2\pi\mathrm{i}}\oint_{1+\Gamma_0}dv\,\frac{v^{z_2+2m}}{2^{z_1-z_2}(1-v)^{z_1+2m+1}}\,e^{t(1-2v)},$$

which coincides with the kernel derived (modulo the conjugation $2^{z_2-z_1}$ and after a simple change of variables) in [BFPS07, Thm. 2.2].

Remark 3.25. It is possible to write a similar formula for TASEP with two-sided infinite initial data (see [MQR17b, Thm. 2.6]), but the result is much more complicated from the point of view of asymptotics than (3.30).

4. From TASEP to the KPZ fixed point

We will only sketch this part. In particular, for simplicity we will mostly only address the case of one-sided initial data for the fixed point, which means

$$\mathfrak{h}_0(\mathbf{x}) = -\infty \quad \forall\ \mathbf{x}>0.$$

This corresponds to right-finite TASEP initial data as in Theorem 3.21, and in fact any such $\mathfrak{h}_0$ can be approximated in UC by right-finite TASEP initial data X_0^ε.

4.1. **Scaling and asymptotics.** Our goal is to take such a sequence of initial data X_0^ε and compute

$$
\begin{aligned}
&\mathbb{P}_{\mathfrak{h}_0}(\mathfrak{h}(\mathbf{t},\mathbf{x}_i) \le \mathbf{a}_i,\ i=1,\dots,M) \\
&\qquad = \lim_{\varepsilon\to 0} \mathbb{P}_{X_0}\Big(X_{2\varepsilon^{-3/2}\mathbf{t}}(\tfrac{1}{2}\varepsilon^{-3/2}\mathbf{t} - \varepsilon^{-1}\mathbf{x}_i - \tfrac{1}{2}\varepsilon^{-1/2}\mathbf{a}_i + 1) > 2\varepsilon^{-1}\mathbf{x}_i - 2,\ i=1,\dots,M\Big)
\end{aligned}
$$

(the equality comes from (1.3), (3.1) and (1.5)). We therefore want to consider Theorem 3.21 with

$$t = 2\varepsilon^{-3/2}\mathbf{t}, \qquad n_i = \tfrac{1}{2}\varepsilon^{-3/2}\mathbf{t} - \varepsilon^{-1}\mathbf{x}_i - \tfrac{1}{2}\varepsilon^{-1/2}\mathbf{a}_i + 1 \tag{4.1}$$

(and with $a_i = 2\varepsilon^{-1}\mathbf{x}_i - 2$).

Lemma 4.1. *Under the scaling* (4.1) *and assuming that* $\varepsilon^{1/2}\left(X_0^\varepsilon(\varepsilon^{-1}\mathbf{x}) + 2\varepsilon^{-1}\mathbf{x} - 1\right) \longrightarrow -\mathfrak{h}_0(-\mathbf{x})$ *as* $\varepsilon \to 0$ *in* LC, *if we set* $z_i = 2\varepsilon^{-1}\mathbf{x}_i + \varepsilon^{-1/2}(u_i + \mathbf{a}_i) - 2$ *and* $y' = \varepsilon^{-1/2}v$, *then we have, as* $\varepsilon \searrow 0$,

$$\mathbf{S}^\varepsilon_{\mathbf{t},\mathbf{x}_i}(v,u_i) := \varepsilon^{-1/2}\mathsf{S}_{t,-n_i}(y',z_i) \longrightarrow \mathbf{S}_{\mathbf{t},\mathbf{x}_i}(v,u_i), \tag{4.2}$$

$$\widetilde{\mathbf{S}}^\varepsilon_{\mathbf{t},-\mathbf{x}_j}(v,u_j) := \varepsilon^{-1/2}\mathcal{S}_{t,n_j}(y',z_j) \longrightarrow \mathbf{S}_{\mathbf{t},-\mathbf{x}_j}(v,u_j), \tag{4.3}$$

$$\bar{\mathbf{S}}^{\varepsilon,\mathrm{epi}(-\mathfrak{h}_0^-)}_{\mathbf{t},-\mathbf{x}_j}(v,u_j) := \varepsilon^{-1/2}\bar{\mathcal{S}}^{\mathrm{epi}(X_0)}_{t,n_j}(y',z_j) \longrightarrow \bar{\mathbf{S}}^{\mathrm{epi}(-\mathfrak{h}_0^-)}_{\mathbf{t},-\mathbf{x}_j}(v,u_j) \tag{4.4}$$

pointwise, where $\mathfrak{h}_0^-(x) = \mathfrak{h}_0(-x)$ *for* $x \ge 0$.

The first two limits follow from standard arguments (although in [MQR17b] we need detailed estimates which are slightly involved). The main point is that the limit of $\bar{\mathbf{S}}^{\varepsilon,\mathrm{epi}(-\mathfrak{h}_0^-)}_{\mathbf{t},-\mathbf{x}_j}$ can be computed naturally: essentially one uses the asymptotics of $\mathbf{S}^\varepsilon_{\mathbf{t},-\mathbf{x}_j}$ together with the fact that, under our scaling, the random walk B goes to the Brownian motion $\mathbf{B}$.

Sketch of the proof of Lemma 4.1. From (3.27),

$$\mathsf{S}_{t,-n_i}(z_i,y) = \frac{1}{2\pi\mathrm{i}}\oint_{\Gamma_0} e^{\varepsilon^{-3/2}F^{(3)} + \varepsilon^{-1}F^{(2)} + \varepsilon^{-1/2}F^{(1)} + F^{(0)}}\,dw,$$

where

$$
\begin{aligned}
F^{(3)} &= \mathbf{t}\left[(2w-1) + \tfrac{1}{2}\log(\tfrac{1-w}{w})\right], \qquad F^{(2)} = -\mathbf{x}_i \log 4w(1-w), \\
F^{(1)} &= (u_i - v - \tfrac{1}{2}\mathbf{a}_i)\log 2w - \tfrac{1}{2}\mathbf{a}_i \log 2(1-w), \qquad F^{(0)} = \log 8w^2.
\end{aligned} \tag{4.5}
$$

The leading term has a double critical point at $w = 1/2$, so we introduce the change of variables $w \mapsto \frac{1}{2}(1 - \varepsilon^{1/2}\tilde{w})$, which leads to

$$\varepsilon^{-3/2}F^{(3)} \approx \tfrac{\mathbf{t}}{3}\tilde{w}^3, \quad \varepsilon^{-1}F^{(2)} \approx \mathbf{x}_i\tilde{w}^2, \quad \varepsilon^{-1/2}F^{(1)} \approx -(u_i - v)\tilde{w}.$$

We also have $F^{(0)} \approx \log(2)$, which cancels the prefactor $1/2$ coming from the change of variables. In view of (1.9), this gives (4.2). (4.3) follows in the same way, now using (3.28).

Now define the scaled walk $\mathbf{B}^\varepsilon(\mathbf{x}) = \varepsilon^{1/2}\left(B_{\varepsilon^{-1}\mathbf{x}} + 2\varepsilon^{-1}\mathbf{x} - 1\right)$ for $\mathbf{x} \in \varepsilon\mathbb{Z}_{\ge 0}$, interpolated linearly in between, and let $\boldsymbol{\tau}^\varepsilon$ be the hitting time by $\mathbf{B}^\varepsilon$ of $\mathrm{epi}(-\mathfrak{h}^\varepsilon(0,\cdot)^-)$. By Donsker's invariance principle [Bil99], $\mathbf{B}^\varepsilon$ converges locally uniformly in distribution to a Brownian motion $\mathbf{B}(\mathbf{x})$ with diffusion coefficient 2, and therefore the hitting time $\boldsymbol{\tau}^\varepsilon$ converges to $\boldsymbol{\tau}$ as well. (3.28) and (4.3) now show that, modulo explicit estimates, (4.4) should hold. □

Theorem 4.2. (One-sided fixed point formulas) *Let* $\mathfrak{h}_0 \in \mathrm{UC}$ *with* $\mathfrak{h}_0(\mathbf{x}) = -\infty$ *for* $\mathbf{x} > 0$. *Then given* $\mathbf{x}_1 < \mathbf{x}_2 < \cdots < \mathbf{x}_M$ *and* $\mathbf{a}_1,\dots,\mathbf{a}_M \in \mathbb{R}$, *we have, for* $\mathfrak{h}(\mathbf{t},\mathbf{x})$ *given as in*

(1.5),

$$\begin{aligned}
&\mathbb{P}_{\mathfrak{h}_0}(\mathfrak{h}(\mathbf{t},\mathbf{x}_1)\le \mathbf{a}_1,\dots,\mathfrak{h}(\mathbf{t},\mathbf{x}_M)\le \mathbf{a}_M)\\
&\qquad= \det\Big(\mathbf{I}-\chi_{\mathbf{a}}\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t},\mathrm{ext}}\chi_{\mathbf{a}}\Big)_{L^2(\{\mathbf{x}_1,\dots,\mathbf{x}_M\}\times\mathbb{R})} &(4.6)\\
&\qquad= \det\Big(\mathbf{I}-\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t},\mathbf{x}_M}+\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t},\mathbf{x}_M}e^{(\mathbf{x}_1-\mathbf{x}_M)\partial^2}\bar\chi_{\mathbf{a}_1}e^{(\mathbf{x}_2-\mathbf{x}_1)\partial^2}\bar\chi_{\mathbf{a}_2}\cdots e^{(\mathbf{x}_M-\mathbf{x}_{M-1})\partial^2}\bar\chi_{\mathbf{a}_M}\Big)_{L^2(\mathbb{R})}, &(4.7)
\end{aligned}$$

where $\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t},\mathbf{x}}$ *was defined in Definition 1.14.*

Proof. We have $n_i<n_j$ for small ε if and only if $\mathbf{x}_j<\mathbf{x}_i$ and in this case we have, under our scaling,

$$\varepsilon^{-1/2}Q^{n_j-n_i}(z_i,z_j)\longrightarrow e^{(\mathbf{x}_i-\mathbf{x}_j)\partial^2}(u_i,u_j),$$

as $\varepsilon\searrow 0$. From this and the above lemma we obtain the following limiting formula:

$$\mathbb{P}_{\mathfrak{h}_0}(\mathfrak{h}(\mathbf{t},\mathbf{x}_1)\le \mathbf{a}_1,\dots,\mathfrak{h}(\mathbf{t},\mathbf{x}_M)\le \mathbf{a}_M)=\det(\mathbf{I}-\bar\chi_{\mathbf{a}}\mathbf{K}_{\lim}\bar\chi_{\mathbf{a}})_{L^2(\{\mathbf{x}_1,\dots,\mathbf{x}_M\}\times\mathbb{R})}$$

with

$$\mathbf{K}_{\lim}(\mathbf{x}_i,u_i;\mathbf{x}_j,u_j)=-e^{(\mathbf{x}_i-\mathbf{x}_j)\partial^2}(u_i,u_j)\mathbf{1}_{\mathbf{x}_i>\mathbf{x}_j}+(\mathbf{S}_{\mathbf{t},\mathbf{x}_i})^*\bar{\mathbf{S}}^{\mathrm{epi}(-\mathfrak{h}_0^-)}_{\mathbf{t},-\mathbf{x}_j}(u_i,u_j).$$

We may turn the above projections $\bar\chi_{-\mathbf{a}}$ into $\chi_{\mathbf{a}}$ by changing variables $u_i\longmapsto -u_i$ in the kernel. If we additionally replace the Fredholm determinant of the kernel by that of its adjoint to get

$$\det\Big(\mathbf{I}-\chi_a\widetilde{\mathbf{K}}_{\lim}\chi_a\Big)\qquad\text{with}\quad \widetilde{\mathbf{K}}_{\lim}(u_i,u_j)=\mathbf{K}_{\lim}(\mathbf{x}_j,-u_j;\mathbf{x}_i,-u_i).$$

Now

$$\mathbf{S}_{\mathbf{t},\mathbf{x}}(-u,v)=(\mathbf{S}_{\mathbf{t},\mathbf{x}})^*(-v,u)\qquad\text{and}\qquad \bar{\mathbf{S}}^{\mathrm{epi}(-\mathfrak{h}_0^-)}_{\mathbf{t},\mathbf{x}}(v,-u)=(\bar{\mathbf{S}}^{\mathrm{hypo}(\mathfrak{h}_0^-)}_{\mathbf{t},\mathbf{x}})^*(u,-v),$$

so

$$\widetilde{\mathbf{K}}_{\lim}=-e^{(\mathbf{x}_j-\mathbf{x}_i)\partial^2}\mathbf{1}_{\mathbf{x}_i<\mathbf{x}_j}+(\bar{\mathbf{S}}^{\mathrm{hypo}(\mathfrak{h}_0^-)}_{\mathbf{t},-\mathbf{x}_i})^*\mathbf{S}_{\mathbf{t},\mathbf{x}_j}=\mathbf{K}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t},\mathrm{ext}}.$$

This gives the extended kernel formula 4.6. The path integral version 4.7 follows from the framework of [BCR15]. □

4.2. **From one-sided to two-sided formulas.** In the last result we assumed $\mathfrak{h}_0(\mathbf{x})=-\infty$ for $\mathbf{x}>0$. Obtaining from this a formula for general $\mathfrak{h}_0\in\mathrm{UC}$ involves two separate arguments:

1. For $\mathfrak{h}_0\in\mathrm{UC}$ let $\mathfrak{h}_0^L(\mathbf{x})=\mathfrak{h}_0(\mathbf{x})\mathbf{1}_{\mathbf{x}\le L}-\infty\cdot\mathbf{1}_{\mathbf{x}>L}$. Then we need to compute the limit of (4.6)/(4.7) as $L\to\infty$ with initial data $\mathfrak{h}_0^L$. To this end we use shift invariance (which follows directly from that of TASEP) to translate this into a problem involving a (shifted) initial condition which is $-\infty$ on the positive axis. The kernels appearing in the Fredholm determinants now involve some shifts by L, but in view of Remark 1.15 it is possible to rewrite them in such a way that they look exactly like those in (4.6)/(4.7) but with $\mathfrak{h}_0$ replaced by $\mathfrak{h}_0^L$, and so essentially all one needs is to show that $\bar{\mathbf{S}}^{\mathrm{hypo}(\mathfrak{h}_0^L)}_{\mathbf{t},\mathbf{x}}\longrightarrow\bar{\mathbf{S}}^{\mathrm{hypo}(\mathfrak{h}_0)}_{\mathbf{t},\mathbf{x}}$, which was shown essentially in [QR16]. See [MQR17b, Sec. 3.4] for the details.
2. We need to show that the limit we get in the previous point is in fact the same as the limit we would get for TASEP with an initial height profile h_0^ε going to $\mathfrak{h}_0$ in UC as $\varepsilon\to 0$. This amounts to considering a truncated version $h_0^{\varepsilon,L}$ of h_0^ε which goes to $\mathfrak{h}_0^L$, taking $\varepsilon\to0$ and $L\to\infty$, and then showing that the limits can be interchanged. This can be justified by showing that the error we incur at the level of TASEP by replacing h_0^ε by $h_0^{\varepsilon,L}$ can be bounded by something that goes to 0 with $L\to\infty$, uniformly in ε. This is the content of the finite speed of propagation result in [MQR17b, Lem. 3.2].

These arguments lead in [MQR17b, Sec. 3.4] to Theorem 1.19 above.

4.3. **Continuum limit.** The last step in order to get to (1.15) is to compute a continuum limit in the $\mathbf{a}_i$'s of the path integral formula (1.17) on the full line. By this we mean that we take $\mathfrak{g} \in \mathrm{UC}$, let $\mathbf{x}_1, \dots, \mathbf{x}_M$ be a fine mesh of $[-R, R]$, take $M \to \infty$ with $\mathbf{a}_i = \mathfrak{g}(\mathbf{x}_i)$ to obtain a continuum statistics formula, and finally take $R \to \infty$. To this end one notes that, with these choices, $e^{(\mathbf{x}_1 - \mathbf{x}_M)\partial^2}$ becomes $e^{-2R\partial^2}$ (which makes sense when applied after $\mathbf{K}_{\mathbf{t}}^{\mathrm{hypo}(\mathfrak{h}_0)}$) while

$$\bar{\chi}_{\mathbf{a}_1} e^{(\mathbf{x}_2 - \mathbf{x}_1)\partial^2} \bar{\chi}_{\mathbf{a}_2} \cdots e^{(\mathbf{x}_M - \mathbf{x}_{M-1})\partial^2} \bar{\chi}_{\mathbf{a}_M}(u_1, u_2) \xrightarrow[\varepsilon \to 0]{} \mathbb{P}_{\mathbf{B}(\ell_1) = u_1}\big(\mathbf{B}(s) \leq \mathfrak{g}(s)\ \forall\, s \in [\ell_1, \ell_2],\, \mathbf{B}(\ell_2) \in du_2\big)/du_2,$$

that is, the transition probability for $\mathbf{B}$ not to hit $\mathrm{epi}(\mathfrak{g})$. This establishes the connection with the hit operator $\mathbf{K}_{\mathbf{t}}^{\mathrm{epi}(\mathfrak{g})}$. For the details, and in particular the computation of the $R \to \infty$ limit (which is essentially contained already in [QR16]) see [MQR17b, Sec. 3.5].

References

[BFP10] J. Baik, P. L. Ferrari, and S. Péché. Limit process of stationary TASEP near the characteristic line. *Comm. Pure Appl. Math.* 63.8 (2010), pp. 1017–1070.

[Bet31] H. A. Bethe. On the theory of metals, i. eigenvalues and eigenfunctions of a linear chain of atoms. *Zeits. Phys.* 74 (1931), pp. 205–226.

[Bil99] P. Billingsley. *Convergence of probability measures.* Second. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999, pp. x+277.

[BCR15] A. Borodin, I. Corwin, and D. Remenik. Multiplicative functionals on ensembles of non-intersecting paths. *Ann. Inst. H. Poincaré Probab. Statist.* 51.1 (2015), pp. 28–58.

[BF08] A. Borodin and P. L. Ferrari. Large time asymptotics of growth models on space-like paths. I. PushASEP. *Electron. J. Probab.* 13 (2008), no. 50, 1380–1418.

[BFP07] A. Borodin, P. L. Ferrari, and M. Prähofer. Fluctuations in the discrete TASEP with periodic initial configurations and the Airy_1 process. *Int. Math. Res. Pap. IMRP* (2007), Art. ID rpm002, 47.

[BFPS07] A. Borodin, P. L. Ferrari, M. Prähofer, and T. Sasamoto. Fluctuation properties of the TASEP with periodic initial configuration. *J. Stat. Phys.* 129.5-6 (2007), pp. 1055–1080.

[BFS08a] A. Borodin, P. L. Ferrari, and T. Sasamoto. Large time asymptotics of growth models on space-like paths. II. PNG and parallel TASEP. *Comm. Math. Phys.* 283.2 (2008), pp. 417–449.

[BFS08b] A. Borodin, P. L. Ferrari, and T. Sasamoto. Transition between Airy_1 and Airy_2 processes and TASEP fluctuations. *Comm. Pure Appl. Math.* 61.11 (2008), pp. 1603–1629.

[Cor12] I. Corwin. The Kardar-Parisi-Zhang equation and universality class. *Random Matrices Theory Appl.* 1 (2012).

[CQR13] I. Corwin, J. Quastel, and D. Remenik. Continuum statistics of the Airy_2 process. *Comm. Math. Phys.* 317.2 (2013), pp. 347–362.

[CQR15] I. Corwin, J. Quastel, and D. Remenik. Renormalization fixed point of the KPZ universality class. *J. Stat. Phys.* 160.4 (2015), pp. 815–834.

[EM98] B. Eynard and M. L. Mehta. Matrices coupled in a chain. I. Eigenvalue correlations. *J. Phys. A* 31.19 (1998), pp. 4449–4456.

[Fer15] P. L. Ferrari. Dimers and orthogonal polynomials: connections with random matrices. In: *Dimer models and random tilings.* Vol. 45. Panor. Synthèses. Soc. Math. France, Paris, 2015, pp. 47–79.

[HH85] D. A. Huse and C. L. Henley. Pinning and roughening of domain walls in Ising systems due to random impurities. *Phys. Rev. Lett.* 54.25 (1985), pp. 2708–2711.

[Joh03] K. Johansson. Discrete polynuclear growth and determinantal processes. *Comm. Math. Phys.* 242.1-2 (2003), pp. 277–329.

[KPZ86] M. Kardar, G. Parisi, and Y.-C. Zhang. Dynamical scaling of growing interfaces. *Phys. Rev. Lett.* 56.9 (1986), pp. 889–892.

[Lig85] T. M. Liggett. *Interacting particle systems*. Vol. 276. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. New York: Springer-Verlag, 1985, pp. xv+488.

[MQR17a] K. Matetski, J. Quastel, and D. Remenik. In preparation. 2017.

[MQR17b] K. Matetski, J. Quastel, and D. Remenik. *The KPZ fixed point*. 2017. arXiv: 1701.00018v1.

[NR16] G. B. Nguyen and D. Remenik. *Extreme statistics of non-intersecting Brownian paths*. 2016. arXiv: 1611.10294.

[NR15] G. B. Nguyen and D. Remenik. *Non-intersecting Brownian bridges and the Laguerre Orthogonal Ensemble*. 2015. arXiv: 1505.01708.

[PS02] M. Prähofer and H. Spohn. Scale invariance of the PNG droplet and the Airy process. *J. Stat. Phys.* 108.5-6 (2002), pp. 1071–1106.

[Qua11] J. Quastel. The Kardar-Parisi-Zhang equation. In: *Current developments in mathematics, 2011*. Int. Press, Somerville, MA, 2011.

[QR14] J. Quastel and D. Remenik. Airy processes and variational problems. In: *Topics in Percolative and Disordered Systems*. Ed. by A. Ramírez, G. Ben Arous, P. A. Ferrari, C. Newman, V. Sidoravicius, and M. E. Vares. Vol. 69. Springer Proceedings in Mathematics & Statistics. 2014, pp. 121–171.

[QR16] J. Quastel and D. Remenik. *How flat is flat in random interface growth?* 2016. arXiv: 1606.09228.

[QR13] J. Quastel and D. Remenik. Local behavior and hitting probabilities of the Airy_1 process. *Probability Theory and Related Fields* 157.3-4 (2013), pp. 605–634.

[QS15] J. Quastel and H. Spohn. *The one-dimensional KPZ equation and its universality class*. 2015. arXiv: 1503.06185.

[Sas05] T. Sasamoto. Spatial correlations of the 1D KPZ surface on a flat substrate. *Journal of Physics A: Mathematical and General* 38.33 (2005), p. L549.

[Sch97] G. M. Schütz. Exact solution of the master equation for the asymmetric exclusion process. *J. Statist. Phys.* 88.1-2 (1997), pp. 427–445.

[Sim05] B. Simon. *Trace ideals and their applications*. Second. Vol. 120. Mathematical Surveys and Monographs. American Mathematical Society, 2005, pp. viii+150.

[TW94] C. A. Tracy and H. Widom. Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.* 159.1 (1994), pp. 151–174.

[TW96] C. A. Tracy and H. Widom. On orthogonal and symplectic matrix ensembles. *Comm. Math. Phys.* 177.3 (1996), pp. 727–754.

Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático, Universidad de Chile, Av. Beauchef 851, Torre Norte, Piso 5, Santiago, Chile

E-mail address: dremenik@dim.uchile.cl