HAL
archives-ouvertes.fr

DLR equations and rigidity for the Sine-beta process

David Dereudre, Adrien Hardy, Thomas Leblé, Mylène Maïda

DLR equations and rigidity for the Sine-beta process

David Dereudre [*],   Adrien Hardy [†],   Thomas Leblé [‡],   Mylène Maïda [§]

September 18, 2018

This paper is dedicated to the memory of our colleague Hans-Otto Georgii (1944-2017). He played a crucial role in the rigorous development of the theory of Gibbs measures.

Abstract

We investigate Sine_β, the universal point process arising as the thermodynamic limit of the microscopic scale behavior in the bulk of one-dimensional log-gases, or β-ensembles, at inverse temperature $\beta > 0$. We adopt a statistical physics perspective, and give a description of Sine_β using the Dobrushin-Landford-Ruelle (DLR) formalism by proving that it satisfies the DLR equations: the restriction of Sine_β to a compact set, conditionally to the exterior configuration, reads as a Gibbs measure given by a finite log-gas in a potential generated by the exterior configuration. In short, Sine_β is a natural infinite Gibbs measure at inverse temperature $\beta > 0$ associated with the logarithmic pair potential interaction. Moreover, we show that Sine_β is number-rigid and tolerant in the sense of Ghosh-Peres, i.e. the number, but not the position, of particles lying inside a compact set is a deterministic function of the exterior configuration. Our proof of the rigidity differs from the usual strategy and is robust enough to include more general long range interactions in arbitrary dimension.

1   Introduction and main results

1.1   The log-gas and the Sine-beta process

The (finite, one-dimensional) log-gas is a random system of n particles confined on the real line $\mathbb{R}$, or the unit circle $\mathbb{T}$, interacting via a repulsive pair potential given by the logarithm of the inverse distance between particles. Physically, it represents a statistical gas of identically charged particles living in a one-dimensional environment and interacting according to the laws of two-dimensional electrostatics. For a fixed value of the *inverse temperature* parameter $\beta > 0$, the distribution of the n-tuple of particles is given by the canonical Gibbs measure for this interaction, whose density reads as

$$\frac{d\,\text{Log-gas}}{d\,\text{Lebesgue}^{\otimes n}} \propto \exp\left(-\beta \times \text{Logarithmic interaction energy of the particles}\right), \qquad (1.1)$$

---

[*]Laboratoire Paul Painlevé, Université de Lille, Cité Scientifique, 59655 Villeneuve d'Ascq Cedex, France. Email: david.dereudre@univ-lille.fr

[†]Laboratoire Paul Painlevé, Université de Lille, Cité Scientifique, 59655 Villeneuve d'Ascq Cedex, France. Email: adrien.hardy@univ-lille.fr

[‡]Courant Institute of Mathematical Sciences, 251 Mercer Street, New York University, New York, NY 10012-1110, USA. Email:thomasl@math.nyu.edu

[§]Laboratoire Paul Painlevé, Université de Lille, Cité Scientifique, 59655 Villeneuve d'Ascq Cedex, France. Email: mylene.maida@univ-lille.fr

where "$\propto$" means "up to a multiplicative normalizing constant".

Random matrices and β-ensembles. From the statistical physics point of view, the log-gas is interesting because of the singular and long-range nature of its interaction potential. Another important motivation comes from the link between log-gases and random matrix theory; we refer to [Forrester, 2010] for an extensive treatment of this connection. There are several models of random $n \times n$ matrices whose random eigenvalues exhibit a joint density of the form,

$$\frac{\mathrm{d}\,\mathrm{Eigenvalues}}{\mathrm{d}\,\mathrm{Lebesgue}^{\otimes n}} \propto \exp\left(-\beta\left(\sum_{j<l} -\log|x_j - x_l|\right)\right) \prod_{j=1}^n \omega(x_j) = \prod_{j<\ell}^n |x_j - x_\ell|^\beta \prod_{j=1}^n \omega(x_j)\,,$$
(1.2)

where ω is an appropriate weight function supported on (a subset of) $\mathbb{R}$ or $\mathbb{T}$.

⋄ The Gaussian ensembles correspond to $\omega(x) = \mathrm{e}^{-x^2}$. For the specific values $\beta = 1$ or 2 or 4, one recovers the celebrated orthogonal/unitary/symplectic invariant random matrix ensembles.

⋄ The Wishart, or Laguerre ensembles correspond to $\omega(x) = x^\alpha \mathrm{e}^{-x} \mathbf{1}_{\mathbb{R}_+}(x)$.

⋄ The MANOVA, or Jacobi ensembles correspond to $\omega(x) = x^{\alpha_1}(1 - x)^{\alpha_2}\mathbf{1}_{[0,1]}(x)$.

These three models are random Hermitian matrices with eigenvalues on $\mathbb{R}$. Another example is given by the Circular β-ensemble, hereafter denoted CβE, which is a random unitary matrix with eigenvalues on the unit circle $\mathbb{T}$ and joint law (1.2) with $\omega(x) = 1$.

When $\beta = 2$, an integrable structure comes into play: the eigenvalues process is a determinantal point process, allowing exact computations for most quantities of interest, such as the correlation functions that one can express in terms of determinants. Similarly, when $\beta = 1$ or 4, some computations are still tractable due to a Pfaffian point process structure, although at the price of more involved formulas. In contrast, the present paper deals with "β arbitrary".

The Sine$_\beta$ **process.** Under an appropriate scaling within the bulk (i.e. the interior) of the spectrum, chosen so that the typical distance between consecutive points is of order 1, the large n limit of the random point process of the eigenvalues exists and is called the Sine$_\beta$ process. Stated otherwise, Sine$_\beta$ is the *microscopic thermodynamic limit in the bulk* of one-dimensional log-gases, and it is the probability law of a certain random infinite point configuration on $\mathbb{R}$.

The Sine$_\beta$ process is *universal* in the sense it only depends on the inverse temperature $\beta > 0$ and not on the initial weight ω, for a large family of weights ω, see Bourgade et al. [2014, 2012]. It is also "universal in the interaction", in the sense that it appears when considering pair potential interactions that only have a logarithmic singularity at zero, see [Venker, 2013].

In the $\beta = 1, 2, 4$ cases, this limiting process is rather well understood due to the determinantal/Pfaffian structure, see e.g. [Deift and Gioev, 2009] and references therein. However, in the general $\beta > 0$ setting, the mere existence of this limit is a difficult result, which was obtained, together with a rather involved description of the limiting object, by Valkó and Virág [2009] for the Gaussian β-ensemble and by Killip and Stoiciu [2009] for the Circular β-ensemble. The fact that these two descriptions coincide is checked e.g. in [Nakano, 2014].

These descriptions of Sine$_\beta$ use systems of coupled stochastic differential equations to derive the number of eigenvalues/particles falling in a given interval. This turns out to be tractable enough to study fine properties of the point process, such as to obtain large gap probability estimates [Valkó and Virág, 2010], a Central Limit Theorem [Kritchevski et al., 2012] and large deviation and maximum deviation estimates [Holcomb and Valkó, 2015, 2017, Holcomb and Paquette, 2018] for the number of points in an interval, as well as the Poissonian behavior of Sine$_\beta$ as $\beta \to 0$ [Allez and Dumaz, 2014].

More recently, the process has been characterized by Valkó and Virág [2017] as the spectrum of a random infinite-dimensional operator. In particular this allows a better understanding on the β-dependency of the process [Valkó and Virág, 2018].

However, among particularly relevant features of Sine$_\beta$, its correlation functions, and even the asymptotic behavior of its two-point correlation function, remain unknown for generic $\beta > 0$; see however [Forrester, 1993, Chapter 13] for special cases.

The goal of this work is to study Sine$_\beta$ from a statistical physics perspective, so as to obtain an alternative description as an infinite Gibbs measure, characterized by canonical Dobrushin-Landford-Ruelle equations. In short, Sine$_\beta$ is the natural infinite Gibbs measure at inverse temperature $\beta > 0$ associated with the logarithmic pair potential interaction.

1.2 Sine-beta as an infinite Gibbs measure

1.2.1 Context for the DLR formalism

For any fixed number n of particles, the canonical Gibbs measure of the Log-gas mentioned in (1.1) minimizes the quantity

$$\beta \times \text{Expected logarithmic energy} + \text{Entropy with respect to Lebesgue}^{\otimes n}$$

among all probability laws of random n-point configuration. This is a famous variational principle for Gibbs measures, see e.g. [Friedli and Velenik, 2017, Section 6.9] for a discussion. On the other hand, in the infinite-volume setting, it is shown in [Leblé and Serfaty, 2017, Corollary 1.2] that Sine$_\beta$ minimizes a free energy functional of the type,

$$\beta \times \text{Expected renormalized energy} + \text{Entropy with respect to Poisson}, \qquad (1.3)$$

among laws of stationary point processes, where Poisson is the Poisson point process with intensity 1 on $\mathbb{R}$. Here "renormalized energy" is a way to define the logarithmic energy of infinite configurations at microscopic scale, see Section 2.4. It is thus natural to ask whether one can obtain a description of Sine$_\beta$ as an infinite Gibbs measure. In view of (1.1), the naive guess would be that

$$\frac{d\,\text{Sine}_\beta}{d\,\text{Poisson}} \propto \exp\left(-\beta \times \text{Renormalized energy}\right), \qquad (1.4)$$

which is well-known to be illusory because any stationary process absolutely continuous with respect to a Poisson process is the said Poisson process itself. The Dobrushin-Landford-Ruelle (DLR) formalism provides the correct setting to possibly recast (1.4) in a *local* way; we refer e.g. to the book [Georgii, 2011] for a general presentation in the lattice case, see also [Dereudre, 2017] for a pedagogical introduction in the setting of point processes. Informally, we will show that, given any bounded Borel set $\Lambda \subset \mathbb{R}$ and any

configuration γ_{Λ^c} outside of Λ, the law of the configuration γ_Λ in Λ drawn from Sine_β knowing the exterior configuration γ_{Λ^c} can be written as

$$\frac{\mathrm{d}\left(\mathrm{Sine}_\beta \text{ in } \Lambda \,\Big|\, \gamma_{\Lambda^c}\right)}{\mathrm{d}\,\mathrm{Lebesgue}^{\otimes N}}(\gamma_\Lambda)$$
$$\propto \exp\left(-\beta \times \text{Logarithmic energy of } \gamma_\Lambda \text{ with itself and with } \gamma_{\Lambda^c}\right),$$

where the number of points N of the configuration γ_Λ is almost surely determined by the exterior configuration γ_{Λ^c}.

1.2.2 Terminology and notation

A point *configuration* on $\mathbb{R}$ is a locally finite subset of $\mathbb{R}$ allowing multiples points. Formally, we identify a configuration γ with an integer valued Radon measure on $\mathbb{R}$ that we still denote γ. That is we write $\gamma = \sum_{x \in \gamma} \delta_x$, and if f is a test function, we let $\int f \mathrm{d}\gamma := \sum_{x \in \gamma} f(x)$. When $\gamma \in \mathsf{Conf}(\mathbb{R})$ is such that $\gamma(\{x\}) \in \{0,1\}$ for every $x \in \mathbb{R}$, we say that γ is a *simple* configuration. The space of point configurations $\mathsf{Conf}(\mathbb{R})$, seen as a subspace of the Radon measures, is equipped with the topology coming by duality with the space of continuous functions $\mathbb{R} \to \mathbb{R}$ with compact support, making $\mathsf{Conf}(\mathbb{R})$ a Polish space, see e.g. [Kallenberg, 1983, Section 15.7]. It is also the smallest topology that makes the mapping $\gamma \mapsto \gamma(B)$ continuous for every bounded Borel set B. We then endow $\mathsf{Conf}(\mathbb{R})$ with its Borel σ-algebra. In the following, by a *point process* we mean a probability measure on simple configurations on $\mathbb{R}$, namely a probability measure on $\mathsf{Conf}(\mathbb{R})$ such that $\mathbb{P}(\gamma \text{ is simple}) = 1$. Let us introduce some additional notation: for any Borel set $\Lambda \subset \mathbb{R}$,

$\diamond$ the set of all configurations in Λ is denoted by $\mathsf{Conf}(\Lambda)$

$\diamond$ $\gamma \in \mathsf{Conf}(\mathbb{R}) \mapsto \gamma_\Lambda \in \mathsf{Conf}(\Lambda)$ stands for the restriction of a configuration to Λ

$\diamond$ $|\gamma| = \gamma(\mathbb{R}) \in \mathbb{N} \cup \{\infty\}$ stands for the number of points of a configuration γ.

1.2.3 Statement of the results

The central result of this paper is the next theorem.

Theorem 1.1. For any $\beta > 0$ and any bounded Borel set $\Lambda \subset \mathbb{R}$ the following holds.
(A) Rigidity. There exists a measurable function $\mathsf{Number}_\Lambda : \mathsf{Conf}(\Lambda^c) \to \mathbb{N}$ such that

$$N := |\gamma_\Lambda| = \mathsf{Number}_\Lambda(\gamma_{\Lambda^c}) \quad \text{for } \mathrm{Sine}_\beta\text{-almost every (a.e.) } \gamma.$$

(B) Definiteness of the exterior potential. For any $x \in \Lambda$ and Sine_β-a.e. configuration γ, the following limit exists and is positive,

$$\omega(x|\gamma_{\Lambda^c}) := \lim_{p \to \infty} \prod_{\substack{u \in \gamma_{\Lambda^c} \\ |u| \leq p}} \left|1 - \frac{x}{u}\right|^\beta. \tag{1.5}$$

Moreover, the partition function

$$Z(\gamma_{\Lambda^c}) := \int_{\Lambda^N} \prod_{j < \ell}^N |x_j - x_\ell|^\beta \prod_{j=1}^N \omega(x_j|\gamma_{\Lambda^c}) \prod_{j=1}^N \mathrm{d}x_j \tag{1.6}$$

is finite and positive.

4

(C) DLR equations. For any bounded measurable function $f : \mathsf{Conf}(\mathbb{R}) \to \mathbb{R}$, we have

$$\mathbb{E}_{\text{Sine}_\beta}[f] = \int \left[\int_{\Lambda^N} f(\{x_1, \ldots, x_N\} \cup \gamma_{\Lambda^c}) \, \rho_{\Lambda^c}(x_1, \ldots, x_N) \prod_{j=1}^N \mathrm{d}x_j \right] \text{Sine}_\beta(\mathrm{d}\gamma)$$

where, with the notation of (1.5), (1.6), we let ρ_{Λ^c} be defined by

$$\rho_{\Lambda^c}(x_1, \ldots, x_N) := \frac{1}{Z(\gamma_{\Lambda^c})} \prod_{j<k}^N |x_j - x_k|^\beta \prod_{j=1}^N \omega(x_j | \gamma_{\Lambda^c}). \tag{1.7}$$

Number-rigidity: Part (A) of Theorem 1.1 states that, conditionally on the configuration outside a given bounded Borel set Λ, the number of points drawn by Sine_β inside Λ is deterministic. This property has been recently put forward under the notion of *number-rigidity* of point processes, starting from the pioneering work of Ghosh and Peres [2017]. Thus, Sine_β is number-rigid for any $\beta > 0$. The notion of rigid processes can be compared to the older notion of (non) *hereditary* processes, see e.g. [Dereudre, 2017, Definition 1] for a presentation. Roughly speaking, if P is not hereditary, it means that *certain* points cannot be deleted with positive probability, whereas number-rigidity implies that *no* point can be deleted. In particular, although Sine_β has finite specific relative entropy with respect to the Poisson point process, they are very different on this aspect - a Poisson process being, of course, far from rigid.

Shortly before the present work was completed, the rigidity property for Sine_β has been proven independently by Chhaibi and Najnudel [2018]. Their proof follows the strategy introduced by Ghosh and Peres [2017], namely to show that the variance of linear statistics for a smooth approximation of the characteristic function of a bounded interval can be made arbitrary small. To do so, they use variance estimates for polynomial test functions that were proven in [Jiang and Matsumoto, 2015] for the CβE, and proceed by approximation. The latter work relies on exact computations involving Jack's special functions which are tied to the specific structure of the Circular ensemble.

In contrast, our proof for the number-rigidity only involves material from classical statistical physics and seems more flexible since it only relies on a weak form of DLR equations, the so-called *canonical DLR equations*, and Campbell measures arguments. This may be of independent interest to prove number-rigidity for a larger class of point processes, in particular when the two-point correlation functions are not explicit, or simply not asymptotically tractable. To the best of our knowledge, this is the first alternative strategy with respect to the Ghosh-Peres method to prove number-rigidity. In particular, we prove the more general result that for large a class of long range interactions on $\mathbb{R}^d$, any solution of the canonical DLR equations is indeed number-rigid, see Theorem 3.18.

The exterior potential: The existence of the limit (1.5) is non-trivial and follows from quite subtle cancellations. In fact, it was expected in [Ghosh and Lebowitz, 2017, Section 12] that it cannot be defined properly leading to the belief that the DLR equations were not reachable in this setting.

Tolerance: There are other notions of rigidity than number-rigidity, such as *barycenter-rigidity* or *super-rigidity*. Barycenter-rigidity states that the barycenter of the configuration inside a domain is a deterministic function of the exterior configuration, and

super-rigidity expresses the fact that the interior configuration is completely prescribed by the exterior. On the other hand, the notion of *tolerance*, introduced in [Ghosh and Peres, 2017], states that, roughly speaking, the number of points is the only rigid quantity prescribed by the exterior. It follows from part (C) of Theorem 1.1 that Sine_β is tolerant, which is a new result for general $\beta > 0$. More precisely, we have:

Corollary 1.2 (Sine_β is tolerant)**.** For any bounded Borel set Λ and Sine_β-a.e. point configuration γ, the law of the particles drawn from Sine_β inside Λ given the exterior configuration γ_{Λ^c} is mutually absolutely continuous with respect to the N-fold Lebesgue measure, where N is the number of points in Λ (which is prescribed, by rigidity).

We also obtain the following result (which will follow from Corollary 3.3).

Corollary 1.3. For any disjoint bounded Borel sets $B_1, \ldots, B_k \subset \mathbb{R}$ with positive Lebesgue measure and any integers $n_1, \ldots, n_k$, we have

$$\mathrm{Sine}_\beta \left(|\gamma_{B_1}| = n_1, \ldots, |\gamma_{B_k}| = n_k \right) > 0.$$

Relation to previous results: As mentioned above, when $\beta = 2$ the process benefits from an integrable structure with explicit correlations functions: it is the determinantal point process associated with the sine kernel. Using this structure, previous to the work of Chhaibi and Najnudel [2018], Ghosh [2015] obtained the rigidity for $\beta = 2$. Moreover, parts (B) and (C) of the theorem, and thus the tolerance, have been obtained by Bufetov [2016] for a class of number-rigid determinantal point processes, including the Sine_2 process.

1.3   Related questions and perspectives

Fluctuations of smooth linear statistics. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a smooth and compactly supported test function. One may define the fluctuation of φ as the random variable

$$\mathrm{Fluct}[\varphi](\gamma) := \int \varphi(x) \left(\gamma(\mathrm{d}x) - \mathrm{d}x \right),$$

and ask for the behavior, as $\ell \to \infty$, of

$$\mathrm{Fluct}[\varphi_\ell](\gamma) := \int \varphi \left(\frac{x}{\ell} \right) \left(\gamma(\mathrm{d}x) - \mathrm{d}x \right). \tag{1.8}$$

when the random configuration γ has law Sine_β. Having in mind similar results for β-ensembles, see e.g. [Shcherbina, 2014], [Bekerman and Lodhia, 2018], [Bekerman et al., 2018], and since Theorem 1.1(C) shows that Sine_β is conditionally a β-ensemble, we could expect the fluctuation in (1.8) to converge in law *without normalization* to a centered Gaussian random variable with standard deviation proportional to the fractional Sobolev $H^{1/2}$ norm of the test function φ. This Central Limit Theorem is proven in [Leblé, 2018] for φ smooth enough, using Theorem 1.1 as a key input.

Uniqueness. It is natural to ask the following:
  (a) Is Sine_β the only stationary process satisfying the DLR equations?

The answer to this question is positive when $\beta = 2$, as a consequence of the work [Kuijlaars and Miña-Díaz, 2017] where, for Sine$_2$ a.e $\gamma \in \mathsf{Conf}(\mathbb{R})$, the asymptotic of the conditional measure $\rho_{\Lambda^c}(x_1, \ldots, x_N)$ has been shown to be Sine$_2$ in the limit where $\Lambda := [-R, R]$ and $R \to \infty$. We expect the answer to be positive for any value of β.

(b) Is Sine$_\beta$ the only minimizer of the free energy functional (1.3)? For a rigorous definition of this functional, see [Leblé and Serfaty, 2017].

(c) Are minimizers of the free energy the same as the solutions to the DLR equations? We expect all two answers to be positive for any $\beta > 0$. Indeed, uniqueness of infinite-volume Gibbs measures is usually expected in dimension one, see e.g. [Friedli and Velenik, 2017, Section 6.5.5] for such a result (that is not applicable here because our interaction is not short-range). For log-gases in dimension 2, namely Coulomb gases, or in higher dimension, it might happen however that the uniqueness of minimizers/solutions to DLR equations, even up to symmetries, fails to hold for certain values of β.

1.4 Strategy for the proof and plan of the paper

In order to prove Theorem 1.1, we first prove part (B), together with a weaker version of part (C), which form the *canonical DLR equations*, where we further condition on the number $|\gamma_\Lambda|$ of particles lying in Λ. These two results form Theorem 2.1, which is proven in Section 2. We start from the Circular ensemble, for which a, finite n, periodic version of the canonical DLR equations holds, and we perform several approximations. As a technical ingredient, we use *discrepancy estimates*, i.e. controls on the difference $|\gamma_\Lambda| - |\Lambda|$ between the number of points of a typical configuration γ in a bounded set Λ, and the size of Λ (in the sense of its Lebesgue measure).

Next, in Section 3, we leverage on Theorem 2.1 in order to obtain the rigidity result of Theorem 1.1 part (A). In fact, we prove a possibly more general result: Any stationary point process P satisfying the canonical DLR equations is number-rigid, see Theorem 3.2. The proof of rigidity goes by contradiction: if P were not rigid, then we could deduce from the canonical DLR equations a particular structure for its Campbell measures that turns out to be absurd due to the long range nature of the logarithmic interaction. This result is stated and proven for the logarithmic interaction and in dimension one, but our proof is robust enough to yield that the same result, that we state in Theorem 3.18, holds for more general long range interactions in dimension $d \geq 1$.

2 Canonical DLR equations

We consider here and prove a weaker version of Theorem 1.1, which we refer to as the *canonical DLR equations*, which involves conditioning on the number of particles lying inside Λ. The term *canonical* refers to the fact that the number of particles is fixed, as in the *canonical* ensemble of statistical physics or the *canonical* Gibbs measure, in contrast with e.g. a *grand canonical* setting.

Theorem 2.1. For any $\beta > 0$ and any bounded Borel set $\Lambda \subset \mathbb{R}$ the following holds true.

(B) Definiteness of the exterior potential. For any $x \in \Lambda$ and Sine_β-a.e. γ, the following limit exists and is positive,

$$\omega(x|\gamma_{\Lambda^c}) := \lim_{p \to \infty} \prod_{\substack{u \in \gamma_{\Lambda^c} \\ |u| \le p}} \left|1 - \frac{x}{u}\right|^\beta. \tag{2.1}$$

Moreover, for Sine_β-a.e. γ, the partition function

$$Z(\gamma_{\Lambda^c}, |\gamma_\Lambda|) := \int_{\Lambda^{|\gamma_\Lambda|}} \prod_{j<\ell}^{|\gamma_\Lambda|} |x_j - x_\ell|^\beta \prod_{j=1}^{|\gamma_\Lambda|} \omega(x_j|\gamma_{\Lambda^c}) \mathrm{d}x_j \tag{2.2}$$

is finite and positive.

(C*) Canonical DLR equations. Let $f : \mathsf{Conf}(\mathbb{R}) \to \mathbb{R}$ be a bounded, measurable function. We have the identity

$$\mathbb{E}_{\mathrm{Sine}_\beta}[f] = \int \left[\int f(\{x_1, \dots, x_{|\gamma_\Lambda|}\} \cup \gamma_{\Lambda^c}) \, \rho_{\Lambda^c}(x_1, \dots, x_{|\gamma_\Lambda|}) \prod_{j=1}^{|\gamma_\Lambda|} \mathrm{d}x_j \right] \mathrm{Sine}_\beta(\mathrm{d}\gamma),$$

with ρ_{Λ^c} defined similarly to (1.7) by

$$\rho_{\Lambda^c}(x_1, \dots, x_{|\gamma_\Lambda|}) := \frac{1}{Z(\gamma_{\Lambda^c}, |\gamma_\Lambda|)} \prod_{j<k}^{|\gamma_\Lambda|} |x_j - x_k|^\beta \prod_{j=1}^{|\gamma_\Lambda|} \omega(x_j|\gamma_{\Lambda^c}). \tag{2.3}$$

Parts (B) of Theorem 2.1 and Theorem 1.1 are the same statement. It will be thus enough to combine Part (A) of Theorem 1.1 with Theorem 2.1 in order to obtain Part (C) of Theorem 1.1 and to conclude the proof - this shall be done in Section 3 and will strongly rely on the *canonical* form of the DLR equations.

The remainder of this section is devoted to the proof of Theorem 2.1.

Some terminology.
 ⋄ We say that a measurable function f is *local* if there exists a compact subset $K \subset \mathbb{R}$ such that $f(\gamma) = f(\gamma_K)$ for every $\gamma \in \mathsf{Conf}(\mathbb{R})$.
 ⋄ We use the notation Λ_p for the line segment $[-\frac{p}{2}, \frac{p}{2}]$.
 ⋄ If n is an integer, and Λ a bounded Borel set, we denote by $\mathbf{B}_{n,\Lambda}(\mathrm{d}\eta)$ the law of the Bernoulli process drawing a random configuration of n independent points uniformly in Λ.

For the rest of this section, we fix $\Lambda \subset \mathbb{R}$ a bounded Borel set and $f : \mathsf{Conf}(\mathbb{R}) \to \mathbb{R}$ a bounded Borel *local* function. The integer p used below is always assumed large enough so that $\Lambda \subset \Lambda_p$.

2.1 Move functions, Gibbs kernels, and proof of Theorem 2.1(B)

First, we introduce *move functions* and *Gibbs kernels*, and use them to rephrase Theorem 2.1.

2.1.1 Heuristics for the move functions

The basic idea is the following: for a given bounded Borel set Λ we want to define a Gibbs measure, formally written $\mathsf{Gibbs}_\Lambda(\,\cdot\,|\mathsf{Ext})$ on interior configurations Int in Λ, given an exterior configuration Ext in Λ^c, namely

$$\mathsf{Gibbs}_\Lambda(\mathsf{Int}|\mathsf{Ext}) \propto \exp\left(-\beta\,\mathsf{Energy}\left(\mathsf{Int}\cup\mathsf{Ext}\right)\right).$$

The physical energy reads

$$\mathsf{Energy}\left(\mathsf{Int}\cup\mathsf{Ext}\right) = \mathsf{Int}\times\mathsf{Int} + \mathsf{Ext}\times\mathsf{Ext} + 2\cdot\mathsf{Ext}\times\mathsf{Int},$$

where the product sign means "interact with", but the exterior configuration Ext could be infinite, yielding two infinite terms in the definition of the energy. Since Ext is fixed, the interaction term $\mathsf{Ext}\times\mathsf{Ext}$ can be absorbed by the normalization constant, and we are left with

$$\mathsf{Gibbs}_\Lambda(\mathsf{Int}|\mathsf{Ext}) \propto \exp\left(-\beta\left(\mathsf{Int}\times\mathsf{Int} + 2\cdot\mathsf{Ext}\times\mathsf{Int}\right)\right). \tag{2.4}$$

To deal with the fact that $\mathsf{Int}\times\mathsf{Ext}$ could correspond to an infinite, or undefined, summation, we introduce the move functions: we fix some reference interior configuration Int_0 and we define $\mathsf{Move}_\Lambda(\mathsf{Int},\mathsf{Ext})$ as the energetic cost to move the points in Λ from Int_0 to those of Int, as felt by the points of Ext, namely:

$$\mathsf{Move}_\Lambda(\mathsf{Int},\mathsf{Ext}) := 2\cdot\mathsf{Ext}\times\mathsf{Int} - 2\cdot\mathsf{Ext}\times\mathsf{Int}_0 = 2\cdot\mathsf{Ext}\times(\mathsf{Int}-\mathsf{Int}_0). \tag{2.5}$$

Thus we can write,

$$2\cdot\mathsf{Ext}\times\mathsf{Int} = \mathsf{Move}_\Lambda(\mathsf{Int},\mathsf{Ext}) + 2\cdot\mathsf{Ext}\times\mathsf{Int}_0. \tag{2.6}$$

The second term in the right-hand side of (2.6) is independent of Int and can again be absorbed by the normalization constant in (2.4), yielding

$$\mathsf{Gibbs}_\Lambda(\mathsf{Int}|\mathsf{Ext}) \propto \exp\left(-\beta\left(\mathsf{Int}\times\mathsf{Int} + \mathsf{Move}_\Lambda(\mathsf{Int},\mathsf{Ext})\right)\right),$$

which is the form of the *Gibbs kernels* (or *Gibbs specifications*) that we introduce more precisely below. If Ext is infinite, it is still not clear why $\mathsf{Move}_\Lambda(\mathsf{Int},\mathsf{Ext})$ would yield a finite quantity, but we can hope for some compensation between the two terms $\mathsf{Ext}\times\mathsf{Int}$ and $\mathsf{Ext}\times\mathsf{Int}_0$. These manipulations are valid in the finite case, and in the case of an infinite exterior configuration the rigorous approach consists in truncating the said configuration outside some large, but finite domain, and to show that the infinite-volume limit of these partial move functions exists, as we shall do next.

2.1.2 Gibbs kernels (finite window)

Definition 2.2 (Logarithmic interaction)**.** We set for convenience[1]

$$g(x) := \begin{cases} -\log|x| & \text{if } x \in \mathbb{R}\setminus\{0\} \\ 0 & \text{if } x = 0 \end{cases} \tag{2.7}$$

---

[1]This allows to include the diagonal in double sums, since $g(x_i - x_i)$ is set to be zero. Of course, g is not continuous at 0, but neither is log...

For any $p \geq 1$, any $\gamma, \eta \in \mathsf{Conf}(\mathbb{R})$, we introduce[2]:

$$\text{(Interaction energy)} \qquad \mathsf{H}_\Lambda(\gamma) := \frac{1}{2} \iint g(x-y) \, \mathrm{d}\gamma_\Lambda^{\otimes 2} \tag{2.8}$$

$$\text{(Move function)} \qquad \mathsf{M}_{\Lambda,\Lambda_p}(\eta,\gamma) := \iint g(x-y) \, \mathrm{d}(\eta_\Lambda - \gamma_\Lambda) \otimes \gamma_{\Lambda_p \backslash \Lambda} \tag{2.9}$$

$$\text{(Normalization)} \qquad Z_{\Lambda,\Lambda_p}(\gamma) := \int \mathrm{e}^{-\beta(\mathsf{H}_\Lambda(\eta) + \mathsf{M}_{\Lambda,\Lambda_p}(\eta,\gamma))} \, \mathbf{B}_{|\gamma_\Lambda|,\Lambda}(\mathrm{d}\eta) \tag{2.10}$$

$$\text{(Gibbs kernel)} \qquad \mathsf{G}_{\Lambda,\Lambda_p}(\mathrm{d}\eta,\gamma) := \frac{1}{Z_{\Lambda,\Lambda_p}(\gamma)} \mathrm{e}^{-\beta(\mathsf{H}_\Lambda(\eta) + \mathsf{M}_{\Lambda,\Lambda_p}(\eta,\gamma))} \, \mathbf{B}_{|\gamma_\Lambda|,\Lambda}(\mathrm{d}\eta). \tag{2.11}$$

The measure $\mathsf{G}_{\Lambda,\Lambda_p}(\mathrm{d}\eta,\gamma)$ is the central object for the DLR formalism, called a *Gibbs kernel*, or *specification*. Moreover we set,

$$f_{\Lambda,\Lambda_p}(\gamma) := \int f(\eta \cup \gamma_{\Lambda_p \backslash \Lambda}) \, \mathsf{G}_{\Lambda,\Lambda_p}(\mathrm{d}\eta,\gamma). \tag{2.12}$$

The right-hand side of (2.12) should be read as follows: take $\gamma \in \mathsf{Conf}(\mathbb{R})$ and a test function f. Keep the configuration γ in $\Lambda_p \backslash \Lambda$ and sample a new set of points η inside Λ with the same number of points as the old one, namely $|\gamma_\Lambda|$. This sampling is done according to the Gibbs measure $\mathsf{G}_{\Lambda,\Lambda_p}(\mathrm{d}\eta,\gamma)$ defined in (2.11), associated to the inverse temperature β and the energy $\mathsf{H}_\Lambda(\eta) + \mathsf{M}_{\Lambda,\Lambda_p}(\eta,\gamma)$, which represents the sum of the interaction energy of the points of η (inside Λ) with themselves and the cost of moving points from the old configuration γ_Λ to the new one η, as felt by the points in $\Lambda_p \backslash \Lambda$ (in particular, the "reference interior configuration" Int_0 as in (2.5) is here chosen to be γ_Λ). Then, combine η and $\gamma_{\Lambda_p \backslash \Lambda}$ into a configuration in Λ_p, and test it against the function f. The quantity $f_{\Lambda,\Lambda_p}(\gamma)$ is the expectation of this operation; let us emphasize that the randomness comes from the re-sampling in Λ, while γ_{Λ^c} is fixed.

2.1.3 Gibbs kernels (infinite window) and reformulation of Theorem 2.1

We will show in Lemma 2.15 below that the limit $p \to \infty$ of the quantities in Definition 2.2 exists. In particular, we obtain:

Lemma 2.3 (Existence of the move functions)**.** For Sine$_\beta$-a.e. configuration γ and any $\eta \in \mathsf{Conf}(\Lambda)$ satisfying $|\gamma_\Lambda| = |\eta|$, the limit

$$\mathsf{M}_{\Lambda,\mathbb{R}}(\eta,\gamma) := \lim_{p \to \infty} \mathsf{M}_{\Lambda,\Lambda_p}(\eta,\gamma) \tag{2.13}$$

exists and is finite. Moreover, for γ fixed, the convergence in (2.13) is uniform in the choice of η.

This allows us to extend the notation of Definition 2.2 to "$p = +\infty$", namely to give a meaning to

$$Z_{\Lambda,\mathbb{R}} := \int \mathrm{e}^{-\beta(\mathsf{H}_\Lambda(\eta) + \mathsf{M}_{\Lambda,\mathbb{R}}(\eta,\gamma))} \, \mathbf{B}_{|\gamma_\Lambda|,\Lambda}(\mathrm{d}\eta). \tag{2.14}$$

$$\mathsf{G}_{\Lambda,\mathbb{R}}(\mathrm{d}\eta,\gamma) := \frac{1}{Z_{\Lambda,\mathbb{R}}(\gamma)} \mathrm{e}^{-\beta(\mathsf{H}_\Lambda(\eta) + \mathsf{M}_{\Lambda,\mathbb{R}}(\eta,\gamma))} \, \mathbf{B}_{|\gamma_\Lambda|,\Lambda}(\mathrm{d}\eta). \tag{2.15}$$

$$f_{\Lambda,\mathbb{R}}(\gamma) := \int f(\eta \cup \gamma_{\Lambda^c}) \, \mathsf{G}_{\Lambda,\mathbb{R}}(\mathrm{d}\eta,\gamma). \tag{2.16}$$

---

[2]In these definitions, the $1/2$ factor has been introduced for aesthetic reasons, so that the usual CβE is an approximation of Sine$_\beta$.

Remark 2.4. At this point, an important observation is that the probability measure $\mathsf{G}_{\Lambda,\mathbb{R}}(\mathrm{d}\eta,\gamma)$ only depends on the exterior γ_{Λ^c} and the number of points $|\gamma_\Lambda|$. Indeed, if one changes γ_Λ to another configuration with the same number of points, then the move functions is changed by an additive constant which can be incorporated into a new partition function.

Moreover, Lemma 2.3 already yields part (B) of the theorem.

Proof of Theorem 2.1, part (B). Since Sine$_\beta$ is stationary, we know that 0 is not a point of γ for Sine$_\beta$-a.e γ. Moreover, since $|\gamma_\Lambda| = |\eta|$, one can write

$$\mathsf{M}_{\Lambda,\Lambda_p}(\eta,\gamma) = \sum_{\substack{u\in\gamma_{\Lambda^c} \\ |u|\leq p}} \left(\sum_{x\in\gamma_\Lambda} - \sum_{x\in\eta} \right) \log\left| 1 - \frac{x}{u} \right|,$$

where the sum over the points of η is to be understood with multiplicity. This yields

$$\mathrm{e}^{-\beta\mathsf{M}_{\Lambda,\Lambda_p}(\eta,\gamma)} = \frac{\prod_{x\in\eta}\left(\prod_{\substack{u\in\gamma_{\Lambda^c} \\ |u|\leq p}} |1 - \frac{x}{u}|^\beta \right)}{\prod_{y\in\gamma}\left(\prod_{\substack{u\in\gamma_{\Lambda^c} \\ |u|\leq p}} |1 - \frac{y}{u}|^\beta \right)} = \frac{\prod_{x\in\eta}\left(\prod_{\substack{u\in\gamma_{\Lambda^c} \\ |u|\leq p}} |1 - \frac{x}{u}|^\beta \right)}{\mathrm{e}^{-\beta\mathsf{M}_{\Lambda,\Lambda_p}(|\gamma_\Lambda|\delta_0,\gamma)}}, \tag{2.17}$$

Letting $p\to\infty$, and using Lemma 2.3 we obtain the existence of the weight $\omega(\cdot|\gamma_{\Lambda^c})$, as defined in (2.1).

Clearly, $Z_{\Lambda,\Lambda_p}(\gamma) \in (0,+\infty)$ for every p, and the convergence (2.13) is uniform in η, hence $Z_{\Lambda,\mathbb{R}}(\gamma) \in (0,+\infty)$. We may observe that, by definition of $Z(\gamma_{\Lambda^c},|\gamma_\Lambda|)$, as in (2.2), and by (2.17), we have $Z_{\Lambda,\mathbb{R}}(\gamma) = Z(\gamma_{\Lambda^c},|\gamma_\Lambda|))$, which concludes the proof. $\qquad\square$

Finally, let us observe that part (C*) of Theorem 2.1 can be written as

$$\mathbb{E}_{\text{Sine}_\beta}(f - f_{\Lambda,\mathbb{R}}) = 0. \tag{2.18}$$

The remainder of this section is devoted to prove (2.18). To do so, we rely on the convergence of the microscopic statistics of CβE towards Sine$_\beta$.

2.2 Circular ensembles and the associated DLR equations

2.2.1 Circular ensembles as model of log-gases

As mentioned in the introduction, the one-dimensional log-gas is related to the *Circular ensembles* appearing in random matrix theory. In the CβE, see e.g. Forrester [2010], the joint law of the eigenvalues $(\mathrm{e}^{i\theta_1},\ldots,\mathrm{e}^{i\theta_n})$ is given by

$$\frac{1}{C_{n,\beta}} \prod_{j<\ell} \left| \mathrm{e}^{i\theta_j} - \mathrm{e}^{i\theta_\ell} \right|^\beta \prod_{j=1}^n \mathrm{d}\theta_j, \tag{2.19}$$

where $\mathrm{d}\theta_j$ is the Lebesgue measure on $[-\pi,\pi]$. The normalization constant $C_{n,\beta}$ is known:

$$C_{n,\beta} := (2\pi)^n \frac{\Gamma(\frac{\beta}{2}n + 1)}{\Gamma(\frac{\beta}{2} + 1)^n}. \tag{2.20}$$

In this section, the central role is played by the point process $Q_{n,\beta}$, defined as the push-forward of the density in (2.19) by the map

$$(\theta_1, \ldots, \theta_n) \mapsto \left(\frac{n\theta_1}{2\pi}, \ldots, \frac{n\theta_n}{2\pi} \right).$$

The realizations of $Q_{n,\beta}$ are point configurations in $\Lambda_n = [-\frac{n}{2}, \frac{n}{2}]$, that we will use as finite window approximations of Sine_β. It can be equivalently defined as follows.

Definition 2.5 (Periodic logarithmic interaction)**.** For any $n \geq 1$, we set

$$g_n(x) := \begin{cases} -\log \left| 2 \sin \left(\frac{\pi x}{n} \right) \right| & \text{if } x \in \mathbb{R} \setminus \{0\} \\ 0 & \text{if } x = 0 \end{cases}. \tag{2.21}$$

For any $\gamma \in \mathsf{Conf}(\mathbb{R})$, we introduce the interaction energy

$$\mathsf{H}_\Lambda^{n-\mathsf{per}}(\gamma) := \frac{1}{2} \iint g_n(x - y) \mathrm{d}\gamma_\Lambda^{\otimes 2}. \tag{2.22}$$

Definition 2.6 (The canonical Gibbs measure of the finite periodic log-gas)**.** For any $\beta > 0$, we consider the point process of n points in Λ_n given by

$$Q_{n,\beta}(\mathrm{d}\gamma) := \frac{1}{Z_{n,\beta}} \, \mathrm{e}^{-\beta \mathsf{H}_{\Lambda_n}^{n-\mathsf{per}}(\gamma)} \mathbf{B}_{n,\Lambda_n}(\mathrm{d}\gamma), \tag{2.23}$$

where the partition function is given by, see (2.20),

$$Z_{n,\beta} := \int \mathrm{e}^{-\beta \mathsf{H}_{\Lambda_n}^{n-\mathsf{per}}(\gamma)} \mathbf{B}_{n,\Lambda_n}(\mathrm{d}\gamma) = \frac{C_{n,\beta}}{(2\pi)^n} = \frac{\Gamma(\frac{\beta}{2}n + 1)}{\Gamma(\frac{\beta}{2} + 1)^n}. \tag{2.24}$$

The probability measure $Q_{n,\beta}$ can be taken as a model of a finite log-gas: (2.23) gives a rigorous meaning to the informal definition (1.1). Let us emphasize that the choice of a *periodic* logarithmic potential is not the usual one, but it is more convenient for us, and yields the same microscopic limit.

2.2.2 Gibbs kernels (periodic setting)

We introduce the following notation, which should be compared to Definition 2.2. We add the superscript $n - \mathsf{per}$, in order to stress the fact that the n-periodic logarithmic interaction is used.

Definition 2.7. For any $\gamma, \eta \in \mathsf{Conf}(\mathbb{R})$,

$$\text{(Move function)} \quad \mathsf{M}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\eta, \gamma) := \iint g_n(x - y) \, \mathrm{d}(\eta_\Lambda - \gamma_\Lambda) \otimes \gamma_{\Lambda_p \setminus \Lambda} \tag{2.25}$$

$$\text{(Normalization)} \quad Z_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma) := \int \mathrm{e}^{-\beta(\mathsf{H}_\Lambda^{n-\mathsf{per}}(\eta) + \mathsf{M}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\eta,\gamma))} \mathbf{B}_{|\gamma_\Lambda|,\Lambda}(\mathrm{d}\eta) \tag{2.26}$$

$$\text{(Gibbs kernel)} \quad \mathsf{G}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\mathrm{d}\eta, \gamma) := \frac{1}{Z_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma)} \, \mathrm{e}^{-\beta(\mathsf{H}_\Lambda^{n-\mathsf{per}}(\eta) + \mathsf{M}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\eta,\gamma))} \mathbf{B}_{|\gamma_\Lambda|,\Lambda}(\mathrm{d}\eta). \tag{2.27}$$

We moreover use the notation,

$$f_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma) := \int f(\eta \cup \gamma_{\Lambda_p \setminus \Lambda}) \, \mathsf{G}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\mathrm{d}\eta, \gamma). \tag{2.28}$$

2.2.3 DLR equations for the Circular ensemble

The first step towards the canonical DLR equations (2.18) is the following result.

Proposition 2.8 (canonical DLR equations for the finite log-gas)**.** For any bounded Borel set $\Lambda \subset \Lambda_n$ and any bounded Borel function $f : \mathsf{Conf}(\Lambda_n) \to \mathbb{R}$, we have

$$\mathbb{E}_{Q_{n,\beta}}(f - f^{n-\mathsf{per}}_{\Lambda,\Lambda_n}) = 0. \tag{2.29}$$

Proof. Let us write for convenience $\Lambda^c := \Lambda_n \setminus \Lambda$. The definitions (2.27)–(2.23) of $f^{n-\mathsf{per}}_{\Lambda,\Lambda_n}$ and $Q_{n,\beta}$ yield

$$Z_{n,\beta}\, \mathbb{E}_{Q_{n,\beta}}(f^{n-\mathsf{per}}_{\Lambda,\Lambda_n})$$
$$= \iint f(\eta \cup \gamma_{\Lambda^c}) \frac{1}{Z^{n-\mathsf{per}}_{\Lambda,\Lambda_n}(\gamma)} e^{-\beta(\mathsf{H}^{n-\mathsf{per}}_{\Lambda}(\eta) + \mathsf{M}^{n-\mathsf{per}}_{\Lambda,\Lambda_n}(\eta,\gamma))} e^{-\beta \mathsf{H}^{n-\mathsf{per}}_{\Lambda_n}(\gamma)} \mathbf{B}_{|\gamma_\Lambda|,\Lambda}(\mathrm{d}\eta) \mathbf{B}_{n,\Lambda_n}(\mathrm{d}\gamma).$$

Next, we use the following algebraic identity: for any $\eta \in \mathsf{Conf}(\Lambda)$,

$$\mathsf{H}^{n-\mathsf{per}}_{\Lambda_n}(\gamma) + \mathsf{M}^{n-\mathsf{per}}_{\Lambda,\Lambda_n}(\eta,\gamma) + \mathsf{H}^{n-\mathsf{per}}_{\Lambda}(\eta) = \mathsf{H}^{n-\mathsf{per}}_{\Lambda_n}(\eta \cup \gamma_{\Lambda^c}) + \mathsf{H}^{n-\mathsf{per}}_{\Lambda}(\gamma). \tag{2.30}$$

It can be easily checked from the definitions; indeed, using the informal notation "$\times$" for "interact with", we can write

$$(\eta \cup \gamma_{\Lambda^c})^{\times 2} = \eta^{\times 2} + \gamma_{\Lambda^c}^{\times 2} + 2 \cdot \eta \times \gamma_{\Lambda^c}$$
$$\gamma^{\times 2} = (\gamma_\Lambda \cup \gamma_{\Lambda^c})^{\times 2} = \gamma_\Lambda^{\times 2} + \gamma_{\Lambda^c}^{\times 2} + 2 \cdot \gamma_\Lambda \times \gamma_{\Lambda^c},$$

and substracting the second line from the first one, we obtain (2.30). Combined with the relation,

$$Z^{n-\mathsf{per}}_{\Lambda,\Lambda_n}(\gamma) = Z^{n-\mathsf{per}}_{\Lambda,\Lambda_n}(\eta \cup \gamma_{\Lambda^c})\, e^{-\beta \mathsf{M}_{\Lambda,\Lambda_n}(\eta,\gamma)} = Z^{n-\mathsf{per}}_{\Lambda,\Lambda_n}(\eta \cup \gamma_{\Lambda^c})\, e^{\beta \mathsf{M}_{\Lambda,\Lambda_n}(\gamma,\eta \cup \gamma_{\Lambda^c})},$$

this yields

$$Z_{n,\beta}\, \mathbb{E}_{Q_{n,\beta}}(f^{n-\mathsf{per}}_{\Lambda,\Lambda_n}) = \iint f(\eta \cup \gamma_{\Lambda^c}) \frac{1}{Z^{n-\mathsf{per}}_{\Lambda,\Lambda_n}(\eta \cup \gamma_{\Lambda^c})}$$
$$\times e^{-\beta(\mathsf{H}^{n-\mathsf{per}}_{\Lambda_n}(\eta \cup \gamma_{\Lambda^c}) + \mathsf{M}_{\Lambda,\Lambda_n}(\gamma,\eta \cup \gamma_{\Lambda^c}) + \mathsf{H}^{n-\mathsf{per}}_{\Lambda}(\gamma))} \mathbf{B}_{|\gamma_\Lambda|,\Lambda}(\mathrm{d}\eta) \mathbf{B}_{n,\Lambda_n}(\mathrm{d}\gamma).$$

Making the change of variables $(\eta, \gamma) \mapsto (\zeta, \xi)$ with $\xi := \eta \cup \gamma_{\Lambda^c}$ and $\zeta := \gamma_\Lambda$, we obtain

$$Z_{n,\beta}\, \mathbb{E}_{Q_{n,\beta}}(f^{n-\mathsf{per}}_{\Lambda,\Lambda_n})$$
$$= \iint f(\xi) \frac{1}{Z^{n-\mathsf{per}}_{\Lambda,\Lambda_n}(\xi)} e^{-\beta(\mathsf{H}^{n-\mathsf{per}}_{\Lambda_n}(\xi) + \mathsf{M}_{\Lambda,\Lambda_n}(\zeta,\xi) + \mathsf{H}^{n-\mathsf{per}}_{\Lambda}(\zeta))} \mathbf{B}_{|\xi_\Lambda|,\Lambda}(\mathrm{d}\zeta) \mathbf{B}_{n,\Lambda_n}(\mathrm{d}\xi)$$
$$= Z_{n,\beta}\, \mathbb{E}_{Q_{n,\beta}}(f),$$

and the proposition is proved. $\qquad\square$

2.2.4 Local convergence of the circular ensembles to Sine-beta

As mentioned in the introduction, the point process $Q_{n,\beta}$ is known to converge towards Sine_β as $n \to \infty$. This convergence was first studied by Killip and Stoiciu [2009]. They have shown that for any C^∞ function $\psi : \mathbb{R} \to \mathbb{R}$ with compact support,

$$\lim_{n \to \infty} \mathbb{E}_{Q_{n,\beta}} \left[\exp \left(\int \psi \, d\gamma \right) \right] = \mathbb{E}_{\text{Sine}_\beta} \left[\exp \left(\int \psi \, d\gamma \right) \right]. \tag{2.31}$$

To be more precise, the limiting process was called $C_\beta E$ in [Killip and Stoiciu, 2009]; almost at the same time, Valkó and Virág [2009] have shown a similar result for the bulk limit of the Gaussian β-ensemble, with a limiting process called Sine_β. A bit later, it has been observed that Sine_β and $C_\beta E$ are identical, so that the result from [Killip and Stoiciu, 2009] can now be stated under the form (2.31); see [Nakano, 2014, Corollary 1.7] or [Valkó and Virág, 2017, Theorem 28]. For our purpose, we will need a slightly stronger convergence result that we state in the next proposition.

Proposition 2.9. The sequence of point processes $(Q_{n,\beta})_{n \geq 1}$ converges to the point process Sine_β in the topology of local convergence: for any bounded, Borel and local test function $\varphi : \text{Conf}(\mathbb{R}) \to \mathbb{R}$, we have

$$\lim_{n \to \infty} \mathbb{E}_{Q_{n,\beta}}[\varphi] = \mathbb{E}_{\text{Sine}_\beta}[\varphi].$$

Proof. The convergence (2.31) implies the weak convergence $Q_{n,\beta} \to \text{Sine}_\beta$, see [Killip and Stoiciu, 2009, Definition 1.2 and 1.3] and [Daley and Vere-Jones, 2008, Proposition 11.1.VIII].

Now, we show that the sequence $(Q_{n,\beta})_{n \geq 1}$ has an accumulation point in the local topology. Indeed, consider the relative entropy of two point processes P, Q defined by

$$I(P|Q) := \int \log \frac{dP}{dQ} \, dP$$

when P has a density with respect to Q and set $I(P|Q) := +\infty$ otherwise. Let Π_{Λ_n} be the Poisson point process of intensity 1 on Λ_n. According to [Georgii and Zessin, 1993, Proposition 2.6], it is enough to check that

$$\sup_{n \in \mathbb{N}^*} \frac{1}{n} I(Q_{n,\beta}|\Pi_{\Lambda_n}) < \infty. \tag{2.32}$$

$$\begin{aligned}
I(Q_{n,\beta}|\Pi_{\Lambda_n}) &= \int \log \frac{dQ_{n,\beta}}{d\mathbf{B}_{n,\Lambda_n}} dQ_{n,\beta} + \int \log \frac{d\mathbf{B}_{n,\Lambda_n}}{d\Pi_{\Lambda_n}} dQ_{n,\beta} \\
&= -\log Z_{n,\beta} - \beta \, \mathbb{E}_{Q_{n,\beta}} \left[\mathsf{H}_{\Lambda_n}^{n-\text{per}}(\gamma) \right] - \log \left(e^{-n} \frac{n^n}{n!} \right).
\end{aligned}$$

Recalling (2.24) and using Stirling formula, we see that

$$\lim_{n \to \infty} \frac{1}{n} \log Z_{n,\beta} - \frac{\beta}{2} \log \frac{n}{2\pi} \tag{2.33}$$

exists and is finite. Moreover,

$$\mathbb{E}_{Q_{n,\beta}}\left[\mathsf{H}_{\Lambda_n}^{n-\mathsf{per}}(\gamma)\right] = \frac{1}{Z_{n,\beta}}\int \mathsf{H}_{\Lambda_n}^{n-\mathsf{per}}(\gamma)\mathrm{e}^{-\beta\mathsf{H}_{\Lambda_n}^{n-\mathsf{per}}(\gamma)}\mathbf{B}_{n,\Lambda_n}(\mathrm{d}\gamma)$$

$$= -\frac{\mathrm{d}}{\mathrm{d}\beta}\log Z_{n,\beta} = -\frac{\mathrm{d}}{\mathrm{d}\beta}\log\frac{\Gamma(\frac{\beta}{2}n+1)}{\Gamma(\frac{\beta}{2}+1)^n}$$

which yields that

$$\frac{1}{n}\mathbb{E}_{Q_{n,\beta}}\left[\mathsf{H}_{\Lambda_n}^{n-\mathsf{per}}(\gamma)\right] + \frac{1}{2}\log\frac{n}{2\pi} = -\frac{1}{2}\psi\left(\frac{\beta}{2}n+1\right) + \frac{1}{2}\psi\left(\frac{\beta}{2}+1\right) + \frac{1}{2}\log n - \frac{1}{2}\log(2\pi),$$

with $\psi(x) := (\log\Gamma(x))'$ the so-called digamma function. Moreover using that, $\psi(x) = \log x - \frac{1}{2x} + o\left(\frac{1}{x}\right)$ as $x \to \infty$, see [Abramowitz and Stegun, 1964, 6.3.18], we obtain that $\log n - \psi(\frac{\beta}{2}n+1)$ remains bounded. We get

$$\sup_n\left|\frac{1}{n}\mathbb{E}_{Q_{n,\beta}}\left[\mathsf{H}_{\Lambda_n}^{n-\mathsf{per}}(\gamma)\right] + \frac{1}{2}\log\frac{n}{2\pi}\right| < +\infty. \tag{2.34}$$

Putting (2.33) and (2.34) together, we obtain (2.32), which proves the claim.

Finally, since the convergence in the local topology is stronger than the weak convergence of point processes, see [Georgii and Zessin, 1993, Section 2.1], any accumulation point of $(Q_{n,\beta})_{n\geq 1}$ has to be Sine_β. Moreover, (2.32) also provides that Sine_β is a point process, namely charges only simple configurations. Indeed, this yields that Sine_β has finite specific entropy and thus has local densities with respect to the Poisson process, see [Georgii, 2011, Chapter15]. The proof of the proposition is therefore complete.

$\square$

2.3   From the finite, periodic DLR equations to Theorem 2.1 (C*)

To prove Theorem 2.1 (C*), namely that

$$\mathbb{E}_{\mathrm{Sine}_\beta}(f - f_{\Lambda,\mathbb{R}}) = 0,$$

we start from Proposition 2.8

$$\mathbb{E}_{Q_{n,\beta}}(f - f_{\Lambda,\Lambda_n}^{n-\mathsf{per}}) = 0,$$

use the convergence of $Q_{n,\beta}$ to Sine_β as expressed by Proposition 2.9, and perform several approximations.

⋄ First, we show that one can replace $f_{\Lambda,\Lambda_n}^{n-\mathsf{per}}$ by $f_{\Lambda,\Lambda_p}^{n-\mathsf{per}}$ in the DLR equations for $Q_{n,\beta}$. The contribution of the exterior configuration is indeed negligible when forgetting about the configuration in $\Lambda_n \setminus \Lambda_p$.

⋄ Next, we prove that one can further replace $f_{\Lambda,\Lambda_p}^{n-\mathsf{per}}$ by f_{Λ,Λ_p}, which means we can replace the periodic logarithmic interaction by the usual one up to negligible terms.

⋄ Finally, we let $n \to \infty$ and replace $Q_{n,\beta}$ by Sine_β, using Proposition 2.9. Moreover, we replace f_{Λ,Λ_p} by $f_{\Lambda,\mathbb{R}}$ in the remaining DLR equations.

This shall complete the proof of Theorem 2.1 up to the proof of technical estimates on the discrepancy, which will be deferred to Section 2.5; preliminary material for proving these estimates is provided in Section 2.4.

2.3.1 Step 1: Truncation errors in the periodic DLR setting

The DLR equations (2.29) obtained for $Q_{n,\beta}$ involve the Gibbs kernel $f_{\Lambda,\Lambda_n}^{n-\mathsf{per}}$, defined in (2.27), where the index n appears twice: as the period of the interaction, and as the size of the window. The following estimate allows us to decouple size and period.

Proposition 2.10 (Truncation error, periodic case)**.** Let $\varepsilon > 0$. For any p large enough (depending on Λ, ε), for any $n \geq p$, and for any bounded measurable test function f on $\mathsf{Conf}(\mathbb{R})$, we have

$$\left| \mathbb{E}_{Q_{n,\beta}}(f_{\Lambda,\Lambda_n}^{n-\mathsf{per}} - f_{\Lambda,\Lambda_p}^{n-\mathsf{per}}) \right| \leq \varepsilon \|f\|_\infty \,.$$

The only difference between $f_{\Lambda,\Lambda_n}^{n-\mathsf{per}}$ and $f_{\Lambda,\Lambda_p}^{n-\mathsf{per}}$ lies in the size of the exterior configuration that is taken into account. To prove Proposition 2.10, we thus need to control the difference of periodic move functions over different large windows. This motivates the following definition.

Definition 2.11 (Configurations with small truncation error)**.** For any $\delta > 0$, we denote by $\mathsf{A}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\delta)$ the set

$$\mathsf{A}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\delta) := \left\{ \gamma \in \mathsf{Conf}(\mathbb{R}) : \sup_{\substack{\eta \in \mathsf{Conf}(\Lambda) \\ |\eta|=|\gamma_\Lambda|}} \left| \mathsf{M}_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\eta,\gamma) - \mathsf{M}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\eta,\gamma) \right| \leq \delta \right\}. \qquad (2.35)$$

In plain words, if γ belongs to $\mathsf{A}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\delta)$, then we can change γ_Λ into any other configuration η in Λ *with the same number of points*, and the energy cost of this operation as felt by the points in $\Lambda_n \backslash \Lambda_p$ is always less than δ. The main ingredient for the proof of Proposition 2.10 is that $\mathsf{A}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\delta)$ has large probability under $Q_{n,\beta}$.

Lemma 2.12 (The truncation error is often small, finite case)**.** For any $\varepsilon, \delta > 0$, we have

$$Q_{n,\beta}\left(\mathsf{A}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\delta) \right) \geq 1 - \varepsilon$$

provided that $p \leq n$ are large enough, depending on $\varepsilon, \delta, \Lambda$.

Roughly speaking "the far exterior does not count". This would be obvious for a short-ranged interaction, but in the case of the logarithm we need to show that some effective cancellations occur. The proof of Lemma 2.12 is deferred to Section 2.5.2, and we now prove Proposition 2.10, using Lemma 2.12

Proof of Proposition 2.10. For any $\gamma \in \mathsf{A}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\delta)$ and $\eta \in \mathsf{Conf}(\Lambda)$ satisfying $|\eta| = |\gamma_\Lambda|$,

$$|Z_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma) - Z_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\gamma)| \leq (\mathrm{e}^{\beta\delta} - 1)Z_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma).$$

Then, writing $\Lambda^c := \Lambda_n \setminus \Lambda$ for convenience, we have for any $\gamma \in \mathsf{A}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\delta)$,

$$
\begin{aligned}
&\left| f_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\gamma) - f_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma) \right| \\
&= \left| \int f(\eta \cup \gamma_{\Lambda^c}) \, \mathsf{G}_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\mathrm{d}\eta,\gamma) - \int f(\eta \cup \gamma_{\Lambda^c}) \, \mathsf{G}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\mathrm{d}\eta,\gamma) \right| \\
&\leq \left| \int f(\eta \cup \gamma_{\Lambda^c}) \left(\mathrm{e}^{-\beta(\mathsf{M}_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\eta,\gamma) - \mathsf{M}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\eta,\gamma))} - 1 \right) \mathsf{G}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\mathrm{d}\eta,\gamma) \right| \\
&\quad + \left| \int f(\eta \cup \gamma_{\Lambda^c}) \left(\frac{Z_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma) - Z_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\gamma)}{Z_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\gamma)} \right) \mathsf{G}_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\mathrm{d}\eta,\gamma) \right| \\
&\leq (\mathrm{e}^{\beta\delta} - 1) \int |f(\eta \cup \gamma_{\Lambda^c})| \mathsf{G}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\mathrm{d}\eta,\gamma) + (\mathrm{e}^{2\beta\delta} - 1) \int |f(\eta \cup \gamma_{\Lambda^c})| \mathsf{G}_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\mathrm{d}\eta,\gamma) \\
&\leq 2(\mathrm{e}^{\beta\delta} - 1) \|f\|_\infty.
\end{aligned}
$$

Given $\varepsilon > 0$, assume that δ satisfies $2(\mathrm{e}^{\beta\delta} - 1) \leq \varepsilon$ and assume further that $p \leq n$ are large enough so that $Q_{n,\beta}(\mathsf{A}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\delta)) \geq 1 - \varepsilon$, which is possible thanks to Lemma 2.12. By using that $\|f_{\Lambda,\Lambda_p}^{n-\mathsf{per}}\|_\infty \leq \|f\|_\infty$ for any $p \leq n$, we thus obtain

$$
\begin{aligned}
\left| \mathbb{E}_{Q_{n,\beta}}(f_{\Lambda,\Lambda_n}^{n-\mathsf{per}} - f_{\Lambda,\Lambda_p}^{n-\mathsf{per}}) \right| &\leq \left| \mathbb{E}_{Q_{n,\beta}}((f_{\Lambda,\Lambda_n}^{n-\mathsf{per}} - f_{\Lambda,\Lambda_p}^{n-\mathsf{per}}) \mathbf{1}_{\mathsf{A}_{\Lambda,\Lambda_p}^n(\delta)}) \right| \\
&\quad + 2\|f\|_\infty Q_{n,\beta}\left(\mathsf{Conf}(\mathbb{R}) \setminus \mathsf{A}_{\Lambda,\Lambda_p}^n(\delta) \right) \\
&\leq 3\varepsilon \|f\|_\infty
\end{aligned}
$$

and the lemma follows since ε is arbitrary. $\qquad\square$

2.3.2 Step 2: From periodic to non-periodic interaction

Next, we show that one can replace the periodic potential g_n by the logarithmic potential g at a small cost.

Proposition 2.13 (From periodic to non-periodic potentials). Let $\varepsilon > 0$. We have

$$
\left| \mathbb{E}_{Q_{n,\beta}}(f_{\Lambda,\Lambda_p} - f_{\Lambda,\Lambda_p}^{n-\mathsf{per}}) \right| \leq \varepsilon \|f\|_\infty,
$$

provided that p is large enough (depending on Λ, ε) and $n \geq p$ is large enough (depending on p).

This is fairly intuitive: in the Gibbs kernel f_{Λ,Λ_p}, all the interactions take place between points that are at distance at most p. The precise value of the period n of the interaction is thus not really important, because for $|x| \leq p \ll n$, we have

$$
\log \left| \sin(\frac{\pi x}{n}) \right| \approx \log \left| \frac{\pi x}{n} \right| \approx \log |x|,
$$

up to an additive constant in the energy, which is irrelevant for a Gibbs specification. The only issue is that this approximation comes with a certain negligible cost *for each pair of points*, so the main ingredient that we will use in the proof of Proposition 2.13 is the

17

fact that it is unlikely under $Q_{n,\beta}$ to have too many points in a given bounded set. More precisely, if we set

$$\mathsf{B}_p := \left\{ \gamma \in \mathsf{Conf}(\mathbb{R}) : \ |\gamma_\Lambda| \le p, \ |\gamma_{\Lambda_p}| \le p^2 \right\},$$

the following estimate holds true.

Lemma 2.14 (No overcrowding)**.** For any $\varepsilon > 0$, for p large enough (depending on Λ, ε), and for $n \ge p$, we have

$$Q_{n,\beta}\left(\mathsf{B}_p \right) \ge 1 - \varepsilon.$$

The proof of Lemma 2.14 is deferred to Section 2.4.3, and we now prove Proposition 2.13 using Lemma 2.14.

Proof of Proposition 2.13. Let $\delta > 0$. For any n large enough (depending on δ and p) we have, for any $x \ne y \in \Lambda_p$,

$$\left| g_n(x-y) + \log \frac{2\pi}{n} - g(x-y) \right| = \left| \log \left| \frac{2\sin\frac{\pi(x-y)}{n}}{\frac{2\pi(x-y)}{n}} \right| \right| \le \frac{\delta}{p^3}.$$

Let $\gamma \in \mathsf{B}_p$ and $\eta \in \mathsf{Conf}(\Lambda)$ satisfying $|\eta| = |\gamma_\Lambda|$ and assume they are simple. Then, for n large enough (depending on δ, p),

$$\left| \mathsf{H}_\Lambda^{n-\mathsf{per}}(\eta) + \frac{1}{2}|\gamma_\Lambda|(|\gamma_\Lambda| - 1)\log\frac{2\pi}{n} - \mathsf{H}_\Lambda(\eta) \right| \le \frac{\delta}{p^3} \int \mathrm{d}\gamma_\Lambda^{\otimes 2} \le \frac{\delta}{p} \le \delta \tag{2.36}$$

$$\left| \mathsf{M}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\eta, \gamma) - \mathsf{M}_{\Lambda,\Lambda_p}(\eta, \gamma) \right| \le \frac{\delta}{p^3} 2|\gamma_\Lambda| \int \mathrm{d}\gamma_{\Lambda_p \setminus \Lambda} \le 2\delta, \tag{2.37}$$

where we used for the second string of inequalities that $|\eta| = |\gamma_\Lambda|$. It follows,

$$\left| Z_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma)\, \mathrm{e}^{-\frac{\beta}{2}|\gamma_\Lambda|(|\gamma_\Lambda| - 1)\log\frac{2\pi}{n}} - Z_{\Lambda,\Lambda_p}(\gamma) \right| \le (\mathrm{e}^{3\beta\delta} - 1)Z_{\Lambda,\Lambda_p}(\gamma).$$

Moreover, by writing

$$f_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma)$$
$$= \int \frac{f(\eta \cup \gamma_\Lambda)}{Z_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma)\, \mathrm{e}^{-\frac{\beta}{2}|\gamma_\Lambda|(|\gamma_\Lambda| - 1)\log\frac{2\pi}{n}}} \mathrm{e}^{-\beta(\mathsf{H}_\Lambda^{n-\mathsf{per}}(\eta) + \frac{1}{2}|\gamma_\Lambda|(|\gamma_\Lambda| - 1)\log\frac{2\pi}{n} + \mathsf{M}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\eta,\gamma))} \mathbf{B}_{|\gamma_\Lambda|,\Lambda}(\mathrm{d}\eta)$$

and using again (2.36)–(2.37) together with Lemma 2.14, the proposition is obtained by following the same lines than in the proof of Proposition 2.10. $\square$

2.3.3 Step 3: Truncation errors in the infinite DLR setting

The results of this section are valid not only for Sine_β, but for any stationary point process P on $\mathbb{R}$ with finite expected renormalized energy $\mathbb{E}_P[\mathbb{W}(\gamma)]$. We refer the reader to Section 2.4 for a precise definition but for now it is enough to keep in mind that $\mathbb{E}_{\mathrm{Sine}_\beta}[\mathbb{W}(\gamma)] < \infty$.

Lemma 2.15 (Definiteness of the move functions, infinite case)**.** Let P be a stationary point process on $\mathbb{R}$ satisfying $\mathbb{E}_P[\mathbb{W}(\gamma)] < \infty$. Then, for P-a.e. $\gamma \in \mathsf{Conf}(\mathbb{R})$ and every $\eta \in \mathsf{Conf}(\Lambda)$ satisfying $|\eta| = |\gamma_\Lambda|$, the limit

$$\mathsf{M}_{\Lambda,\mathbb{R}}(\eta,\gamma) = \lim_{p\to\infty} \mathsf{M}_{\Lambda,\Lambda_p}(\eta,\gamma)$$

exists and is finite, and the convergence is uniform for such η's.

Lemma 2.15 is proven in Section 2.5.3

We now state a result concerning the truncation error in the infinite, non-periodic setting.

Proposition 2.16 (Truncation error, infinite setting)**.** Let P be a stationary point process on $\mathbb{R}$ satisfying $\mathbb{E}_P[\mathbb{W}(\gamma)] < \infty$ and let $\varepsilon > 0$. For any p large enough (depending on Λ, ε, and P), for any bounded measurable function f, we have

$$\left| \mathbb{E}_P(f_{\Lambda,\mathbb{R}} - f_{\Lambda,\Lambda_p}) \right| \leq \varepsilon \|f\|_\infty \,.$$

Proposition 2.16 should be compared to Proposition 2.10. As for the proof of the latter, we rely on a result saying that the truncation error is often small.

The following definition is the counterpart of Definition 2.11 in the infinite setting.

Definition 2.17 (Infinite configurations with small truncation error)**.** For any $\delta > 0$, set

$$\mathsf{A}_{\Lambda,\Lambda_p}(\delta) := \left\{ \gamma \in \mathsf{Conf}(\mathbb{R}) : \sup_{\substack{\eta \in \mathsf{Conf}(\Lambda) \\ |\eta| = |\gamma_\Lambda|}} \left| \mathsf{M}_{\Lambda,\mathbb{R}}(\eta,\gamma) - \mathsf{M}_{\Lambda,\Lambda_p}(\eta,\gamma) \right| \leq \delta \right\} \,.$$

Lemma 2.18 (The truncation error is often small, infinite case)**.** For any $\varepsilon, \delta > 0$, we have for every p large enough (depending on $P, \varepsilon, \delta, \Lambda$),

$$P\left(\mathsf{A}_{\Lambda,\Lambda_p}(\delta) \right) \geq 1 - \varepsilon. \tag{2.38}$$

Lemma 2.18 is the counterpart of Lemma 2.12 in the infinite, non-periodic setting. Its proof is postponed to Section 2.5.3.

Proof of Proposition 2.16. Using Lemma 2.18, the proposition is obtained exactly as in the proof of Proposition 2.10. $\qquad\square$

2.3.4 Proof of the canonical DLR equations

We may now give the proof of the canonical DLR equations for Sine_β.

Proof of Theorem 2.1 (C*). Let $\varepsilon > 0$ and f be a bounded Borel local function on $\mathsf{Conf}(\mathbb{R})$. We write,

$$\left| \mathbb{E}_{\mathrm{Sine}_\beta} \left(f - f_{\Lambda, \mathbb{R}} \right) \right| \leq \left| \mathbb{E}_{\mathrm{Sine}_\beta}(f) - \mathbb{E}_{Q_{n,\beta}}(f) \right| \tag{A}$$

$$+ \left| \mathbb{E}_{Q_{n,\beta}}(f - f_{\Lambda, \Lambda_n}^{n-\mathsf{per}}) \right| \tag{B}$$

$$+ \left| \mathbb{E}_{Q_{n,\beta}}(f_{\Lambda, \Lambda_n}^{n-\mathsf{per}}) - \mathbb{E}_{Q_{n,\beta}}(f_{\Lambda, \Lambda_p}^{n-\mathsf{per}}) \right| \tag{C}$$

$$+ \left| \mathbb{E}_{Q_{n,\beta}}(f_{\Lambda, \Lambda_p}^{n-\mathsf{per}}) - \mathbb{E}_{Q_{n,\beta}}(f_{\Lambda, \Lambda_p}) \right| \tag{D}$$

$$+ \left| \mathbb{E}_{Q_{n,\beta}}(f_{\Lambda, \Lambda_p}) - \mathbb{E}_{\mathrm{Sine}_\beta}(f_{\Lambda, \Lambda_p}) \right| \tag{E}$$

$$+ \left| \mathbb{E}_{\mathrm{Sine}_\beta}(f_{\Lambda, \Lambda_p} - f_{\Lambda, \mathbb{R}}) \right|. \tag{F}$$

Proposition 2.8 states that (B) $= 0$, Proposition 2.10 that (C) $\leq \varepsilon \|f\|_\infty$, Proposition 2.13 that (D) $\leq \varepsilon \|f\|_\infty$, and Proposition 2.16 that (F) $\leq \varepsilon \|f\|_\infty$, provided that p is chosen large enough (depending on Λ, ε) and n is large enough (depending on p). Moreover, since f and f_{Λ, Λ_p} are both bounded Borel local functions, it follows from Proposition 2.9, that (A) and (E) can be made arbitrary small provided that n is large enough (depending on f, Λ, p), and the proof is complete under the extra assumption that f is a local function.

In order to extend the result to arbitrary, possibly non local, bounded Borel functions, we proceed as follows. Let $\mathcal{M}$ be the class of all measurable events A such that $\mathbf{1}_A$ satisfies the DLR equations, and let Π be the class of all measurable events A which are local in the sense that $\mathbf{1}_A$ is a local function as above. So far, we have proven that $\Pi \subset \mathcal{M}$. We want to prove that $\mathcal{M}$ is the whole Borel σ-algebra. The set Π is clearly stable under finite intersections. Moreover, we can check that $\mathcal{M}$ is a monotone class, using monotone convergence and the linearity of DLR equations. By the monotone class theorem, $\mathcal{M}$ contains $\sigma(\Pi)$, the σ-algebra generated by Π. Next, consider the countable collection of open sets $\{\gamma \in \mathsf{Conf}(\mathbb{R}) : |\gamma \cap (a, b)| < c\}_{a,b,c \in \mathbb{Q}} \subset \Pi$, which generate the topology of $\mathsf{Conf}(\mathbb{R})$. Since this collection is countable the σ-algebra it generates, which is included in $\sigma(\Pi)$, is the whole Borel σ-algebra. This finally shows that $\mathcal{M}$ is the whole Borel σ-algebra of $\mathsf{Conf}(\mathbb{R})$, and we have thus obtained the DLR equations for any indicator function of a Borel subset of $\mathsf{Conf}(\mathbb{R})$, which by linearity the DLR equations extends to every simple function, and finally, by density, to every bounded measurable function. $\qquad \square$

2.4 Renormalized energy and discrepancy estimates

2.4.1 Renormalized energy

We gather here the definition of the *renormalized energy*, which is a way to define the logarithmic energy of an infinite point configuration, and some useful properties. A first version of this object was introduced by Sandier and Serfaty [2012] but we use here the variation introduced in [Petrache and Serfaty, 2017]. In the present work, we do not work directly with the energy, we mostly make use of the connection between the renormalized energy and discrepancy estimates, as explained in the next paragraph.

The following definitions can be found, with more details and justification for existence, e.g. in [Leblé and Serfaty, 2017, Section 2.6].

Definition 2.19 (Compatible electric fields)**.** For any $\gamma \in \mathsf{Conf}(\mathbb{R})$, a vector field $E : \mathbb{R}^2 \to \mathbb{R}^2$ is said to be *an electric field compatible with γ*, and we write $E \in \mathsf{Comp}(\gamma)$, if it

satisfies:
$$-\mathrm{div}E = 2\pi(\gamma - \delta_{\mathbb{R}}),$$
in the sense of distributions, where by definition, the action of the measure $\delta_{\mathbb{R}}$ on a smooth and compactly supported function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ is $\int_{\mathbb{R}} \varphi(\cdot, 0)\mathrm{d}x$.

Definition 2.20 (Renormalized energy of an infinite point configuration)**.** Given any configuration $\gamma \in \mathsf{Conf}(\mathbb{R})$ and $E \in \mathsf{Comp}(\gamma)$, we first consider for any $\eta \in (0,1)$ the *regularized field* E_η: we set for any $z = (x, y) \in \mathbb{R}^2$,

$$E_\eta(z) := E(z) + \sum_{p \in \gamma} \nabla f_\eta(x - p, y), \quad \text{where } f_\eta(z) := \mathbf{1}_{|z| \geq \eta} \log \left| \frac{z}{\eta} \right| \text{ for } z \in \mathbb{R}^2.$$

Then, the *renormalized energy of γ* is defined by

$$\mathbb{W}(\gamma) := \inf_{E \in \mathsf{Comp}(\gamma)} \left\{ \lim_{\eta \to 0} \left(\limsup_{R \to \infty} \frac{1}{R} \int_{[-R,R] \times \mathbb{R}} |E_\eta(x,y)|^2 \mathrm{d}x\mathrm{d}y \right) + 2\pi \log \eta \right\}.$$

For a periodic configuration, the renormalized energy can be computed explicitly in terms of the periodic logarithmic energy of the configuration in a fundamental domain.

Proposition 2.21 (Energy of a periodic point configuration)**.** Given a configuration γ of n distinct points $\gamma := \{\gamma_1, \ldots, \gamma_n\} \in \mathsf{Conf}(\Lambda_n)$, let $\gamma^{n-\mathsf{per}} \in \mathsf{Conf}(\mathbb{R})$ be the n-periodic configuration defined by

$$\gamma^{n-\mathsf{per}} := \bigcup_{k_1, \ldots, k_n \in \mathbb{Z}} \{\gamma_1 + k_1 n, \ldots, \gamma_n + k_n n\}.$$

Then, we have

$$\mathbb{W}(\gamma^{n-\mathsf{per}}) = \frac{\pi}{n} \left(2\mathsf{H}_{\Lambda_n}^{n-\mathsf{per}}(\gamma) + n \log \frac{n}{2\pi} \right).$$

Proof. We refer to [Petrache and Serfaty, 2017, Proposition 1.5, $d = 1$] or [Borodin and Serfaty, 2013, Proposition 2.10] □

2.4.2 Discrepancy estimates

We introduce an important quantity for our purpose: the discrepancy, as well as bounds on the average discrepancy for log-gases.

Definition 2.22 (Discrepancy)**.** The *discrepancy* of $\gamma \in \mathsf{Conf}(\mathbb{R})$ relative to a bounded Borel set $\Lambda \subset \mathbb{R}$ of Lebesgue measure $|\Lambda|$ is defined by

$$\mathrm{Discr}_\Lambda(\gamma) := |\gamma_\Lambda| - |\Lambda|.$$

A crucial fact for our purpose is that a bound on the renormalized energy translates into a discrepancy estimate, see e.g. [Leblé and Serfaty, 2017, Section 3.2].

Lemma 2.23 (Energy bound yields discrepancy estimate)**.** There exists $C > 0$ such that, for any stationary point process P on $\mathbb{R}$ satisfying $\mathbb{E}_P[\mathbb{W}(\gamma)] < \infty$ and any bounded Borel set $\Lambda \subset \mathbb{R}$, we have

$$\mathbb{E}_P \left[\mathrm{Discr}_\Lambda^2(\gamma) \right] \leq C \left(C + \mathbb{E}_P \left[\mathbb{W}(\gamma) \right] \right) |\Lambda|. \tag{2.39}$$

Proof. This is [Leblé and Serfaty, 2017, Lemma 3.2]. □

In particular, for the periodic log-gas $Q_{n,\beta}$, we obtain the following bound.

Lemma 2.24. There exists a constant $C_\beta > 0$ depending on β only such that, for any bounded Borel set $\Lambda \subset \mathbb{R}$, and $n \geq 1$ large enough so that $\Lambda \subset \Lambda_n$, we have

$$\mathbb{E}_{Q_{n,\beta}}\left[\text{Discr}_\Lambda^2(\gamma)\right] \leq C_\beta|\Lambda| \quad \text{and} \quad \mathbb{E}_{Q_{n,\beta}}\left[|\gamma_\Lambda|^2\right] \leq 2|\Lambda|(C_\beta + |\Lambda|).$$

Proof. As $Q_{n,\beta}$ is the law of a stationary point process, one can apply Lemma 2.23 together with Proposition 2.21 to obtain, provided that $\Lambda \subset \Lambda_n$,

$$\mathbb{E}_{Q_{n,\beta}}\left[\text{Discr}_\Lambda^2(\gamma)\right] = \mathbb{E}_{Q_{n,\beta}}\left[\text{Discr}_\Lambda^2(\gamma^{n-\text{per}})\right]$$
$$\leq C\left(C + \frac{\pi}{n}\mathbb{E}_{Q_{n,\beta}}\left[2\mathsf{H}_{\Lambda_n}^{n-\text{per}}(\gamma) + n\log\frac{n}{2\pi}\right]\right)|\Lambda|.$$

Using (2.34), the first inequality follows. The second inequality is obtained from the first one by writing

$$\mathbb{E}_{Q_{n,\beta}}\left[|\gamma_\Lambda|^2\right] \leq 2\left(\mathbb{E}_{Q_{n,\beta}}\left[\text{Discr}_\Lambda^2(\gamma)\right] + |\Lambda|^2\right).$$
□

We will also use the following asymptotic behavior for the discrepancy.

Lemma 2.25. If $\gamma \in \text{Conf}(\mathbb{R})$ satisfies $\mathbb{W}(\gamma) < \infty$ then, as $k \to \infty$,

$$\text{Discr}_{[0,k]}(\gamma) = o(k).$$

Proof. This is a consequence of [Petrache and Serfaty, 2017, Lemma 2.1]. □

2.4.3 Proof of Lemma 2.14

An easy consequence of Lemma 2.24 is the following.

Proof of Lemma 2.14. Lemma 2.24 and Markov inequality yield the existence of $C_\Lambda > 0$ such that, for any $1 \leq p \leq n$ and n large enough so that $\Lambda \subset \Lambda_n$,

$$Q_{n,\beta}(\text{Conf}(\mathbb{R}) \setminus \mathsf{B}_p) \leq Q_{n,\beta}(|\gamma_\Lambda| > p) + Q_{n,\beta}(|\gamma_{\Lambda_p}| > p^2)$$
$$\leq \frac{1}{p^2}\mathbb{E}_{Q_{n,\beta}}(|\gamma_\Lambda|^2) + \frac{1}{p^4}\mathbb{E}_{Q_{n,\beta}}(|\gamma_{\Lambda_p}|^2) \leq \frac{C_\Lambda}{p^2},$$

and the lemma follows. □

2.5 Auxiliary proofs

We now provide proofs for Lemmas 2.12, 2.15, and 2.18.

In this section, we always assume $p \geq n$ are large enough so that $3\Lambda \subset \Lambda_p \subset \Lambda_n$. We use the following notation:

⋄ The distance of $x \in \mathbb{R}$ to a subset $\Lambda \subset \mathbb{R}$ is $\text{dist}(x, \Lambda) := \inf_{y \in \Lambda} |x - y|$, and the distance from a subset $I \subset \mathbb{R}$ to Λ is $\text{dist}(I, \Lambda) := \inf_{x \in I} \text{dist}(x, \Lambda)$.

◇ For any $\gamma_1, \gamma_2 \in \mathsf{Conf}(\mathbb{R})$ with $|\gamma_1| = |\gamma_2| = M < \infty$, say $\gamma_j = \sum_{i=1}^M \delta_{\gamma_j^i}$, we set

$$W_1(\gamma_1, \gamma_2) := \inf_{\sigma \in \mathfrak{S}_M} \sum_{i=1}^M |\gamma_1^i - \gamma_2^{\sigma(i)}| \tag{2.40}$$

where $\mathfrak{S}_M$ is the set of permutations of $\{1, \dots, M\}$. Note that the definition does not depend on the indexing.

◇ We denote by Leb the Lebesgue measure of $\mathbb{R}$ and by Leb_Λ its restriction to $\Lambda \subset \mathbb{R}$.

2.5.1 Intermediary results

Lemma 2.26 (The electrostatic potential generated when moving points)**.** Take any configurations γ, η in $\mathsf{Conf}(\mathbb{R})$ such that $|\eta| = |\gamma_\Lambda|$. Recalling the definitions (2.7) and (2.21), we set for convenience

$$\Psi_n := g_n * (\eta - \gamma_\Lambda), \quad \Psi := g * (\eta - \gamma_\Lambda). \tag{2.41}$$

Given $\varepsilon > 0$, the following holds true for p large enough (depending on ε) and n large enough (depending on p).

(a)
$$\left| \int_{\Lambda_p} \Psi(s)\mathrm{d}s \right| \leq \varepsilon\, W_1(\eta, \gamma_\Lambda)$$

(b)
$$\left| \int_{\Lambda_n \setminus \Lambda_p} \Psi_n(s)\mathrm{d}s \right| \leq \varepsilon \Big(W_1(\eta, \gamma_\Lambda) + |\gamma_\Lambda| \Big).$$

(c) For any $x \in \Lambda_n \setminus \Lambda_p$,
$$|\Psi_n(x)| \leq \frac{W_1(\eta, \gamma_\Lambda)}{\mathrm{dist}(x, \Lambda)} \qquad \text{and} \qquad |\Psi_n'(x)| \leq \frac{8W_1(\eta, \gamma_\Lambda)}{\mathrm{dist}(x, \Lambda)^2}.$$

(d) For any $x \in \mathbb{R} \setminus \Lambda_p$,
$$|\Psi(x)| \leq \frac{W_1(\eta, \gamma_\Lambda)}{\mathrm{dist}(x, \Lambda)} \qquad \text{and} \qquad |\Psi'(x)| \leq \frac{8W_1(\eta, \gamma_\Lambda)}{\mathrm{dist}(x, \Lambda)^2}.$$

Proof. Let us enumerate the configurations as $\gamma_\Lambda = \sum_{i=1}^M \delta_{\gamma_i}$ and $\eta = \sum_{i=1}^M \delta_{\eta_i}$. To prove (a), we start by writing

$$\int_{\Lambda_p} \Psi(s)\mathrm{d}s = \sum_{i=1}^M \int_{\Lambda_p} \Big(\log|\gamma_i - s| - \log|\eta_i - s| \Big)\mathrm{d}s.$$

Now, set

$$\mathsf{V}(t) := \int_{-1}^1 -\log|t - s|\mathrm{d}s = (1 + t)\log(1 + t) + (1 - t)\log(1 - t).$$

and let $k > 0$ be fixed so that $\Lambda \subset \Lambda_k$. We obtain, by a linear change of variables sending $[-1, 1]$ on $\Lambda_p = [-\frac{p}{2}, \frac{p}{2}]$:

$$\left| \int_{\Lambda_p} \Psi(s)\mathrm{d}s \right| = \left| \frac{p}{2} \sum_{i=1}^M \Big(\mathsf{V}\Big(\frac{2\eta_i}{p}\Big) - \mathsf{V}\Big(\frac{2\gamma_i}{p}\Big) \Big) \right| \leq \sup_{[-\frac{k}{p}, \frac{k}{p}]} |\mathsf{V}'| \sum_{i=1}^M |\eta_i - \gamma_i|.$$

23

Since V' is continuous near the origin, since $\mathsf{V}'(0) = 0$ and since the enumeration of η and γ_Λ is arbitrary, (a) follows by taking p large enough, depending on Λ and ε.

We now turn to (b). Since $g_n * \mathsf{Leb}_{\Lambda_n} = 0$ on Λ_n, see e.g. [Borodin and Serfaty, 2013, Equation (2.49)], we have

$$\int \Psi_n \, \mathsf{Leb}_{\Lambda_n \setminus \Lambda_p} = - \int \Psi_n \, \mathsf{Leb}_{\Lambda_p}.$$

For any fixed p, we have for n large enough and for any $x \neq y \in \Lambda_p$,

$$\left| g_n(x-y) + \log\left(\frac{2\pi}{n}\right) - g(x-y) \right| = \left| \log \left| \frac{2 \sin \frac{\pi(x-y)}{n}}{\frac{2\pi(x-y)}{n}} \right| \right| \leq \frac{1}{p^2}. \tag{2.42}$$

We may thus write

$$\left| \int \Psi_n \, \mathsf{Leb}_{\Lambda_n \setminus \Lambda_p} \right| = \left| \int \Psi_n \, \mathsf{Leb}_{\Lambda_p} \right| \leq \left| \int \Psi \, \mathsf{Leb}_{\Lambda_p} \right| + \frac{|\gamma_\Lambda| p}{p^2},$$

where we have used the fact that η and γ_Λ have the same number of points, hence the contribution of the constant term $\log\left(\frac{2\pi}{n}\right)$ in (2.42) vanishes. Using point (a) of the present lemma, we obtain

$$\left| \int \Psi_n \, \mathsf{Leb}_{\Lambda_n \setminus \Lambda_p} \right| \leq \varepsilon \left(W_1(\eta, \gamma_\Lambda) + |\gamma_\Lambda| \right),$$

for p large enough (depending on Λ, ε) and n large enough (depending on p).

Finally, we prove (c) and (d). For any $x \in \Lambda_n \setminus \Lambda_p$, by applying the mean value theorem to $g_n(x - \cdot)$ between η_i and γ_i, we obtain

$$|g_n(x - \eta_i) - g_n(x - \gamma_i)| \leq |\eta_i - \gamma_i| \frac{\pi}{n} \frac{1}{|\tan\left(\frac{\pi}{n} \operatorname{dist}(x, \Lambda)\right)|} \leq \frac{|\eta_i - \gamma_i|}{\operatorname{dist}(x, \Lambda)},$$

and the first inequality of (c) is obtained by summing over $i \in \{1, \dots, M\}$. We obtain the second inequality of (c), as well as (d), by the same argument but using g_n', g and g' instead of g_n respectively.

$\square$

2.5.2   Proof of Lemma 2.12

We want to show that in the finite periodic model, with high probability, it is possible to move the points in Λ at a small cost.

Proof of Lemma 2.12. In view of (2.41), and in order to obtain the controls on the move functions from discrepancy estimates, it is convenient to work with "move functions with background" defined by

$$\widetilde{\mathsf{M}}_{\Lambda, \Lambda_p}^{n-\mathsf{per}}(\gamma, \eta) := \int_{\Lambda_p \setminus \Lambda} \Psi_n \, \mathrm{d}(\gamma - \mathsf{Leb}). \tag{2.43}$$

They are related to the usual move functions defined in (2.25) as follows:

$$\mathsf{M}_{\Lambda, \Lambda_p}^{n-\mathsf{per}}(\gamma, \eta) = \widetilde{\mathsf{M}}_{\Lambda, \Lambda_p}^{n-\mathsf{per}}(\gamma, \eta) + \int_{\Lambda_p \setminus \Lambda} \Psi_n \, \mathrm{d}\mathsf{Leb}. \tag{2.44}$$

Given any $\gamma \in \mathsf{Conf}(\mathbb{R})$, let us set for convenience,

$$\mathsf{E} := \sup_{\substack{\eta \in \mathsf{Conf}(\Lambda) \\ |\eta| = |\gamma_\Lambda|}} \left| \mathsf{M}_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\gamma, \eta) - \mathsf{M}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma, \eta) \right|,$$

$$\widetilde{\mathsf{E}} := \sup_{\substack{\eta \in \mathsf{Conf}(\Lambda) \\ |\eta| = |\gamma_\Lambda|}} \left| \widetilde{\mathsf{M}}_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\gamma, \eta) - \widetilde{\mathsf{M}}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma, \eta) \right|.$$

Using (2.44) we see that

$$\mathsf{E} \leq \widetilde{\mathsf{E}} + \sup_{\substack{\eta \in \mathsf{Conf}(\Lambda) \\ |\eta| = |\gamma_\Lambda|}} \left| \int_{\Lambda_n \setminus \Lambda_p} \Psi_n \, \mathrm{d}\mathsf{Leb} \right|.$$

Using Lemma 2.26(b) and the fact that $W_1(\eta, \gamma_\Lambda) \leq |\Lambda| \, |\gamma_\Lambda|$, given any $\alpha > 0$, we have, if p is large enough (depending on Λ, α) and n is large enough (depending on p)

$$\mathsf{E} \leq \widetilde{\mathsf{E}} + \alpha |\gamma_\Lambda|.$$

Next, for any $\delta, L > 0$, we have

$$Q_{n,\beta}\Big(\mathsf{Conf}(\mathbb{R}) \setminus \mathsf{A}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\delta) \Big) = Q_{n,\beta}(\mathsf{E} > \delta)$$
$$\leq Q_{n,\beta}(\mathsf{E} > \delta, \ |\gamma_\Lambda| \leq L) + Q_{n,\beta}(|\gamma_\Lambda| > L)$$
$$\leq Q_{n,\beta}(\widetilde{\mathsf{E}} > \delta - \alpha L, \ |\gamma_\Lambda| \leq L) + Q_{n,\beta}(|\gamma_\Lambda| > L)$$
$$\leq \frac{1}{\delta - \alpha L} \mathbb{E}_{Q_{n,\beta}} [\widetilde{\mathsf{E}} \, \mathbf{1}_{|\gamma_\Lambda| \leq L}] + \frac{1}{L^2} \mathbb{E}_{Q_{n,\beta}} [|\gamma_\Lambda|^2].$$

To prove the lemma, it is enough to show that, given any $L > 0$,

$$\mathbb{E}_{Q_{n,\beta}} [\widetilde{\mathsf{E}} \, \mathbf{1}_{|\gamma_\Lambda| \leq L}]$$

can be made arbitrarily small by taking first p, then n, large enough. Indeed, given any $\delta > 0$, by taking p, n, L large enough and $\alpha := \delta/2L$ the lemma would follow from Lemma 2.24.

To prove this claim, we split $\Lambda_n \setminus \Lambda_p$ into the subintervals

$$I_j := \begin{cases} \left(\frac{j}{2}, \frac{j+1}{2} \right] & \text{if } j > 0 \\ \left[\frac{j}{2}, \frac{j+1}{2} \right) & \text{if } j < 0, \end{cases}$$

so as to write

$$\widetilde{\mathsf{M}}_{\Lambda,\Lambda_n}^{n-\mathsf{per}}(\gamma, \eta) - \widetilde{\mathsf{M}}_{\Lambda,\Lambda_p}^{n-\mathsf{per}}(\gamma, \eta) = \left(\sum_{j=-n}^{-p-1} + \sum_{j=p}^{n-1} \right) \int_{I_j} \Psi_n \, \mathrm{d}(\gamma - \mathsf{Leb}). \qquad (2.45)$$

By applying the mean value theorem to Ψ_n and Lemma 2.26(c), we obtain for any $x \in I_j$,

$$\left| \Psi_n(x) - \Psi_n\left(\frac{j}{2} \right) \right| \leq \frac{8 W_1(\eta, \gamma_\Lambda)}{\mathrm{dist}(x, \Lambda)^2},$$

and thus, for any $-n \le j < n$,

$$\left| \int_{I_j} \Psi_n \, \mathrm{d}(\gamma - \mathsf{Leb}) - \Psi_n\left(\frac{j}{2}\right) \mathrm{Discr}_{I_j}(\gamma) \right| \le \frac{8 W_1(\eta, \gamma_\Lambda)}{\mathrm{dist}(I_j, \Lambda)^2} \left(\mathrm{Discr}_{I_j}(\gamma) + 1 \right).$$

Since p is arbitrarily large and Λ is fixed, there exists $c > 0$ such that, for any $|j| \ge p$,

$$\mathrm{dist}(I_j, \Lambda) \ge \frac{j}{c}.$$

Combined with (2.45), we obtain

$$\left| \widetilde{\mathsf{M}}_{\Lambda, \Lambda_n}^{n-\mathsf{per}}(\gamma, \eta) - \widetilde{\mathsf{M}}_{\Lambda, \Lambda_p}^{n-\mathsf{per}}(\gamma, \eta) - \left(\sum_{j=-n}^{-p-1} + \sum_{j=p}^{n-1} \right) \Psi_n\left(\frac{j}{2}\right) \mathrm{Discr}_{I_j}(\gamma) \right|$$

$$\le 8 c^2 W_1(\eta, \gamma_\Lambda) \left(\sum_{j=-n}^{-p-1} + \sum_{j=p}^{n-1} \right) \frac{\mathrm{Discr}_{I_j}(\gamma) + 1}{j^2}. \quad (2.46)$$

By performing a summation by parts, we can write

$$\sum_{j=p}^{n-1} \Psi_n\left(\frac{j}{2}\right) \mathrm{Discr}_{I_j}(\gamma) = \Psi_n\left(\frac{n}{2}\right) \mathrm{Discr}_{(0, \frac{n}{2}]}(\gamma) - \Psi_n\left(\frac{p}{2}\right) \mathrm{Discr}_{(0, \frac{p}{2}]}(\gamma)$$

$$+ \sum_{j=p}^{n-1} \left(\Psi_n\left(\frac{j}{2}\right) - \Psi_n\left(\frac{j+1}{2}\right) \right) \mathrm{Discr}_{(0, \frac{j+1}{2}]}(\gamma). \quad (2.47)$$

Using again the mean value theorem and Lemma 2.26(c), we have

$$\left| \sum_{j=p}^{n-1} \left(\Psi_n\left(\frac{j}{2}\right) - \Psi_n\left(\frac{j+1}{2}\right) \right) \mathrm{Discr}_{(0, \frac{j+1}{2}]}(\gamma) \right| \le 8 c^2 W_1(\eta, \gamma_\Lambda) \sum_{j=p}^{n-1} \frac{\left| \mathrm{Discr}_{(0, \frac{j+1}{2}]}(\gamma) \right|}{j^2}$$

$$\quad (2.48)$$

and similar estimates holds for the sum where j ranges from $-n$ to $-p-1$. Moreover, since $x / \mathrm{dist}(x, \Lambda)$ is bounded when $x \notin 2\Lambda$, it follows from Lemma 2.26(c) that there exists $\kappa > 0$ independent on η, γ such that

$$\sup_{x \notin 2\Lambda} |x \Psi_n(x)| \le \kappa W_1(\eta, \gamma_\Lambda). \quad (2.49)$$

As a consequence, we obtain from (2.46)–(2.49),

$$\left| \widetilde{\mathsf{M}}_{\Lambda, \Lambda_n}^{n-\mathsf{per}}(\gamma, \eta) - \widetilde{\mathsf{M}}_{\Lambda, \Lambda_p}^{n-\mathsf{per}}(\gamma, \eta) \right| \le 2\kappa W_1(\eta, \gamma_\Lambda) \left(\frac{\left| \mathrm{Discr}_{(0, \frac{n}{2}]}(\gamma) \right|}{n} + \frac{\left| \mathrm{Discr}_{(0, \frac{p}{2}]}(\gamma) \right|}{p} \right)$$

$$+ 2\kappa W_1(\eta, \gamma_\Lambda) \left(\frac{\left| \mathrm{Discr}_{(-\frac{n}{2}, 0]}(\gamma) \right|}{n} + \frac{\left| \mathrm{Discr}_{(-\frac{p}{2}, 0]}(\gamma) \right|}{p} \right)$$

$$+ 8 c^2 W_1(\eta, \gamma_\Lambda) \left(\sum_{j=p}^{n-1} \frac{\left| \mathrm{Discr}_{(0, \frac{j+1}{2}]}(\gamma) \right|}{j^2} + \sum_{j=-n}^{-p-1} \frac{\left| \mathrm{Discr}_{(\frac{j+1}{2}, 0]}(\gamma) \right|}{j^2} \right)$$

$$+ 8 c^2 W_1(\eta, \gamma_\Lambda) \left(\sum_{j=-n}^{-p-1} + \sum_{j=p}^{n-1} \right) \frac{\mathrm{Discr}_{I_j}(\gamma) + 1}{j^2}$$

$$=: W_1(\eta, \gamma_\Lambda) \, \mathrm{Err}_{n,p}(\gamma), \quad (2.50)$$

Given any $L > 0$, we obtain from (2.50) and the upper bound $W_1(\eta, \gamma_\Lambda) \leq |\gamma_\Lambda| |\Lambda|$ that

$$\mathbb{E}_{Q_{n,\beta}}[\widetilde{\mathsf{E}}\, \mathbf{1}_{|\gamma_\Lambda| \leq L}] \leq L |\Lambda|\, \mathbb{E}_{Q_{n,\beta}}[\mathrm{Err}_{n,p}(\gamma)]. \tag{2.51}$$

Finally, we use Cauchy-Schwarz inequality and Lemma 2.24, to obtain

$$\mathbb{E}_{Q_{n,\beta}}\left(\frac{|\mathrm{Discr}_{(0,\frac{n}{2}]}(\gamma)|}{n}\right) \leq \sqrt{\frac{C_\beta}{2n}}.$$

and

$$\mathbb{E}_{Q_{n,\beta}}\left(\sum_{j=p}^{n-1} \frac{\left|\mathrm{Discr}_{(0,\frac{j+1}{2}]}(\gamma)\right|}{j^2}\right) \leq \sqrt{\frac{C_\beta}{2}} \sum_{j=p}^{\infty} \frac{\sqrt{j+1}}{j^2},$$

and this yields that $\mathbb{E}_{Q_{n,\beta}}[\mathrm{Err}_{n,p}(\gamma)]$ can be made arbitrarily small when $p \leq n$ are large enough. The lemma follows from (2.51).

$\square$

2.5.3 Proof of Lemma 2.15 and Lemma 2.18

Proof. Let P be a stationary point process such that $\mathbb{E}_P[\mathbb{W}(\gamma)] < \infty$. We start by showing that, for P-a.e. $\gamma \in \mathsf{Conf}(\mathbb{R})$ and every $\eta \in \mathsf{Conf}(\Lambda)$ such that $|\eta| = |\gamma_\Lambda|$, the sequence $\{\mathsf{M}_{\Lambda,\Lambda_p}(\eta, \gamma)\}_{p \geq 1}$ is a Cauchy sequence, thus proving Lemma 2.15.

As in the proof of Lemma 2.12, we introduce the move functions with a background,

$$\widetilde{\mathsf{M}}_{\Lambda,\Lambda_p}(\gamma, \eta) := \int_{\Lambda_p \setminus \Lambda} \Psi\, \mathrm{d}(\gamma - \mathsf{Leb}).$$

Since by definition,

$$\mathsf{M}_{\Lambda,\Lambda_p}(\gamma, \eta) = \widetilde{\mathsf{M}}_{\Lambda,\Lambda_p}(\gamma, \eta) + \int \Psi\, \mathsf{Leb}_{\Lambda_p \setminus \Lambda},$$

it follows from Lemma 2.26(a) that it is enough to show that $\{\widetilde{\mathsf{M}}_{\Lambda,\Lambda_p}(\eta, \gamma)\}_{p \geq 1}$ is a Cauchy sequence, uniformly in η. For any $m \geq p$, by following the same steps as in the proof of Lemma 2.12, using estimates on Ψ instead of Ψ_n, we have with $\mathrm{Err}_{m,p}(\gamma)$ defined in (2.50),

$$\left|\widetilde{\mathsf{M}}_{\Lambda,\Lambda_m}(\gamma, \eta) - \widetilde{\mathsf{M}}_{\Lambda,\Lambda_p}(\gamma, \eta)\right| \leq W_1(\eta, \gamma)\, \mathrm{Err}_{m,p}(\gamma) \leq |\gamma_\Lambda| |\Lambda|\, \mathrm{Err}_{m,p}(\gamma). \tag{2.52}$$

Now, using Lemma 2.23, we obtain

$$\mathbb{E}_P\left(\sum_{j=p}^{\infty} \frac{\left|\mathrm{Discr}_{(0,\frac{j+1}{2}]}(\gamma)\right|}{j^2}\right) \leq \sqrt{C(C + \mathbb{E}_P(\mathbb{W}))} \sum_{j=p}^{\infty} \frac{\sqrt{j+1}}{j^2} < \infty$$

so that,

$$\sum_{j=p}^{\infty} \frac{\left|\mathrm{Discr}_{(0,\frac{j+1}{2}]}(\gamma)\right|}{j^2} < +\infty \qquad P\text{-a.s}$$

and in particular,

$$\lim_{p \to \infty} \sum_{j=p}^{\infty} \frac{\left|\mathrm{Discr}_{(0,\frac{j+1}{2}]}(\gamma)\right|}{j^2} = 0 \qquad P\text{-a.s.}$$

By similar arguments, we have

$$\lim_{p \to \infty} \sum_{j=p}^{\infty} \frac{|\mathrm{Discr}_{I_j}(\gamma)| + 1}{j^2} = 0 \qquad P\text{-a.s},$$

and the same holds true for the sums involving negative j's. Moreover, since Lemma 2.25 yields that

$$\lim_{m \to \infty} \frac{\left|\mathrm{Discr}_{(0, \frac{m}{2}]}(\gamma)\right|}{m} = 0 \qquad P\text{-a.s},$$

we have obtained that

$$\lim_{p \to \infty} \lim_{m \to \infty} \mathrm{Err}_{m,p}(\gamma) = 0 \qquad P\text{-a.s}$$

and our claim follows from (2.52).

Moreover, the previous estimates show that,

$$\left|\widetilde{\mathsf{M}}_{\Lambda, \mathbb{R}}(\gamma, \eta) - \widetilde{\mathsf{M}}_{\Lambda, \Lambda_p}(\gamma, \eta)\right| \leq |\gamma_\Lambda| |\Lambda| \, \mathrm{Err}_{\infty, p}(\gamma)$$

where

$$\mathrm{Err}_{\infty, p}(\gamma) := 2\kappa \left(\frac{\left|\mathrm{Discr}_{(0, \frac{p}{2}]}(\gamma)\right|}{p} + \frac{\left|\mathrm{Discr}_{(-\frac{p}{2}, 0]}(\gamma)\right|}{p} \right) \tag{2.53}$$

$$+ 8c^2 \left(\sum_{j=p}^{\infty} \frac{\left|\mathrm{Discr}_{(0, \frac{j+1}{2}]}(\gamma)\right|}{j^2} + \sum_{j=-\infty}^{-p-1} \frac{\left|\mathrm{Discr}_{(\frac{j+1}{2}, 0]}(\gamma)\right|}{j^2} \right)$$

$$+ 8c^2 \left(\sum_{j=-\infty}^{-p-1} + \sum_{j=p}^{\infty} \right) \frac{\mathrm{Discr}_{I_j}(\gamma) + 1}{j^2}$$

$$\tag{2.54}$$

exists P-a.s. Now the proof of (2.38) is exactly the same as in Lemma 2.12 but using the estimates on Ψ instead of Ψ_n. $\qquad \square$

3 Number-rigidity for solutions of canonical DLR equations

In this section we prove that Sine_β is number-rigid in the sense of Ghosh and Peres [2017], that is part (A) of Theorem 1.1. In fact, we prove that any stationary process satisfying (2.18) is number-rigid, which is the main result of this section. We say that a point process is *stationary* when it is invariant under translations of the configurations $\gamma \mapsto \gamma + x := \{x + y : \ y \in \gamma\}$ for any $x \in \mathbb{R}$. Recall also that $f_{\Lambda, \mathbb{R}}$ was introduced in (2.16).

Definition 3.1 (Canonical DLR)**.** Let us fix $\beta > 0$. We say that a stationary point process P on $\mathbb{R}$ satisfies the *canonical DLR equations* if $\mathbb{E}_P[\mathbb{W}(\gamma)] < \infty$ and

$$\mathbb{E}_P(f - f_{\Lambda, \mathbb{R}}) = 0 \qquad \qquad \text{(canonical DLR)}$$

for every bounded Borel set $\Lambda \subset \mathbb{R}$ and every bounded Borel function $f : \mathsf{Conf}(\mathbb{R}) \to \mathbb{R}$.

The assumption that P has finite renormalized energy $\mathbb{E}_P[\mathbb{W}(\gamma)]$ is here to ensure that the move functions and thus $f_{\Lambda,\mathbb{R}}$ are well-defined. Our goal is now to prove:

Theorem 3.2. If P is a stationary point process on $\mathbb{R}$ that satisfies canonical DLR, then P is number-rigid.

Thus, Theorem 1.1(A) follows from Theorem 2.1(C*) and Theorem 3.2. As a consequence of Theorem 1.1(A), Theorem 2.1(C*) upgrades to Theorem 1.1(C) and, since Theorem 1.1(B) has already be proven in the previous section, the proof of our main theorem is complete.

The proof of Theorem 3.2 is based on canonical and grand canonical descriptions of a Gibbs point process via its Campbell measures. These descriptions have been studied intensively in the seventies and eighties, see for instance [Georgii, 1979, Kozlov, 1976, Nguyen and Zessin, 1979, Wakolbinger and Eder, 1984]. The proof goes by contradiction and follows three steps:

Step 1: First, we show in Section 3.2 that any point process satisfying the canonical DLR equations admits a canonical description via its Campbell measures, see Theorem 3.7.

Step 2: Next, if we further assume that the process is ergodic and, for the sake of contradiction, **not** number-rigid, then we show in Section 3.3 that this representation can be extended into a grand canonical version, see Theorem 3.15.

Step 3: Finally, we show that the grand canonical representation yields a contradiction because of the long range of the logarithmic interaction, see Section 3.5. Roughly speaking, we use this representation to move points far away from the origin and show that the configurations obtained should have a much larger weight than it is allowed.

It turns out these three steps can be performed in a much more general setting than the one dimensional logarithmic interaction. This leads to a more general result than Theorem 3.2 that we present in Section 3.5.7, see Theorem 3.18; we focused on the one dimensional log-gas for the sake of the presentation.

Convention: For convenience, if $\mathbf{x}_n = (x_1, \dots, x_n) \in \mathbb{R}^n$, given any $\gamma \in \mathsf{Conf}(\mathbb{R})$ with an abuse of notation we write $\gamma \setminus \mathbf{x}_n$ (resp. $\gamma \cup \mathbf{x}_n$, etc) instead of $\gamma \setminus \{\mathbf{x}_n\}$ (resp. $\gamma \cup \{\mathbf{x}_n\}$, etc). Moreover, given $\gamma \in \mathsf{Conf}(\mathbb{R})$, the sum

$$\sum_{\mathbf{x}_n \subset \gamma}$$

means that we are summing over all ordered n-tuples of points from the configuration γ, namely

$$\sum_{\mathbf{x}_n \subset \gamma} f(\mathbf{x}_n) = \sum_{\substack{x_1, \dots, x_n \in \gamma \\ x_i \neq x_j \text{ for } i \neq j}} f(x_1, \dots, x_n).$$

The reader should have in mind that in the following the configuration γ will be simple since it comes from a point process.

3.1 A first consequence of the canonical DLR equations

The following corollary will be useful in the sequel.

Corollary 3.3. Let P satisfy be a stationary point process on $\mathbb{R}$ that satisfies canonical DLR. Then, for any disjoint bounded Borel sets $B_1, \ldots, B_k \subset \mathbb{R}$ with positive Lebesgue measure and any integers $n_1, \ldots, n_k$, we have

$$P(|\gamma_{B_1}| = n_1, \ldots, |\gamma_{B_k}| = n_k) > 0.$$

In particular, Corollary 1.3 now follows from Theorem 2.1.

Proof. Set $n := \sum_j n_j$. Let $\Lambda \subset \mathbb{R}$ be a bounded Borel set so that $\Lambda \setminus (B_1 \cup \ldots \cup B_k)$ has positive Lebesgue measure and $P(|\gamma_\Lambda| \geq n) > 0$. The latter is ensured as soon as Λ is large enough since P is stationary and in particular $\mathbb{E}_P|\gamma| = +\infty$. If we set $f(\gamma) := \prod_{j=1}^{k} \mathbf{1}_{|\gamma_{B_j}| = n_j}$, then by applying canonical DLR on Λ we have

$$P(|\gamma_{B_1}| = n_1, \ldots, |\gamma_{B_k}| = n_k) = \int \prod_{j=1}^{k} f(\{x_1, \ldots, x_{|\gamma_\Lambda|}\}) \rho_{\Lambda^c}(x_1, \ldots, x_{|\gamma_\Lambda|}) \prod_{i=1}^{|\gamma_\Lambda|} \mathrm{d}x_i \, P(\mathrm{d}\gamma).$$

Now, if $P(|\gamma_{B_1}| = n_1, \ldots, |\gamma_{B_k}| = n_k) = 0$, then for P-a.e. γ,

$$\int \prod_{j=1}^{k} f(\{x_1, \ldots, x_{|\gamma_\Lambda|}\}) \rho_{\Lambda^c}(x_1, \ldots, x_{|\gamma_\Lambda|}) \prod_{i=1}^{|\gamma_\Lambda|} \mathrm{d}x_i = 0,$$

but by the definition of ρ_{Λ^c}, see (2.3), this yields that $|\gamma_\Lambda| < n$ for P-a.e γ, which is not possible. □

3.2 Campbell measures

In this section, we introduce the Campbell measures and prove that the canonical DLR equations yield a result on the representation for these measures.

Definition 3.4 (Campbell measure)**.** The *Campbell measure* of order $n \geq 1$ of a point process P on $\mathbb{R}$ is the measure $\mathsf{C}_P^{(n)}$ on $\mathbb{R}^n \times \mathsf{Conf}(\mathbb{R})$ defined by, for any positive Borel test function f,

$$\mathsf{C}_P^{(n)}(f) := \int \sum_{\mathbf{x}_n \subset \gamma} f(\mathbf{x}_n, \gamma \setminus \mathbf{x}_n) \, P(\mathrm{d}\gamma). \tag{3.1}$$

The Campbell measure encodes the joint distribution of n typical points $\mathbf{x}_n$ and their neighborhood $\gamma \setminus \mathbf{x}_n$ for a random configuration γ with law P. For example, in the case of a Poisson point process Π of intensity measure ρ, the Campbell measure is given by

$$\mathsf{C}_\Pi^{(n)} = \rho^{\otimes n} \otimes \Pi,$$

which is known as the Slivnyak-Mecke Theorem, see e.g. [Moller and Waagepetersen, 2003, Section 3.2].

In the following, we need to introduce the following *cost function.*

Definition 3.5 (Cost of moving n points from 0 in γ)**.** Let $\gamma \in \mathsf{Conf}(\mathbb{R})$ and take $\mathbf{x}_n \subset \gamma$. We consider the cost of moving of the n-tuple $(0, \ldots, 0)$ to $\mathbf{x}_n = (x_1, \ldots, x_n)$ defined by

$$\mathsf{Cost}(\mathbf{x}_n, \gamma) = \sum_{i<j} g(x_i - x_j) + \lim_{p \to \infty} \int_{-p}^{p} \sum_{i=1}^{n} (g(x_i - y) - g(y)) \gamma(\mathrm{d}y) \tag{3.2}$$

provided the limit exists. Note that $\mathsf{Cost}(\mathbf{x}_n, \gamma)$ does not depend on the ordering of $\mathbf{x}_n$.

Remark 3.6. If P be a stationary process satisfying $\mathbb{E}_P[\mathbb{W}(\gamma)] < +\infty$, then the limit (3.2) exists for P-a.e. γ and any $\mathbf{x}_n$ (the proof is the same to that of Proposition 2.15).

The following theorem is a first description of the structure of Campbell measure for point processes satisfying canonical DLR equations.

Theorem 3.7. Let P be a stationary point process on $\mathbb{R}$ satisfying canonical DLR. Then, for every $n \geq 1$, there exists a Borel measure Q_n on $\mathsf{Conf}(\mathbb{R})$ such that $\mathsf{C}_P^{(n)}$ is absolutely continuous with respect to $\mathsf{Leb}^{\otimes n} \otimes \mathsf{Q}_n$ and has density

$$\frac{\mathrm{d}\mathsf{C}_P^{(n)}}{\mathrm{d}\mathsf{Leb}^{\otimes n} \otimes \mathsf{Q}_n}(\mathbf{x}_n, \gamma) = \mathrm{e}^{-\beta \mathsf{Cost}(\mathbf{x}_n, \gamma)}. \tag{3.3}$$

Proof. Consider the tilted measure,

$$\widetilde{\mathsf{C}}_P^{(n)} := \mathrm{e}^{\beta \mathsf{Cost}} \, \mathsf{C}_P^{(n)}. \tag{3.4}$$

It is thus enough to show that $\widetilde{\mathsf{C}}_P^{(n)} = \mathsf{Leb}^{\otimes n} \otimes \mathsf{Q}_n$ for some measure Q_n on $\mathsf{Conf}(\mathbb{R})$.

Let $f : \mathbb{R}^n \times \mathsf{Conf}(\mathbb{R}) \to [0, \infty)$ be a measurable function satisfying $f(\mathbf{x}, \gamma) = 0$ as soon as $\mathbf{x} \notin K$ for some compact set $K \subset \mathbb{R}^d$. Let $\Lambda \subset \mathbb{R}$ be a bounded Borel set such that $K \subset \Lambda^n$. Using the definition (3.1), canonical DLR, and setting

$$F_\Lambda(\mathbf{x}_n, \eta, \gamma) := \mathsf{H}_\Lambda(\eta) + \mathsf{M}_{\Lambda, \mathbb{R}}(\eta, \gamma) - \mathsf{Cost}(\mathbf{x}_n, \eta \cup \gamma_{\Lambda^c} \setminus \mathbf{x}_n),$$

we find

$$\widetilde{\mathsf{C}}_P^{(n)}(f) = \int \sum_{\mathbf{x}_n \subset \gamma_\Lambda} \mathrm{e}^{\beta \mathsf{Cost}(\mathbf{x}_n, \gamma \setminus \mathbf{x}_n)} f(\mathbf{x}_n, \gamma \setminus \mathbf{x}_n) P(\mathrm{d}\gamma)$$

$$= \int \sum_{\mathbf{x}_n \subset \eta} \mathrm{e}^{\beta \mathsf{Cost}(\mathbf{x}_n, \eta \cup \gamma_{\Lambda^c} \setminus \mathbf{x}_n)} f(\mathbf{x}_n, (\eta \setminus \mathbf{x}_n) \cup \gamma_{\Lambda^c})$$

$$\times \frac{1}{Z_{\Lambda, \mathbb{R}}(\gamma)} \mathrm{e}^{-\beta(\mathsf{H}_\Lambda(\eta) + \mathsf{M}_{\Lambda, \mathbb{R}}(\eta, \gamma))} \mathbf{B}_{|\gamma_\Lambda|, \Lambda}(\mathrm{d}\eta) P(\mathrm{d}\gamma)$$

$$= \int \sum_{\mathbf{x}_n \subset \eta} f(\mathbf{x}_n, (\eta \setminus \mathbf{x}_n) \cup \gamma_{\Lambda^c}) \frac{\mathrm{e}^{-\beta F_\Lambda(\mathbf{x}_n, \eta, \gamma)}}{Z_{\Lambda, \mathbb{R}}(\gamma)} \mathbf{B}_{|\gamma_\Lambda|, \Lambda}(\mathrm{d}\eta) P(\mathrm{d}\gamma)$$

$$= \int \mathbf{B}_{n, \Lambda}(\mathrm{d}\mathbf{x}_n) \int \frac{\mathbf{1}_{|\gamma_\Lambda| \geq n} |\gamma_\Lambda|!}{(|\gamma_\Lambda| - n)!} f(\mathbf{x}_n, \tilde{\eta} \cup \gamma_{\Lambda^c}) \frac{\mathrm{e}^{-\beta F_\Lambda(\mathbf{x}_n, \tilde{\eta} \cup \mathbf{x}_n, \gamma)}}{Z_{\Lambda, \mathbb{R}}(\gamma)} \mathbf{B}_{|\gamma_\Lambda| - n, \Lambda}(\mathrm{d}\tilde{\eta}) P(\mathrm{d}\gamma)$$

where we made the change of variables $\tilde{\eta} := \eta \setminus \mathbf{x}_n$. Since one can check by direct computation that the map

$$\mathbf{x}_n \mapsto F_\Lambda(\mathbf{x}_n, \tilde{\eta} \cup \mathbf{x}_n, \gamma) =: F_\Lambda(\tilde{\eta}, \gamma),$$

31

is constant, we obtain the factorization

$$\widetilde{\mathsf{C}}_P^{(n)}(f) = \int f(\mathbf{x}_n, \zeta) \, \mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{x}_n) \, \mathsf{Q}_n^\Lambda(\mathrm{d}\zeta) \tag{3.5}$$

where the measure Q_n^Λ is defined for any measurable map $g : \mathsf{Conf}(\mathbb{R}) \to [0, \infty)$ by,

$$\int g(\zeta) \mathsf{Q}_n^\Lambda(\mathrm{d}\zeta) := \int \frac{\mathbf{1}_{|\gamma_\Lambda| \geq n} \, |\gamma_\Lambda|!}{|\Lambda|^n (|\gamma_\Lambda| - n)!} g(\tilde{\eta} \cup \gamma_{\Lambda^c}) \, \frac{\mathrm{e}^{-\beta F_\Lambda(\tilde{\eta}, \gamma)}}{Z_{\Lambda, \mathbb{R}}(\gamma)} \, \mathbf{B}_{|\gamma_\Lambda| - n, \Lambda}(\mathrm{d}\tilde{\eta}) P(\mathrm{d}\gamma).$$

Finally, since (3.5) holds true after replacing Λ by any bounded Borel set $\Lambda' \supset \Lambda$, we obtain by taking $f := \mathbf{1}_{\Lambda^n} \otimes g$,

$$\int g(\zeta) \mathsf{Q}_n^\Lambda(\mathrm{d}\zeta) = \int g(\zeta) \mathsf{Q}_n^{\Lambda'}(\mathrm{d}\zeta)$$

for any measurable map $g : \mathsf{Conf}(\mathbb{R}) \to [0, \infty)$, and thus $\mathsf{Q}_n := \mathsf{Q}_n^\Lambda$ does not depend on Λ. $\qquad \square$

3.3 Rigidity and ergodicity

We now provide a convenient characterization of number-rigidity. First, in the next lemma we show that, when one wants to prove number-rigidity, it is enough to restrict to an increasing countable family instead of all bounded measurable subsets $\Lambda \subset \mathbb{R}$.

Lemma 3.8. Given any countable family $(B_m)_{m \geq 1}$ of bounded Borel subsets of $\mathbb{R}$ satisfying $\cup_m B_m = \mathbb{R}$, P is number-rigid if and only if, for any $m \geq 1$, there exists a measurable function $\mathsf{Number}_{B_m} : \mathsf{Conf}(B_m^c) \to \mathbb{N}$ such that $|\gamma_{B_m}| = \mathsf{Number}_{B_m}(\gamma_{B_m^c})$ P-a.s.

Proof. Let $\Lambda \subset \mathbb{R}$ be measurable and bounded and $m \geq 1$ such that $\Lambda \subset B_m$. If one assumes that $|\gamma_{B_m}|$ is a measurable function of $\gamma_{B_m^c}$, then by writing $|\gamma_\Lambda| = |\gamma_{B_m}| - |\gamma_{B_m \setminus \Lambda}|$ we see that $|\gamma_\Lambda|$ coincides almost surely with a measurable function of γ_{Λ^c}. $\qquad \square$

Next, we need a few definitions.

Definition 3.9. For any bounded Borel set Λ and for a fixed γ, we introduce the event

$$\mathsf{Ext}_\Lambda^\gamma := \{\eta \in \mathsf{Conf}(\mathbb{R}) : \eta_{\Lambda^c} = \gamma_{\Lambda^c}\}$$

of having the same exterior configuration than γ outside of Λ. We also need the event

$$\mathsf{HasPoints}_{k, \Lambda} := \{\eta \in \mathsf{Conf}(\mathbb{R}) : |\eta_\Lambda| = k\}$$

of having $k \geq 0$ points in Λ and, given a point process P on $\mathbb{R}$ and $n \geq 1$,

$$\mathsf{CanRmv}_{n, \Lambda}^P := \Big\{ \gamma \in \mathsf{Conf}(\mathbb{R}) : \exists k \in \mathbb{N},$$
$$P\left(\mathsf{HasPoints}_{k, \Lambda} \middle| \mathsf{Ext}_\Lambda^\gamma\right) > 0 \ \text{ and } \ P\left(\mathsf{HasPoints}_{k+n, \Lambda} \middle| \mathsf{Ext}_\Lambda^\gamma\right) > 0 \Big\}. \tag{3.6}$$

The latter can be informally understood as the set of configurations γ for which, with positive probability under P, it is possible to generate two configurations with a number of points in Λ which differs by n, conditionally on having the same exterior configuration

given by γ_{Λ^c}. Formally, since we may have $P(\mathsf{Ext}_\Lambda^\gamma) = 0$, one has to proceed more carefully. Let us define $\mathscr{F}_{\Lambda^c}$ as the σ-algebra generated by the random variable $\gamma \mapsto \gamma_{\Lambda^c}$. Consider the conditional expectation $\mathbb{E}_P[\mathbf{1}_{\mathsf{HasPoints}_{k,\Lambda}}|\mathscr{F}_{\Lambda^c}]$, which is a $\mathscr{F}_{\Lambda^c}$-measurable random variable, and thus can be seen as a measurable function from $\mathsf{Conf}(\Lambda^c)$ to $[0,1]$. We set

$$P\left(\mathsf{HasPoints}_{k,\Lambda}\Big|\mathsf{Ext}_\Lambda^\gamma\right) := \mathbb{E}_P[\mathbf{1}_{\mathsf{HasPoints}_{k,\Lambda}}|\mathscr{F}_{\Lambda^c}](\gamma_{\Lambda^c}).$$

Finally, recalling that $\Lambda_m := [-\frac{m}{2}, \frac{m}{2}]$, we consider the events

$$\mathsf{CanRmv}_n^P := \bigcup_{m \geq 1} \mathsf{CanRmv}_{n,\Lambda_m}^P$$

and the set of *non-rigid configurations*

$$\mathsf{CanRmv}_*^P := \bigcup_{n \geq 1} \mathsf{CanRmv}_n^P. \tag{3.7}$$

Proposition 3.10. A point process P on $\mathbb{R}$ is number-rigid if and only if $P(\mathsf{CanRmv}_*^P) = 0$.

Proof. First, let us assume that P is number-rigid, namely that for any bounded Borel set $\Lambda \subset \mathbb{R}$, there exists a measurable function $\mathsf{Number}_\Lambda : \mathsf{Conf}(\Lambda^c) \to \mathbb{N}$ such that

$$|\gamma_\Lambda| = \mathsf{Number}_\Lambda(\gamma_{\Lambda^c}) \quad \text{for } P\text{-a.e. } \gamma.$$

Then for any $m \geq 1$, for P-a.e. γ and $k \neq \mathsf{Number}_{\Lambda_m}(\gamma_{\Lambda_m^c})$,

$$P\left(\mathsf{HasPoints}_{k,\Lambda_m}\big|\mathsf{Ext}_{\Lambda_m}^\gamma\right) = 0,$$

and $P(\mathsf{CanRmv}_{n,\Lambda_m}^P) = 0$ for any $n \geq 1$, from which $P(\mathsf{CanRmv}_*^P) = 0$ follows.

Conversely, let us assume that P is *not* number-rigid. Then by Lemma 3.8 there exists an integer $m \geq 1$, two integers $0 \leq k_1 < k_2$, and an event E such that $P(E) > 0$ and, for any $\gamma \in E$,

$$P\left(\mathsf{HasPoints}_{k_1,\Lambda_m}\big|\mathsf{Ext}_{\Lambda_m}^\gamma\right) > 0, \qquad P\left(\mathsf{HasPoints}_{k_2,\Lambda_m}\big|\mathsf{Ext}_{\Lambda_m}^\gamma\right) > 0.$$

Thus, if $n := k_2 - k_1$, we have $P(\mathsf{CanRmv}_{n,\Lambda_m}^P) > 0$ and hence $P(\mathsf{CanRmv}_*^P) > 0$. $\qquad\square$

We next make use in a crucial way of the notion of ergodic point processes (which are in this work always assumed to be stationary). An event $E \subset \mathsf{Conf}(\mathbb{R})$ is P-a.s. translation-invariant, if for P-a.e. $\gamma \in \mathsf{Conf}(\mathbb{R})$ and every $x \in \mathbb{R}$, $\mathbf{1}_E(\gamma) = \mathbf{1}_E(\gamma + x)$.

Definition 3.11 (Ergodic point process). We say that a point process P on $\mathbb{R}$ is *ergodic* if it is stationary and, for any P-a.s. translation-invariant event E, we have $P(E) \in \{0, 1\}$.

Proposition 3.10 has the following consequence for ergodic point processes.

Corollary 3.12. Let P be an ergodic point process on $\mathbb{R}$ which is *not* number-rigid. Then there exists $n \geq 1$ such that $P(\mathsf{CanRmv}_n^P) = 1$ and furthermore

$$\lim_{m \to \infty} P(\mathsf{CanRmv}_{n,\Lambda_m}^P) = 1. \tag{3.8}$$

33

Proof. If P is an ergodic point process on $\mathbb{R}$ which is not number-rigid, Proposition 3.10, yields $P(\mathsf{CanRmv}_*^P) > 0$ and therefore there exists $n \geq 1$ such that $P(\mathsf{CanRmv}_n^P) > 0$. Since the event CanRmv_n^P is P-a.s. translation-invariant this implies that $P(\mathsf{CanRmv}_n^P) = 1$. Moreover, (3.8) follows by monotone convergence. $\qquad\square$

The interest we have in ergodic processes comes from the following decomposition.

Proposition 3.13. Let P be a point process on $\mathbb{R}$ which is stationary and satisfies canonical DLR. Let $\mathscr{I}$ be the σ-algebra of the translation-invariant Borel sets of $\mathsf{Conf}(\mathbb{R})$. Then there exists a family of point processes $(P_\eta)_{\eta \in \mathsf{Conf}(\mathbb{R})}$ such that for P-a.e. η, the point process P_η is ergodic, satisfies canonical DLR and

$$P_\eta(\cdot) := \mathbb{E}_P[\mathbf{1}.|\mathscr{I}](\eta). \tag{3.9}$$

In particular we have the standard decomposition of P via its Gibbsian ergodic phases

$$P = \int P_\eta \, P(\mathrm{d}\eta). \tag{3.10}$$

Proof. Since $\mathsf{Conf}(\mathbb{R})$ is a Polish space, there exists a regular conditional probability with respect to $\mathscr{I}$ (see [Dudley, 2002, Theorem 10.2.2]), namely there exists a version of the conditional expectation $A \mapsto \mathbb{E}_P[\mathbf{1}_A|\mathscr{I}]$ such that, for any $\eta \in \mathsf{Conf}(\mathbb{R})$,

$$A \mapsto P_\eta(A) := \mathbb{E}_P[\mathbf{1}_A|\mathscr{I}](\eta) \tag{3.11}$$

defines a probability measure on $\mathsf{Conf}(\mathbb{R})$ and we have

$$P(A) = \mathbb{E}_P\big[\,\mathbb{E}_P[\mathbf{1}_A|\mathscr{I}]\,\big] = \int P_\eta(A) \, P(\mathrm{d}\eta).$$

Note that P_η is a stationary point process for P-a.e. η since P is a stationary point process and by definition of $\mathscr{I}$. Moreover, the quantity $\mathbb{E}_{P_\eta}[\mathbb{W}(\gamma)]$ is necessarily finite for P-a.e. η since otherwise $\mathbb{E}_P[\mathbb{W}(\gamma)]$ would be infinite, and since for any $A \in \mathscr{I}$ we have

$$P_\eta(A) = \mathbb{E}_P[\mathbf{1}_A|\mathscr{I}](\eta) = \mathbf{1}_A(\eta) \in \{0,1\},$$

we see that P_η is ergodic. Moreover, let $\overline{\mathscr{F}_{\Lambda^c}}$ be the σ-algebra generated by the random variables $\gamma \mapsto \gamma_{\Lambda^c}$ and $\gamma \mapsto |\gamma_\Lambda|$, and let $\overline{\mathscr{F}_\infty} := \cap_{m \geq 1} \overline{\mathscr{F}_{\Lambda_m^c}}$. In particular, $\mathscr{I} \subset \overline{\mathscr{F}_\infty}$.

Since P satisfies canonical DLR, [Preston, 1976, Theorem 2.2] states that there exists a version of $A \mapsto \mathbb{E}_P[\mathbf{1}_A|\overline{\mathscr{F}_\infty}]$, such that, for any $\xi \in \mathsf{Conf}(\mathbb{R})$, $\mathbb{E}_P[\mathbf{1}.|\overline{\mathscr{F}_\infty}](\xi)$ is a point process on $\mathbb{R}$ that satisfies canonical DLR. By writing

$$P_\eta(A) = \mathbb{E}_P[\,\mathbb{E}_P[\,\mathbf{1}_A|\overline{\mathscr{F}_\infty}\,]|\mathscr{I}\,](\eta) = \int \mathbb{E}_P[\,\mathbf{1}_A|\overline{\mathscr{F}_\infty}\,](\xi) \, P_\eta(\mathrm{d}\xi)$$

we see that P_η satisfies canonical DLR, and thus P can be written as a mixture of ergodic probability measures satisfying canonical DLR. $\qquad\square$

Combined with the previous proposition, the next result will allow us to restrict to ergodic processes to prove Theorem 3.2.

Proposition 3.14. Let us write $P = \int P_\eta \, P(\mathrm{d}\eta)$ as in Proposition 3.13. If P_η is number-rigid for P-a.e. η, then P is number-rigid.

Proof. In view of Lemma 3.8, it is enough to show that, for all $m \geq 1$, the equality

$$|\gamma_{\Lambda_m}| = \mathbb{E}_P \left[|\gamma_{\Lambda_m}| \, \big| \, \mathscr{F}_{\Lambda_m^c} \right],$$

holds P-almost surely.

By assumption, for P-a.e. η, the process P_η is number-rigid, hence we have

$$|\gamma_{\Lambda_m}| = \mathbb{E}_{P_\eta} \left[|\gamma_{\Lambda_m}| \, \big| \, \mathscr{F}_{\Lambda_m^c} \right], \ \ P_\eta\text{-a.s.},$$

thus it suffices to show

$$\mathbb{E}_P \left[|\gamma_{\Lambda_m}| \big| \mathscr{F}_{\Lambda_m^c} \right] = \mathbb{E}_{P_\eta} \left[|\gamma_{\Lambda_m}| \big| \mathscr{F}_{\Lambda_m^c} \right], \ \ P_\eta\text{-a.s.} \tag{3.12}$$

Both sides of (3.12) are $\mathscr{F}_{\Lambda_m^c}$-measurable random variables. Let f be a bounded, $\mathscr{F}_{\Lambda_m^c}$-measurable random variable. We may write

$$\mathbb{E}_{P_\eta} \left[f \mathbb{E}_P \left[|\gamma_{\Lambda_m}| \big| \mathscr{F}_{\Lambda_m^c} \right] \right] = \mathbb{E}_{P_\eta} \left[\mathbb{E}_P \left[f |\gamma_{\Lambda_m}| \big| \mathscr{F}_{\Lambda_m^c} \right] \right],$$

and by definition of P_η we have,

$$\mathbb{E}_{P_\eta} \left[\mathbb{E}_P \left[f |\gamma_{\Lambda_m}| \big| \mathscr{F}_{\Lambda_m^c} \right] \right] = \mathbb{E}_P \left[\mathbb{E}_P \left[f |\gamma_{\Lambda_m}| \big| \mathscr{F}_{\Lambda_m^c} \right] \big| \mathscr{I} \right] (\eta),$$

where $\mathscr{I}$ is as in Proposition 3.13. Since $\mathscr{I} \subset \mathscr{F}_{\Lambda_m^c}$, we have, by the "tower property" of conditional expectation,

$$\mathbb{E}_P \left[\mathbb{E}_P \left[f |\gamma_{\Lambda_m}| \big| \mathscr{F}_{\Lambda_m^c} \right] \big| \mathscr{I} \right] (\eta) = \mathbb{E}_P \left[f |\gamma_{\Lambda_m}| \big| \mathscr{I} \right] (\eta),$$

so we obtain

$$\mathbb{E}_{P_\eta} \left[f \mathbb{E}_P \left[|\gamma_{\Lambda_m}| \big| \mathscr{F}_{\Lambda_m^c} \right] \right] = \mathbb{E}_P \left[f |\gamma_{\Lambda_m}| \big| \mathscr{I} \right] (\eta), \tag{3.13}$$

but by definition of P_η we have

$$\mathbb{E}_P \left[f |\gamma_{\Lambda_m}| \big| \mathscr{I} \right] (\eta) = \mathbb{E}_{P_\eta} \left[f |\gamma_{\Lambda_m}| \right]. \tag{3.14}$$

Combining (3.13) and (3.14) we see that

$$\mathbb{E}_{P_\eta} \left[f \mathbb{E}_P \left[|\gamma_{\Lambda_m}| \big| \mathscr{F}_{\Lambda_m^c} \right] \right] = \mathbb{E}_{P_\eta} \left[f |\gamma_{\Lambda_m}| \right] = \mathbb{E}_{P_\eta} \left[f \mathbb{E}_{P_\eta} \left[|\gamma_{\Lambda_m}| \big| \mathscr{F}_{\Lambda_m^c} \right] \right],$$

where the last equality is simply the definition of a conditional expectation. Since this is true for any f bounded and $\mathscr{F}_{\Lambda_m^c}$-measurable, we get (3.12), which concludes the proof. $\square$

3.4 Ergodic solutions of DLR equations which are not number-rigid

Theorem 3.7 gave a description of the Campbell measures of stationary solutions of canonical DLR. In this section we further assume these solutions are ergodic and not number-rigid and improve on the previous description. Let us recall that the main rigidity result we have in mind, as stated in Theorem 3.2, claims that solutions of the DLR equations are *all* rigid, so the result of the present section should turn out to be empty; the next theorem is the main part of our proof by contradiction.

Theorem 3.15. Let P be an ergodic point process on $\mathbb{R}$ satisfying canonical DLR and assume that P is not number-rigid. Then there exists $n \geq 1$ and a measurable function $\mathsf{create}_n : \mathsf{Conf}(\mathbb{R}) \to [0, \infty)$ such that the Campbell measure $\mathsf{C}_P^{(n)}$ is absolutely continuous with respect to $\mathsf{Leb}^{\otimes n} \otimes P$ and with density

$$\frac{\mathrm{d}\mathsf{C}_P^{(n)}}{\mathrm{d}\mathsf{Leb}^{\otimes n} \otimes P}(\mathbf{x}_n, \gamma) = \mathsf{create}_n(\gamma)\, \mathrm{e}^{-\beta \mathsf{Cost}(\mathbf{x}_n, \gamma)}, \tag{3.15}$$

where Cost is defined in (3.2).

Compared to Theorem 3.7, the important change is that we gained information on the second marginal of $\mathsf{C}_P^{(n)}$: Instead of being some abstract measure Q_n as in (3.3), here we see that it is absolutely continuous with respect to P itself and obtain crucial information on its density.

Proof. Let P be an ergodic point process on $\mathbb{R}$ satisfying canonical DLR which is not number-rigid. Corollary 3.12 provides $n \geq 1$ such that $P(\mathsf{CanRmv}_n^P) = 1$. Recalling Theorem 3.7 it is enough to prove that the measure Q_n of (3.3) is absolutely continuous with respect to P. To do so, let E be an event such that $P(E) = 0$ and prove that $\mathsf{Q}_n(E) = 0$, which is enough to prove the existence of create_n according to the Radon-Nikodym theorem.

Since $\mathsf{Cost}(\mathbf{x}_n, \gamma)$ is finite for P-a.e. γ and $\mathbf{x}_n \in \mathbb{R}^n$, the density (3.3) is positive and it is enough to show that $\mathsf{C}_P^{(n)}([0, 1]^n \times E) = 0$. By definition of Campbell measures, we have

$$\mathsf{C}_P^{(n)}([0, 1]^n \times E) = \int \sum_{\mathbf{x}_n \subset \gamma} \mathbf{1}_{[0,1]^n}(\mathbf{x}_n)\, \mathbf{1}_E(\gamma \setminus \mathbf{x}_n) P(\mathrm{d}\gamma).$$

Since $P(\mathsf{CanRmv}_n) = 1$, by monotonicity we also have $\lim_{m \to \infty} P(\mathsf{CanRmv}_{n, \Lambda_m}) = 1$ and thus, by monotone convergence:

$$\mathsf{C}_P^{(n)}([0, 1]^n \times E) = \lim_{m \to \infty} \int \mathbf{1}_{\mathsf{CanRmv}_{n, \Lambda_m}}(\gamma) \sum_{\mathbf{x}_n \subset \gamma} \mathbf{1}_{[0,1]^n}(\mathbf{x}_n)\, \mathbf{1}_E(\gamma \setminus \mathbf{x}_n) P(\mathrm{d}\gamma).$$

Recalling (3.6), for any integers n, k, we define

$$A_{n,k,\Lambda}^P := \Big\{ \gamma \in \mathsf{Conf}(\mathbb{R}) : \exists\, p \leq k,$$
$$P\left(\mathsf{HasPoints}_{p,\Lambda} \Big| \mathsf{Ext}_\Lambda^\gamma\right) > 0 \ \text{ and } \ P\left(\mathsf{HasPoints}_{p+n,\Lambda} \Big| \mathsf{Ext}_\Lambda^\gamma\right) > 0 \Big\} \tag{3.16}$$

and

$$\mathsf{CanRmv}_{n,k,\Lambda_m} := A_{n,k,\Lambda}^P \setminus A_{n,k-1,\Lambda}^P,$$

so that one can write the disjoint union: $\mathsf{CanRmv}_{n,k,\Lambda_m} = \cup_k \mathsf{CanRmv}_{n,k,\Lambda_m}$. Now, if we set:

$$\mathsf{Term}_{k,m} := \int \mathbf{1}_{\mathsf{HasPoints}_{k,\Lambda_m}}(\gamma) \mathbf{1}_{\mathsf{CanRmv}_{n,k,\Lambda_m}}(\gamma) \sum_{\mathbf{x}_n \subset \gamma} \mathbf{1}_{[0,1]^n}(\mathbf{x}_n)\, \mathbf{1}_E(\gamma \setminus \mathbf{x}_n) P(\mathrm{d}\gamma),$$

we obtain

$$\mathsf{C}_P^{(n)}([0, 1]^n \times E) \leq \lim_{m \to \infty} \sum_{k=0}^{\infty} \mathsf{Term}_{k,m}. \tag{3.17}$$

36

We introduce for convenience

$$\mathsf{Dens}^{\Lambda_m}_{\gamma_{\Lambda_m^c},k}(\eta) := \frac{1}{Z_{\Lambda_m,\mathbb{R}}(\gamma)} \, e^{-\beta(\mathsf{H}_{\Lambda_m}(\eta) + \mathsf{M}_{\Lambda_m,\mathbb{R}}(\eta,\gamma))} \, \mathbf{1}_{|\eta|=k}$$

the density with respect to $\mathbf{B}_{k,\Lambda_m}$ of P conditionally to $\gamma_{\Lambda_m^c}$ and having k points in Λ_m, provided it makes sense. Recalling Remark 2.4, the right hand side does only depend on γ_{Λ_m} through its cardinality, which explains the notation. By using canonical DLR we obtain

$$\mathsf{Term}_{k,m} = \int \mathbf{1}_{\mathsf{CanRmv}_{n,k,\Lambda_m}}(\eta \cup \gamma_{\Lambda_m^c})$$
$$\times \sum_{\mathbf{x}_n \subset \eta} \mathbf{1}_{[0,1]^n}(\mathbf{x}_n) \mathbf{1}_E(\eta \cup \gamma_{\Lambda_m^c} \setminus \mathbf{x}_n) \mathsf{Dens}^{\Lambda_m}_{\gamma_{\Lambda_m^c},k}(\eta) \mathbf{B}_{k,\Lambda_m}(\mathrm{d}\eta) P(\mathrm{d}\gamma).$$

Similarly as in the proof of Theorem 3.7, we set $\tilde{\eta} := \eta \setminus \mathbf{x}_n$ and obtain,

$$\mathsf{Term}_{k,m} = \frac{k!}{m^n(k-n)!} \int \mathbf{1}_{\mathsf{CanRmv}_{n,k,\Lambda_m}}(\tilde{\eta} \cup \mathbf{x}_n \cup \gamma_{\Lambda_m^c}) \mathbf{1}_E(\tilde{\eta} \cup \gamma_{\Lambda_m^c})$$
$$\times \mathsf{Dens}^{\Lambda_m}_{\gamma_{\Lambda_m^c},k}(\tilde{\eta} \cup \mathbf{x}_n) \mathbf{B}_{n,[0,1]}(\mathrm{d}\mathbf{x}_n) \mathbf{B}_{k-n,\Lambda_m}(\mathrm{d}\tilde{\eta}) P(\mathrm{d}\gamma)$$
$$= \frac{k!}{m^n(k-n)!} \int \mathbf{1}_{\mathsf{CanRmv}_{n,k,\Lambda_m}}(\tilde{\eta} \cup \mathbf{x}_n \cup \gamma_{\Lambda_m^c}) \mathbf{1}_E(\tilde{\eta} \cup \gamma_{\Lambda_m^c}) \mathsf{Dens}^{\Lambda_m}_{\gamma_{\Lambda_m^c},k-n}(\tilde{\eta})$$
$$\times \left(\int \frac{\mathsf{Dens}^{\Lambda_m}_{\gamma_{\Lambda_m^c},k}(\tilde{\eta} \cup \mathbf{x}_n)}{\mathsf{Dens}^{\Lambda_m}_{\gamma_{\Lambda_m^c},k-n}(\tilde{\eta})} \mathbf{B}_{n,[0,1]}(\mathrm{d}\mathbf{x}_n) \right) \mathbf{B}_{k-n,\Lambda_m}(\mathrm{d}\tilde{\eta}) P(\mathrm{d}\gamma). \quad (3.18)$$

The integrand inside the parentheses is finite since the denominator is positive on the event where $\tilde{\eta} \cup \mathbf{x}_n \cup \gamma_{\Lambda_m^c} \in \mathsf{CanRmv}_{n,k,\Lambda_m}$. Since canonical DLR yields for any $k, n, m \geq 0$ and $\mathbf{x}_n \in \mathbb{R}^n$,

$$P(E) = \int \mathbf{1}_E(\tilde{\eta} \cup \gamma_{\Lambda_m^c}) \mathsf{Dens}^{\Lambda_m}_{\gamma_{\Lambda_m^c},|\gamma_{\Lambda_m}|}(\tilde{\eta}) \, \mathbf{B}_{|\gamma_{\Lambda_m}|,\Lambda_m}(\mathrm{d}\tilde{\eta}) P(\mathrm{d}\gamma)$$
$$\geq \int \mathbf{1}_{\mathsf{CanRmv}_{n,k,\Lambda_m}}(\tilde{\eta} \cup \mathbf{x}_n \cup \gamma_{\Lambda_m^c}) \mathbf{1}_E(\tilde{\eta} \cup \gamma_{\Lambda_m^c}) \mathsf{Dens}^{\Lambda_m}_{\gamma_{\Lambda_m^c},k-n}(\eta) \, \mathbf{B}_{k-n,\Lambda_m}(\mathrm{d}\tilde{\eta}) P(\mathrm{d}\gamma)$$

and $P(E) = 0$ by assumption, we obtain with (3.18) that $\mathsf{Term}_{k,m} = 0$ for every k, m. Thus $\mathsf{C}_P^{(n)}([0,1]^n \times E) = 0$ and this concludes the proof of Theorem 3.15. $\qquad \square$

Now let us give some properties of the function create_n. In Lemmas 3.16 and 3.17, we work under the same assumptions as in Theorem 3.15.

First, we have the following simple result concerning the effect of translations on create_n appearing in Theorem 3.15, which is a direct consequence of the stationarity.

Lemma 3.16. Under the assumptions of Theorem 3.15, for any $x \in \mathbb{R}$, P-a.e. γ and $\mathsf{Leb}^{\otimes n}$-a.e. $\mathbf{x}_n$, we have

$$\mathsf{create}_n(\gamma) \, e^{-\beta \mathsf{Cost}(\mathbf{x}_n,\gamma)} = \mathsf{create}_n(\gamma - x) \, e^{-\beta \mathsf{Cost}(\mathbf{x}_n + x, \gamma - x)}.$$

The following property of create_n will be the crucial for the forthcoming proof by contradiction of number-rigidity.

Lemma 3.17. Under the assumptions of Theorem 3.15, for P-a.e. γ, $\mathsf{Leb}^{\otimes n}$-a.e. $\mathbf{y}_n$, and any $\mathbf{x}_n \subset \gamma$, we have

$$\mathsf{create}_n(\gamma) = \mathsf{create}_n\left(\gamma \setminus \mathbf{x}_n \cup \mathbf{y}_n\right) \mathrm{e}^{n\beta \sum_{j=1}^{n} g(y_j) - g(x_j)}. \tag{3.19}$$

Proof. It is enough to show that, for Q_n-a.e. γ and $\mathsf{Leb}^{\otimes n}$-a.e. $\mathbf{x}_n, \mathbf{y}_n$, we have

$$\mathsf{create}_n(\gamma \cup \mathbf{x}_n)\, \mathrm{e}^{n\beta \sum_{j=1}^{n} g(x_j)} = \mathsf{create}_n\left(\gamma \cup \mathbf{y}_n\right)\, \mathrm{e}^{n\beta \sum_{j=1}^{n} g(y_j)}. \tag{3.20}$$

Indeed, let $h : \mathbb{R}^n \times \mathbb{R}^n \times \mathsf{Conf}(\mathbb{R}) \to [0, \infty)$ be a measurable test function. We have by definition of $\mathsf{C}_P^{(n)}$, Theorem 3.7 and (3.20),

$$\int \sum_{\mathbf{x}_n \subset \gamma} h(\mathbf{x}_n, \mathbf{y}_n, \gamma)\, \mathsf{create}_n(\gamma)\, P(\mathrm{d}\gamma)\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{y}_n)$$

$$= \int \mathsf{C}_P^{(n)}\left[(\mathbf{x}_n, \gamma) \mapsto h(\mathbf{x}_n, \mathbf{y}_n, \gamma \cup \mathbf{x}_n)\mathsf{create}_n(\gamma \cup \mathbf{x}_n)\right]\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{y}_n)$$

$$= \int h(\mathbf{x}_n, \mathbf{y}_n, \gamma \cup \mathbf{x}_n)\mathsf{create}_n(\gamma \cup \mathbf{x}_n)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n, \gamma)}\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{x}_n)\mathsf{Q}_n(\mathrm{d}\gamma)\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{y}_n)$$

$$= \int h(\mathbf{x}_n, \mathbf{y}_n, \gamma \cup \mathbf{x}_n)\mathsf{create}_n(\gamma \cup \mathbf{y}_n)\mathrm{e}^{n\beta \sum_{j=1}^n g(y_j) - g(x_j)}\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n, \gamma)}\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{x}_n)\mathsf{Q}_n(\mathrm{d}\gamma)\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{y}_n)$$

$$= \int \mathsf{C}_P^{(n)}\left[(\mathbf{x}_n, \gamma) \mapsto h(\mathbf{x}_n, \mathbf{y}_n, \gamma \cup \mathbf{x}_n)\mathsf{create}_n(\gamma \cup \mathbf{y}_n)\mathrm{e}^{n\beta \sum_{j=1}^n g(y_j) - g(x_j)}\right]\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{y}_n)$$

$$= \int \sum_{\mathbf{x}_n \subset \gamma} h(\mathbf{x}_n, \mathbf{y}_n, \gamma)\, \mathsf{create}_n(\gamma \cup \mathbf{y}_n \setminus \mathbf{x}_n)\mathrm{e}^{n\beta \sum_{j=1}^n g(y_j) - g(x_j)}\, P(\mathrm{d}\gamma)\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{y}_n),$$

from which our claim follows.

Let $f : \mathbb{R}^n \times \mathbb{R}^n \times \mathsf{Conf}(\mathbb{R}) \to [0, \infty)$ be any measurable function, and set

$$f_1(\mathbf{x}_n, \gamma) := \sum_{\mathbf{y}_n \subset \gamma} f(\mathbf{x}_n, \mathbf{y}_n, \gamma \setminus \mathbf{y}_n), \qquad f_2(\mathbf{y}_n, \gamma) := \sum_{\mathbf{x}_n \subset \gamma} f(\mathbf{x}_n, \mathbf{y}_n, \gamma \setminus \mathbf{x}_n).$$

We have (writing $(\mathbf{x}_n, \mathbf{y}_n) = \mathbf{z}_{2n} \in \mathbb{R}^{2n}$),

$$\mathsf{C}_P^{(2n)}(f) = \int \sum_{\mathbf{z}_{2n} \subset \gamma} f(\mathbf{z}_{2n}, \gamma \setminus \mathbf{z}_{2n})P(\mathrm{d}\gamma)$$

$$= \int \sum_{\mathbf{x}_n \subset \gamma}\left(\sum_{\mathbf{y}_n \in \gamma \setminus \mathbf{x}_n} f(\mathbf{x}_n, \mathbf{y}_n, \gamma \setminus (\mathbf{x}_n \cup \mathbf{y}_n))\right)P(\mathrm{d}\gamma)$$

$$= \mathsf{C}_P^{(n)}(f_1)$$

and similarly $\mathsf{C}_P^{(2n)}(f) = \mathsf{C}_P^{(n)}(f_2)$. Next, using Theorem 3.15 and then Theorem 3.7, we obtain

$$\mathsf{C}_P^{(n)}(f_1) = \int \mathsf{create}_n(\gamma)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n, \gamma)}\left[\sum_{\mathbf{y}_n \in \gamma} f(\mathbf{x}_n, \mathbf{y}_n, \gamma \setminus \mathbf{y}_n)\right]\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{x}_n)P(\mathrm{d}\gamma)$$

$$= \int \mathsf{C}_P^{(n)}\left[(\mathbf{y}_n, \gamma) \mapsto \mathsf{create}_n(\gamma \cup \mathbf{y}_n)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n, \gamma \cup \mathbf{y}_n)}f(\mathbf{x}_n, \mathbf{y}_n, \gamma)\right]\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{x}_n)$$

$$= \int \mathsf{create}_n(\gamma \cup \mathbf{y}_n)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n, \gamma \cup \mathbf{y}_n) - \beta\mathsf{Cost}(\mathbf{y}_n, \gamma)}$$

$$\times f(\mathbf{x}_n, \mathbf{y}_n, \gamma)\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{x}_n)\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{y}_n)\mathsf{Q}_n(\mathrm{d}\gamma).$$

The same computation with f_2 yields

$$\mathsf{C}_P^{(n)}(f_2) = \int \mathsf{create}_n(\gamma \cup \mathbf{x}_n)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{y}_n,\gamma\cup\mathbf{x}_n)-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma)}$$
$$\times f(\mathbf{x}_n,\mathbf{y}_n,\gamma)\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{x}_n)\mathsf{Leb}^{\otimes n}(\mathrm{d}\mathbf{y}_n)\mathsf{Q}_n(\mathrm{d}\gamma).$$

Thus, since the test function f is arbitrary, we have for Q_n-a.e. γ and $\mathsf{Leb}^{\otimes n}$-a.e. $\mathbf{x}_n, \mathbf{y}_n$,

$$\mathsf{create}_n(\gamma \cup \mathbf{x}_n)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{y}_n,\gamma\cup\mathbf{x}_n)-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma)} = \mathsf{create}_n(\gamma \cup \mathbf{y}_n)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma\cup\mathbf{y}_n)-\beta\mathsf{Cost}(\mathbf{y}_n,\gamma)},$$

from which, after using the definition of Cost and several simplifications, we obtain (3.20).
$\qquad\square$

3.5 Proof of Theorem 3.2

We are finally in position to provide a proof for Theorem 3.2. Let P be a stationary point process satisfying canonical DLR.

3.5.1 Restriction to ergodic processes.

First, as a consequence of Proposition 3.13 and Proposition 3.14, we assume without loss of generality that P is further ergodic in the assumptions of Theorem 3.2 to prove this theorem.

3.5.2 Strategy of the proof.

The proof goes by contradiction.

Assume that the ergodic point process P is **not** number-rigid. By Theorem 3.15, there exists $n \geq 1$ and a function create_n such that (3.15) holds true; from now this n is fixed. For any $m \geq 1$ and $K \geq 0$, we denote by N_m^K the event

$$\mathsf{N}_m^K := \left\{ \gamma \in \mathsf{Conf}(\mathbb{R}) : |\gamma_{[m,m+1]}| \leq K \right\} \tag{3.21}$$

and consider the test function $\mathsf{Test}_{m,K}$ defined by

$$\mathsf{Test}_{m,K}(\mathbf{x}_n,\gamma) = \mathbf{1}_{[0,1]^n}(\mathbf{x}_n)\mathbf{1}_{\mathsf{N}_0^K}(\gamma)\mathbf{1}_{\mathsf{N}_m^K}(\gamma) \sum_{\mathbf{y}_n \subset \gamma} \mathbf{1}_{[m,m+1]^n}(\mathbf{y}_n). \tag{3.22}$$

In the following K is fixed and always assumed to be larger than $2n$.

We compute, using that $\mathbf{x}_n \in [0,1]^n$ and $m \geq 1$,

$$\mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K}) = \int \sum_{\mathbf{x}_n \subset \gamma} \mathbf{1}_{[0,1]^n}(\mathbf{x}_n)\mathbf{1}_{\mathsf{N}_0^K}(\gamma \setminus \mathbf{x}_n)\mathbf{1}_{\mathsf{N}_m^K}(\gamma \setminus \mathbf{x}_n) \sum_{\mathbf{y}_n \subset \gamma\setminus\mathbf{x}_n} \mathbf{1}_{[m,m+1]^n}(\mathbf{y}_n)P(\mathrm{d}\gamma)$$

$$= \int \mathbf{1}_{\mathsf{N}_0^{K+n}}(\gamma)\mathbf{1}_{\mathsf{N}_m^K}(\gamma) \sum_{\mathbf{x}_n \subset \gamma} \mathbf{1}_{[0,1]^n}(\mathbf{x}_n) \sum_{\mathbf{y}_n \subset \gamma} \mathbf{1}_{[m,m+1]^n}(\mathbf{y}_n)P(\mathrm{d}\gamma).$$

Thus, forgetting about the truncation $\mathbf{1}_{\mathsf{N}_0^{K+n}}(\gamma)\mathbf{1}_{\mathsf{N}_m^K}(\gamma)$, we should think of $\mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K})$ as encoding the correlation between the number of n-tuples from γ with law P falling into $[0,1]$ and $[m,m+1]$ respectively. The main idea to reach the contradiction is that

$\mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K})$ is bounded from above independently on m, since we have the rough upper bound,

$$\mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K}) \leq (K+n)^n \times K^n, \tag{3.23}$$

but we will prove that its Cesáro series diverges by using the long range of the logarithmic interaction.

3.5.3 Manipulations on the test function - Part I

By using Theorem 3.15, introducing phantom variables $\mathbf{z}_n$ and then using Lemma 3.17, we can write

$$\mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K})$$
$$= \int \mathsf{create}_n(\gamma)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma)}\mathbf{1}_{\mathsf{N}_0^K}(\gamma)\mathbf{1}_{\mathsf{N}_m^K}(\gamma)\sum_{\mathbf{y}_n\subset\gamma}\mathbf{1}_{[m,m+1]^n}(\mathbf{y}_n)P(\mathrm{d}\gamma)\mathsf{Leb}_{[0,1]}^{\otimes n}(\mathrm{d}\mathbf{x}_n)$$
$$= \int \mathsf{create}_n(\gamma)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma)}\mathbf{1}_{\mathsf{N}_0^K}(\gamma)\mathbf{1}_{\mathsf{N}_m^K}(\gamma)\sum_{\mathbf{y}_n\subset\gamma}\mathbf{1}_{[m,m+1]^n}(\mathbf{y}_n)P(\mathrm{d}\gamma)\mathsf{Leb}_{[0,1]}^{\otimes n}(\mathrm{d}\mathbf{x}_n)\mathsf{Leb}_{[1,2]}^{\otimes n}(\mathrm{d}\mathbf{z}_n)$$
$$= \int \mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma)}\mathbf{1}_{\mathsf{N}_0^K}(\gamma)\mathbf{1}_{\mathsf{N}_m^K}(\gamma)\sum_{\mathbf{y}_n\subset\gamma}\mathsf{create}_n(\gamma\setminus\mathbf{y}_n\cup\mathbf{z}_n)\mathbf{1}_{[m,m+1]^n}(\mathbf{y}_n)\mathrm{e}^{n\beta\sum_{j=1}^n g(z_j)-g(y_j)}$$
$$\times P(\mathrm{d}\gamma)\mathsf{Leb}_{[0,1]}^{\otimes n}(\mathrm{d}\mathbf{x}_n)\mathsf{Leb}_{[1,2]}^{\otimes n}(\mathrm{d}\mathbf{z}_n).$$

Next, we use that for any $\mathbf{z}_n \in [1,2]^n$ and $\mathbf{y}_n \in [m,m+1]^n$ we have

$$\mathrm{e}^{n\beta\sum_{j=1}^n g(z_j)-g(y_j)} \geq \mathrm{e}^{-Cg(m)}, \tag{3.24}$$

for any m large enough and $C > 0$ only depending on n, β. This yields

$$\mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K}) \geq \mathrm{e}^{-Cg(m)}\int \mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma)}\mathbf{1}_{\mathsf{N}_0^K}(\gamma)\mathbf{1}_{\mathsf{N}_m^K}(\gamma)$$
$$\times \sum_{\mathbf{y}_n\subset\gamma}\mathsf{create}_n(\gamma\setminus\mathbf{y}_n\cup\mathbf{z}_n)\mathbf{1}_{[m,m+1]^n}(\mathbf{y}_n)P(\mathrm{d}\gamma)\mathsf{Leb}_{[0,1]}^{\otimes n}(\mathrm{d}\mathbf{x}_n)\mathsf{Leb}_{[1,2]}^{\otimes n}(\mathrm{d}\mathbf{z}_n) \quad (3.25)$$

for any m sufficiently large.

3.5.4 Manipulations on the test function - Part II

The integrand in (3.25) can also be interpreted in terms of a Campbell measure, and using again Theorem 3.15, we have for any fixed $\mathbf{x}_n, \mathbf{z}_n$,

$$\int \mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma)}\mathbf{1}_{\mathsf{N}_0^K}(\gamma)\mathbf{1}_{\mathsf{N}_m^K}(\gamma)\sum_{\mathbf{y}_n\subset\gamma}\mathsf{create}_n(\gamma\setminus\mathbf{y}_n\cup\mathbf{z}_n)\mathbf{1}_{[m,m+1]^n}(\mathbf{y}_n)P(\mathrm{d}\gamma)$$
$$= \int \mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma\cup\mathbf{y}_n)-\beta\mathsf{Cost}(\mathbf{y}_n,\gamma)}\mathbf{1}_{\mathsf{N}_0^K}(\gamma\cup\mathbf{y}_n)\mathbf{1}_{\mathsf{N}_m^K}(\gamma\cup\mathbf{y}_n)\mathsf{create}_n(\gamma\cup\mathbf{z}_n)\mathsf{create}_n(\gamma)$$
$$\times P(\mathrm{d}\gamma)\mathsf{Leb}_{[m,m+1]}^{\otimes n}(\mathrm{d}\mathbf{y}_n)$$
$$= \int \mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma\cup\mathbf{y}_n)-\beta\mathsf{Cost}(\mathbf{y}_n,\gamma)}\mathbf{1}_{\mathsf{N}_0^K}(\gamma)\mathbf{1}_{\mathsf{N}_m^{K-n}}(\gamma)\mathsf{create}_n(\gamma\cup\mathbf{z}_n)\mathsf{create}_n(\gamma) \tag{3.26}$$
$$\times P(\mathrm{d}\gamma)\mathsf{Leb}_{[m,m+1]}^{\otimes n}(\mathrm{d}\mathbf{y}_n).$$

Moreover, recalling Definition 3.5, we have for any $\mathbf{x}_n \in [0,1]^n$ and $\mathbf{y}_n \in [m, m+1]^n$,

$$\left|\mathsf{Cost}(\mathbf{x}_n, \gamma \cup \mathbf{y}_n) - \mathsf{Cost}(\mathbf{x}_n, \gamma)\right| = \left|\sum_{i,j=1}^{n} g(y_j - x_i) - g(y_j)\right| \le c \tag{3.27}$$

for some $c > 0$ depending only on n, β. Thus, together with (3.25)–(3.26), we obtain

$$\mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K}) \ge \mathrm{e}^{-(Cg(m)+c)} \int \mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma) - \beta\mathsf{Cost}(\mathbf{y}_n,\gamma)} \mathbf{1}_{\mathsf{N}_0^K}(\gamma) \mathbf{1}_{\mathsf{N}_m^{K-n}}(\gamma)$$
$$\times \mathsf{create}_n(\gamma \cup \mathbf{z}_n)\mathsf{create}_n(\gamma) P(\mathrm{d}\gamma) \mathsf{Leb}_{[0,1]}^{\otimes n}(\mathrm{d}\mathbf{x}_n)\mathsf{Leb}_{[m,m+1]}^{\otimes n}(\mathrm{d}\mathbf{y}_n)\mathsf{Leb}_{[1,2]}^{\otimes n}(\mathrm{d}\mathbf{z}_n). \tag{3.28}$$

Finally, in order to isolate the dependence in m in the remaining integral, let us set

$$\mathsf{II}_{m,K}(\gamma) := \int \mathsf{create}_n(\gamma)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{y}_n,\gamma)}\mathbf{1}_{\mathsf{N}_m^{K-n}}(\gamma)\mathsf{Leb}_{[m,m+1]}^{\otimes n}(\mathrm{d}\mathbf{y}_n) \tag{3.29}$$

so that the previous inequality reads

$$\mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K}) \ge \mathrm{e}^{-(Cg(m)+c)} \int \mathsf{II}_{m,K}(\gamma)\, \mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma)}\mathbf{1}_{\mathsf{N}_0^K}(\gamma)\mathsf{create}_n(\gamma \cup \mathbf{z}_n)$$
$$\times P(\mathrm{d}\gamma)\mathsf{Leb}_{[0,1]}^{\otimes n}(\mathrm{d}\mathbf{x}_n)\mathsf{Leb}_{[1,2]}^{\otimes n}(\mathrm{d}\mathbf{z}_n). \tag{3.30}$$

3.5.5 An ergodic argument

We now make change the change of variables $\mathbf{y}_n \mapsto \mathbf{y}_n + m$ in (3.29) to obtain from Lemma 3.16,

$$\mathsf{II}_{m,K}(\gamma) = \int \mathsf{create}_n(\gamma)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{y}_n+m,\gamma)}\mathbf{1}_{\mathsf{N}_m^{K-n}}(\gamma)\mathsf{Leb}_{[0,1]}^{\otimes n}(\mathrm{d}\mathbf{y}_n)$$
$$= \int \mathsf{create}_n(\gamma - m)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{y}_n,\gamma-m)}\mathbf{1}_{\mathsf{N}_0^{K-n}}(\gamma - m)\mathsf{Leb}_{[0,1]}^{\otimes n}(\mathrm{d}\mathbf{y}_n)$$
$$= \mathsf{II}_{0,K}(\gamma - m).$$

Since P is ergodic by assumption, the ergodic theorem implies that P-a.s.

$$\lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} \mathsf{II}_{m,K}(\gamma) = \mathbb{E}_P\big[\mathsf{II}_{0,K}\big].$$

Note that $\mathbb{E}_P\big[\mathsf{II}_{0,K}\big] > 0$, since

$$\mathbb{E}_P\big[\mathsf{II}_{0,K}\big] = \int \mathsf{create}_n(\eta)\mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{y}_n,\eta)}\mathbf{1}_{\mathsf{N}_0^{K-n}}(\eta)\mathsf{Leb}_{[0,1]}^{\otimes n}(\mathrm{d}\mathbf{y}_n)P(\mathrm{d}\eta)$$
$$= \mathsf{C}_P^{(n)}\left(\mathbf{1}_{[0,1]^n}(\mathbf{y}_n)\mathbf{1}_{\mathsf{N}_0^{K-n}}(\gamma)\right)$$
$$= \int \sum_{\mathbf{y}_n \subset \gamma} \mathbf{1}_{[0,1]^n}(\mathbf{y}_n)\mathbf{1}_{\mathsf{N}_0^K}(\gamma \setminus \mathbf{y}_n)P(\mathrm{d}\gamma) \tag{3.31}$$
$$\ge P(|\gamma_{[0,1]}| = n) > 0$$

by Corollary 3.3; we used that $K \ge n$ by assumption. Fatou's lemma then yields

$$\liminf_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} \mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K}) \ge \left(\liminf_{M\to\infty}\frac{1}{M}\sum_{m=1}^{M}\mathrm{e}^{-(Cg(m)+c)}\right)\mathbb{E}_P\big[\mathsf{II}_{0,K}\big]$$
$$\times \int \mathrm{e}^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma)}\mathsf{create}_n(\gamma \cup \mathbf{z}_n)\mathbf{1}_{\mathsf{N}_0^K}(\gamma)P(\mathrm{d}\gamma)\mathsf{Leb}_{[0,1]}^{\otimes n}(\mathrm{d}\mathbf{x}_n)\mathsf{Leb}_{[1,2]}^{\otimes n}(\mathbf{z}_n). \tag{3.32}$$

3.5.6 The remaining integral and conclusion

If the integral in (3.32) vanishes, then

$$e^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma)}\mathsf{create}_n(\gamma\cup\mathbf{z}_n)\mathbf{1}_{\mathsf{N}_0^K}(\gamma)=0, \tag{3.33}$$

$P\otimes\mathsf{Leb}_{[0,1]}^{\otimes n}\otimes\mathsf{Leb}_{[1,2]}^{\otimes n}$-almost everywhere. Let us fix $m_0\geq 1$. Note that when we proved the lower bound (3.28), the only inequalities we have used were (3.24) and (3.27). Since the converse inequality also holds in (3.24) after changing the constant C to another positive constant and (3.27) is both-sided, the exact same line of arguments yields $\kappa>0$ only depending on m_0,n,β such that

$$\mathsf{C}_P^{(n)}(\mathsf{Test}_{m_0,K})\leq\kappa\int\mathsf{II}_{m_0,K}(\gamma)\,e^{-\beta\mathsf{Cost}(\mathbf{x}_n,\gamma)}\mathsf{create}_n(\gamma\cup\mathbf{z}_n)\mathbf{1}_{\mathsf{N}_0^K}(\gamma)$$
$$\times P(\mathrm{d}\gamma)\mathsf{Leb}_{[0,1]}^{\otimes n}(\mathrm{d}\mathbf{x}_n)\mathsf{Leb}_{[1,2]}^{\otimes n}(\mathrm{d}\mathbf{z}_n). \tag{3.34}$$

Recalling (3.31), we have $\mathbb{E}_P[\mathsf{II}_{m_0,K}]=\mathbb{E}_P[\mathsf{II}_{0,K}]\leq(K+n)^n<\infty$ and thus (3.33)–(3.34) imply together that $\mathsf{C}_P^{(n)}(\mathsf{Test}_{m_0,K})=0$. However, we have the rough lower bound

$$\mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K})\geq P\left(|\gamma_{[0,1]}|=n\cap|\gamma_{[m,m+1]}|=n\right),$$

as soon as $K\geq 2n$, and Corollary 3.3 then yields that $\mathsf{C}_P^{(n)}(\mathsf{Test}_{m_0,K})>0$.

As a consequence, the remaining integral in (3.32) is positive and, because $g(m)=-\log(m)\to-\infty$ as $m\to\infty$, it follows from (3.32) that

$$\liminf_{M\to\infty}\frac{1}{M}\sum_{m=1}^{M}\mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K})=\infty. \tag{3.35}$$

Recalling the upper bound (3.23) on $\mathsf{C}_P^{(n)}(\mathsf{Test}_{m,K})$, we finally reached a contradiction. The proof of Theorem 3.2 is therefore complete.

3.5.7 A more general statement for number-rigidity

A careful examination of the proof of Theorem 3.2 reveals that we did not use much of the properties of the vector space $\mathbb{R}$ nor of the logarithmic interacting potential g. More precisely, the properties of g are only used in equalities (3.24), (3.27) and when stating that

$$\liminf_{M\to\infty}\frac{1}{M}\sum_{m=1}^{M}e^{-(Cg(m)+c)}=\infty.$$

As for the compact sets $[0,1]$, $[1,2]$ and $[m,m+1]$ appearing in the proof, they can be replaced by arbitrary disjoint compacts sets K_0,K_1,K_m with unit Lebesgue volume and such that the distance from K_m to K_0 and K_1 goes to infinity with m.

Thus, let us consider more general interaction potentials $g:\mathbb{R}^d\to\mathbb{R}\cup\{+\infty\}$ for $d\geq 1$ and redefine (2.8)–(2.16) accordingly by using this new g in their definition (where we now set $\Lambda_m:=[-\frac{m}{2},\frac{m}{2}]^d$). Let us also assume there exists a non-trivial class $\mathscr{C}$ of stationary point processes for which the move functions exists, namely so that the results of Lemma 2.15 holds for any $P\in\mathscr{C}$ and, having in mind the proof of Proposition 3.13,

that $\mathscr{C}$ is stable by desintegration: if $P \in \mathscr{C}$ can be written as (3.10), then $P_\eta \in \mathscr{C}$ for P-a.e. η (recall that for the one dimensional logarithmic interaction $\mathscr{C}$ were the class of stationary point processes that have finite renormalized energy). We then say that a stationary point process P on $\mathbb{R}^d$ satisfies the *canonical DLR equations with respect to g* if $P \in \mathscr{C}$ and canonical DLR holds with the new definition for $f_{\Lambda, \mathbb{R}^d}$. In this more general setting, cosmetic modifications of the proof of Theorem 3.2 leads to the following result.

Theorem 3.18. Let $d \geq 1$ and $g : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a measurable function which satisfies $g(x) \to -\infty$ when $\|x\| \to \infty$ and assume that there exists a compact set $K \subset \mathbb{R}^d$ such that g is continuous on $\mathbb{R}^d \setminus K$. If P is a stationary point process on $\mathbb{R}^d$ which satisfies the canonical DLR equations with respect to g, then P is number-rigid.

In particular, this results applies to the logarithmic potential $g(x) = -\log \|x\|$ on $\mathbb{R}^d$ for any $d \geq 1$, including the $d = 2$ Coulomb interaction. The question of identifying an appropriate class $\mathscr{C}$ for this setting will be investigated in another work. Note that Theorem 3.18 does not cover, however, the Coulomb interaction $g(x) = \|x\|^{-(d-2)}$ in dimension $d \geq 3$ or, more generally, the Riesz interactions $g(x) = \|x\|^{-s}$ for any $s \in \mathbb{R}$.

References

M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

R. Allez and L. Dumaz. From Sine kernel to Poisson statistics. *Electron. J. Probab.*, 19 (114):25pp., 2014.

F. Bekerman and A. Lodhia. Mesoscopic central limit theorem for general β-ensembles. *Annales de l'Institut Henri Poincaré (to appear). Preprint arXiv:1605.05206*, 2018.

F. Bekerman, T. Leblé, and S. Serfaty. CLT for fluctuations of β-ensembles with general potential. *Electronic journal of probability (to appear). Preprint arXiv:1706.09663*, 2018.

A. Borodin and S. Serfaty. Renormalized energy concentration in random matrices. *Comm. Math. Phys.*, 320(1):199–244, 2013.

P. Bourgade, L. Erdös, and H.-T. Yau. Bulk universality of general β-ensembles with non-convex potential. *J. Math. Phy., special issue in honor of E. Lieb's 80th birthday,*, 53, 2012.

P. Bourgade, L. Erdös, and H.-T. Yau. Universality of general β-ensembles. *Duke Math. J.*, 163(6):1127–1190, 2014.

A. I Bufetov. Conditional measures of determinantal point processes. *arXiv preprint arXiv:1605.01400*, 2016.

R. Chhaibi and J. Najnudel. Rigidity of the Sine$_\beta$ process. *arXiv preprint arXiv:1804.01216*, 2018.

D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. II.* Probability and its Applications (New York). Springer, New York, second edition, 2008.

P. Deift and D. Gioev. *Random matrix theory: invariant ensembles and universality*, volume 18. Courant Lecture Notes in Mathematics. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, Rhode Island, 2009.

D. Dereudre. Introduction to the theory of Gibbs point processes. *arXiv preprint arXiv:1701.08105*, 2017.

R. M. Dudley. *Real analysis and probability*, volume 74. Revised reprint of the 1989 original. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge., 2002.

P. J. Forrester. Exact integral formulas and asymptotics for the correlations in the $1/r^2$ quantum many-body system. *Phys. Lett. A*, 179(2):127–130, 1993.

P. J. Forrester. *Log-gases and random matrices (LMS-34)*. Princeton University Press, 2010.

S. Friedli and Y. Velenik. *Statistical mechanics of lattice systems: a concrete mathematical introduction.* Cambridge University Press, 2017.

H.-O. Georgii. *Canonical Gibbs measures*, volume 760 of *Lecture Notes in Mathematics.* Springer, Berlin, 1979. ISBN 3-540-09712-0. Some extensions of de Finetti's representation theorem for interacting particle systems.

H.-O. Georgii. *Gibbs measures and phase transitions*, volume 9 of *De Gruyter Studies in Mathematics.* Walter de Gruyter & Co., Berlin, second edition, 2011.

H.-O. Georgii and H. Zessin. Large deviations and the maximum entropy principle for marked point random fields. *Probab. Th. Rel. Fields*, 96(2):177–204, 1993.

S. Ghosh. Determinantal processes and completeness of random exponentials: the critical case. *Probab. Th. Rel. Fields*, 163(3-4):643–665, 2015.

S. Ghosh and J.L. Lebowitz. Fluctuations, large deviations and rigidity in hyperuniform systems: A brief survey. *Indian J. Pure Appl. Math.*, 48(4):609–631, 2017.

S. Ghosh and Y. Peres. Rigidity and tolerance in point processes: Gaussian zeros and Ginibre eigenvalues. *Duke Math. J.*, 166(10):1789–1858, 2017.

D. Holcomb and E. Paquette. The maximum deviation of the Sine$_\beta$ counting process. *Preprint. ArXiv:1801.08989*, 2018.

D. Holcomb and B. Valkó. Large deviations for the Sine$_\beta$ and Sch$_\tau$ processes. *Probab. Th. Rel. Fields*, 163(1-2):339–378, 2015.

D. Holcomb and B. Valkó. Overcrowding asymptotics for the Sine$_\beta$ process. *Ann. Inst. H. Poincaré Probab. Statist.*, 53(3):1181–1195, 2017.

T. Jiang and S. Matsumoto. Moments of traces of circular beta-ensembles. *Ann. Probab.*, 43(6):3279–3336, 2015.

O. Kallenberg. *Random Measures.* Berlin, Akademie-Verlag, 3rd edition edition, 1983.

R. Killip and M. Stoiciu. Eigenvalue statistics for CMV matrices: from Poisson to clock via random matrix ensembles. *Duke Math. J.*, 146(3):361–399, 2009.

O. K. Kozlov. A Gibbsian description of point random fields. *Teor. Verojatnost. i Primenen.*, 21(2):348–365, 1976. ISSN 0040-361x.

E. Kritchevski, B. Valkó, and B. Virág. The scaling limit of the critical one-dimensional random Schrödinger operator. *Comm. Math. Phys.*, 314:1–32, 2012.

A. Kuijlaars and E. Miña-Díaz. Universality for conditional measures of the Sine point process. *arXiv preprint arXiv:1703.02349*, 2017.

T. Leblé. CLT for fluctuations of linear statistics in the Sine-beta process. *arXiv preprint arXiv:1809.03448*, 2018.

T. Leblé and S. Serfaty. Large deviation principle for empirical fields of Log and Riesz gases. *Invent. math.*, 210(3):645–757, 2017.

Jesper Moller and Rasmus Plenge Waagepetersen. *Statistical inference and simulation for spatial point processes.* Chapman and Hall/CRC, 2003.

F. Nakano. Level statistics for one-dimensional Schrödinger operators and Gaussian beta ensemble. *J. Stat. Phys.*, 156(1):66–93, 2014.

X.-X. Nguyen and H. Zessin. Integral and differential characterizations of the Gibbs process. *Math. Nachr.*, 88:105–115, 1979.

M. Petrache and S. Serfaty. Next order asymptotics and renormalized energy for Riesz interactions. *J. Inst. Math. Jussieu*, 16(3):501–569, 2017.

C. Preston. *Random fields.* Lecture Notes in Mathematics, Vol. 534. Springer-Verlag, Berlin-New York, 1976.

E. Sandier and S. Serfaty. From the Ginzburg-Landau model to vortex lattice problems. *Comm. Math. Phys.*, 313(3):635–743, 2012.

M. Shcherbina. Change of variables as a method to study general β-models: bulk universality. *J. Math. Phy.*, 55(4):043504, 2014.

B. Valkó and B. Virág. Continuum limits of random matrices and the Brownian carousel. *Invent. math.*, 177(3):463–508, 2009.

B. Valkó and B. Virág. Large gaps between random eigenvalues. *Ann. Probab.*, 38(3):1263–1279, 2010.

B. Valkó and B. Virág. The Sine$_\beta$ operator. *Invent. math.*, 209(1):275–327, 2017.

B. Valkó and B. Virág. Operator limit of the circular beta ensemble. *Preprint. ArXiv:1710.06988*, 2018.

M. Venker. Particle systems with repulsion exponent β and random matrices. *Electron. Commun. Probab.*, 18(83):1–12, 2013.

A. Wakolbinger and G. Eder. A condition Σ_λ^c for point processes. *Math. Nachr.*, 116: 209–232, 1984.