Ergod. Th. & Dynam. Sys., (2023), **43**, 1847–1908 © The Author(s), 2022. Published by Cambridge University Press.
doi:10.1017/etds.2022.6

A complete proof that square ice entropy is $\frac{3}{2}\log_2(4/3)$

SILVÈRE GANGLOFF

Faculty of Applied Mathematics of AGH, Mickiewicza, A-3/A-4, Krakow, Poland
(e-mail: sgangloff@agh.edu.pl)

(*Received* 28 *May* 2019 *and accepted in revised form* 11 *January* 2022)

Abstract. In this text, we provide a fully rigorous and complete proof of E.H. Lieb's statement that (topological) entropy of square ice (or six-vertex model, XXZ spin chain for anisotropy parameter $\Delta = 1/2$) is equal to $\frac{3}{2}\log_2(4/3)$.

Key words: entropy, square ice, six-vertex model
2020 Mathematics Subject Classification: 37A35 (Primary); 82B20 (Secondary)

1. *Introduction*

1.1. *Computing entropy of multidimensional subshifts of finite type.* This work is the consequence of a renewal of interest from the fields of symbolic dynamics to entropy computation methods developed in quantum and statistical physics for lattice models. This interest comes from *constructive methods* for multidimensional subshifts of finite type (some equivalent formulation in symbolic dynamics of lattice models) that are involved in the characterization by Hochman and Meyerovitch [**HM10**] of the possible values of topological entropy for these dynamical systems (where the dynamics are provided by the action of the $\mathbb{Z}^2$ shift action) with a recursion-theoretic criterion. The consequences of this theorem are not only that entropy may be algorithmically uncomputable for a multidimensional subshift of finite type (SFT), which was previously proved for cellular automata [**HKC92**], but also strong evidence that the study of these systems as a class is intertwined with computability theory. Moreover, it is an important tool to localize sub-classes for which the entropy is computable in a uniform way, as ones defined by strong dynamical constraints [**PS15**]. Some current research attempts have been made to understand the *frontier between the uncomputability and the computability* of entropy for a multidimensional SFT. For instance, approaching the frontier from the uncomputable domain, the author, together with Sablik [**GS17**], proved that the characterization of Hochman and Meyerovitch stands under a relaxed form of the constraint studied in [**PS15**], which includes notably all exactly solvable models considered in statistical and quantum physics. To approach the frontier from the computable domain, it is natural to attempt to understand (in particular, prove) and extend the computation methods developed for these models.

1.2. *Content of this text.* Our study in the present text focuses on square ice (or equivalently, the six-vertex model or the XXZ spin chain for anisotropy parameter $\Delta = 1/2$). Since it is central among exactly solvable models in quantum physics [**B82**], this work will serve as a ground for further connections between entropy computation methods and constructive methods coming from symbolic dynamics. The entropy of square ice was argued by Lieb [**L67**] to be exactly $\frac{3}{2}\log_2(\frac{4}{3})$. However, his proof was not complete, as it relied on a non-verified hypothesis (the condensation of Bethe roots, see §6). Moreover, various other arguments involved in Lieb's argumentation and later developments have not yet received full rigorous treatment. In this text, we fill the holes and propose a (complete) proof of the following theorem.

THEOREM 1. *The entropy of square ice is equal to* $\frac{3}{2}\log_2(\frac{4}{3})$.

For completeness, we include some exposition of what can be considered as background material. The proof is thus self-contained, except for the use of the coordinate Bethe ansatz, for which we rely on another paper by Duminil-Copin *et al.* [**DGHMT18**].

One can find an overview of the proof in §3, presented after some definitions related to symbolic dynamics and representations of square ice in §2.

2. *Background: square ice and its entropy*

2.1. *Subshifts of finite type*

2.1.1. *Definitions.* Let $\mathcal{A}$ be some finite set, called the *alphabet*. For all $d \geq 1$, the set $\mathcal{A}^{\mathbb{Z}^d}$, whose elements are called *configurations*, is a topological space with the infinite power of the discrete topology on $\mathcal{A}$. Let us denote by σ the *shift* action of $\mathbb{Z}^d$ on this space defined by the following equality for all $\mathbf{u} \in \mathbb{Z}^d$ and x element of the space: $(\sigma^{\mathbf{u}}(x))_{\mathbf{v}} = x_{\mathbf{v}+\mathbf{u}}$. A compact subset X of this space is called a d-dimensional *subshift* when this subset is stable under the action of the shift, which means that for all $\mathbf{u} \in \mathbb{Z}^d$, $\sigma^{\mathbf{u}}(X) \subset X$. For any finite subset $\mathbb{U}$ of $\mathbb{Z}^d$, an element p of $\mathcal{A}^{\mathbb{U}}$ is called a *pattern* on the alphabet $\mathcal{A}$ and on *support* $\mathbb{U}$. We say that this pattern *appears* in a configuration x when there exists a translate $\mathbb{V}$ of $\mathbb{U}$ such that $x_{\mathbb{V}} = p$. We say that it appears in another pattern q on a support containing $\mathbb{U}$ such that the restriction of q on $\mathbb{U}$ is p. We say that it appears in a subshift X when it appears in a configuration of X. Such a pattern is also called *globally admissible* for X. For all $d \geq 1$, the number of patterns on support $\mathbb{U}_N^{(d)} \equiv [\![1, N]\!]^d$ that appear in a d-dimensional subshift X is denoted by $\mathcal{N}_N(X)$. When $d = 2$, the number of patterns on support $\mathbb{U}_{M,N}^{(2)} \equiv [\![1, M]\!] \times [\![1, N]\!]$ that appear in X is denoted by $\mathcal{N}_{M,N}(X)$. A d-dimensional subshift X defined by forbidding patterns in some finite set $\mathcal{F}$ to appear in the configurations, formally

$$X = \{x \in \mathcal{A}^{\mathbb{Z}^d} : \text{for all } \mathbb{U} \subset \mathbb{Z}^d, x_{\mathbb{U}} \notin \mathcal{F}\},$$

is called a subshift of *finite type* (SFT). In a context where the set of forbidden patterns defining the SFT is fixed, a pattern is called *locally admissible* for this SFT when no

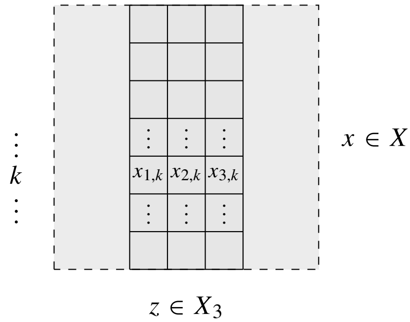


FIGURE 1. Illustration of Definition 2 for $N = 3$.

forbidden pattern appears in it. A *morphism* between two $\mathbb{Z}^d$-subshifts X, Z is a continuous map $\varphi : X \to Z$ such that $\varphi \circ \sigma^{\mathbf{v}} = \sigma^{\mathbf{v}} \circ \varphi$ for all $\mathbf{v} \in \mathbb{Z}^d$ (the map commutes with the shift action). An *isomorphism* is an invertible morphism.

2.1.2. *Topological entropy*

Definition 1. Let X be a d-dimensional subshift. The *topological entropy* of X is defined as

$$h(X) \equiv \inf_{N\geq 1} \frac{\log_2(\mathcal{N}_N(X))}{N^d}.$$

It is a well-known fact in topological dynamics that this infimum is a limit:

$$h(X) = \lim_{N\geq 1} \frac{\log_2(\mathcal{N}_N(X))}{N^d}.$$

It is a topological invariant, meaning that when there is an isomorphism between two subshifts, these two subshifts have the same entropy [**LM95**].

Definition 2. Let X be a bidimensional subshift ($d = 2$). For all $n \geq 1$, we denote by X_N the subshift obtained from X by restricting to the width N infinite strip $\{1, \ldots, N\} \times \mathbb{Z}$. Formally, this subshift is defined on alphabet $\mathcal{A}^N$ and by that $z \in X_N$ if and only if there exists $x \in X$ such that for all $k \in \mathbb{Z}$, $z_k = (x_{1,k}, \ldots, x_{N,k})$. See Figure 1.

In the following, we will use the following proposition.

PROPOSITION 1. *The entropy of X can be computed through the sequence* $(h(X_N))_N$*:*

$$h(X) = \lim_N \frac{h(X_N)}{N}.$$

We include a proof of this statement, for completeness.

Proof. From the definition of X_N,

$$h(X) = \lim_N \lim_M \frac{\log_2(\mathcal{N}_{M,N}(X))}{NM}.$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

We prove this via an upper bound on the $\limsup_N$ and a lower bound on the $\liminf_N$ of the sequence in this formula.

Upper bound by decomposing squares into rectangles. Since for any M, N, k, the set $\mathbb{U}^{(2)}_{kM,kN}$ is the union of MN translates of $\mathbb{U}^{(2)}_k$, a pattern on support $\mathbb{U}^{(2)}_{kM,kN}$ can be seen as an array of patterns on $\mathbb{U}^{(2)}_k$. As a consequence,

$$\mathcal{N}_{kM,kN}(X) \leq (\mathcal{N}_{k,k}(X))^{MN},$$

and using this inequality, we get

$$\begin{aligned}\lim_M \frac{\log_2(\mathcal{N}_{M,N}(X))}{NM} &= \lim_M \frac{\log_2(\mathcal{N}_{kM,kN}(X))}{k^2NM} \\ &\leq \lim_M \frac{\log_2(\mathcal{N}_{k,k}(X))}{k^2} = \frac{\log_2(\mathcal{N}_{k,k}(X))}{k^2}.\end{aligned}$$

As a consequence, for all k,

$$\limsup_N \lim_M \frac{\log_2(\mathcal{N}_{M,N}(X))}{NM} \leq \frac{\log_2(\mathcal{N}_{k,k}(X))}{k^2},$$

and this implies

$$\limsup_N \lim_M \frac{\log_2(\mathcal{N}_{M,N}(X))}{NM} \leq h(X), \tag{1}$$

by taking $k \to +\infty$ in the last inequality.

Lower bound by decomposing rectangles into squares. For all M, N, by considering a pattern on $\mathbb{U}^{(2)}_{MN,NM}$ as an array of patterns on $\mathbb{U}^{(2)}_{M,N}$, we get that

$$\mathcal{N}_{MN,NM}(X) \leq (\mathcal{N}_{M,N}(X))^{MN}.$$

Thus,

$$h(X) = \lim_M \frac{\log_2(\mathcal{N}_{MN,NM}(X))}{M^2N^2} \leq \lim_M \frac{\log_2(\mathcal{N}_{M,N}(X))}{MN}.$$

As a consequence,

$$h(X) \leq \liminf_N \lim_M \frac{\log_2(\mathcal{N}_{M,N}(X))}{NM}. \tag{2}$$

The two inequalities in equations (1) and (2) imply that the sequence $(h(X_N)/N)_N$ converges and that the limit is $h(X)$. □

In the following, for all N and M, we assimilate patterns of X_N on $\mathbb{U}^{(1)}_M$ with patterns of X on $\mathbb{U}^{(2)}_{M,N}$.

2.2. *Representations of square ice.* The square ice can be defined as an isomorphic class of subshifts of finite type, whose elements can be thought of as various representations of the same object. The most widely used is the six-vertex model (whose name derives from the fact that the elements of the alphabet represent vertices of a regular grid) and is presented in §2.2.1. In this text, we will use another representation, presented in §2.2.2, whose configurations consist of drifting discrete curves, representing possible *particle*

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

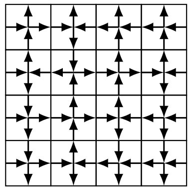
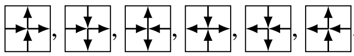
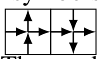
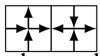


FIGURE 2. An example of locally and thus globally admissible pattern of the six-vertex model.

trajectories. In §2.2.3, we provide a proof that one can restrict to a subset of the total set of patterns considered to compute entropy of square ice.

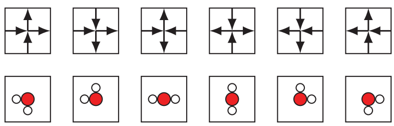
2.2.1. *The six-vertex model.* The *six-vertex model* is the subshift of finite type described as follows.

Symbols: 

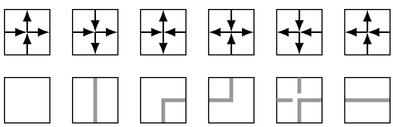
Local rules: Considering two adjacent positions in $\mathbb{Z}^2$, the arrows corresponding to the common edge of the symbols on the two positions have to be directed the same way. For instance, the pattern  is allowed, while  is not.

Global behavior: The symbols draw a lattice whose edges are oriented in such a way that all the vertices have two incoming arrows and two outgoing ones. This is called an Eulerian orientation of the square lattice. See an example of an admissible pattern in Figure 2.

Remark 1. The name of square ice of the considered class of SFT appears clearly when considering the following application on the alphabet of the six-vertex model to local configurations of dihydrogen monoxide:



2.2.2. *Drifting discrete curves.* From the six-vertex model, we derive another representation of square ice through an isomorphism, which consists in transforming the letters via an application π_s on the alphabet of the six-vertex model, described as follows:



For instance, the pattern in Figure 2 corresponds, after application of π_s, to that in Figure 3. In this SFT, the local rules consist in forcing that any segment of the curve in a symbol extends in the positions that it directs to (in the fifth symbol in the above list, we consider that there are only two segments).

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

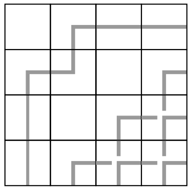


FIGURE 3. Representation of pattern in Figure 2.

In the following, we denote by X^s this SFT.

Remark 2. One can see straightforwardly that locally admissible patterns of this SFT are always globally admissible, since any locally admissible pattern can be extended into a configuration by extending the curves in a straight way.

2.2.3. *Entropy of X^s and cylindrical stripes subshifts of square ice.* Consider some alphabet $\mathcal{A}$, and X a bidimensional subshift of finite type on this alphabet. For all $N \geq 1$, we set $\Pi_N = \mathbb{Z}/(N\mathbb{Z}) \times \mathbb{Z}$. Let us also denote by $\pi_N : [\![1, N]\!] \times \mathbb{Z} \to \Pi_N$ and $\phi_N : \mathcal{A}^{[\![1,N]\!]\times\mathbb{Z}} \to \mathcal{A}^{\Pi_N}$ the canonical projections. Formally, for all $\mathbf{u} \in [\![1, N]\!] \times \mathbb{Z}$ and $x \in \mathcal{A}^{[\![1,N]\!]\times\mathbb{Z}}$,

$$(\phi_N(x))_{\pi_N(\mathbf{u})} = x_{\mathbf{u}}.$$

We say that a pattern p on support $\mathbb{U} \subset [\![1, N]\!] \times \mathbb{Z}$ appears in a configuration $\overline{x}$ on Π_N when there exists a configuration in X_N whose image by π_N is $\overline{x}$, and there exists an element $\mathbf{u} \in \Pi_N$ such that for all $\mathbf{v} \in \mathbb{U}$, $\overline{x}_{\mathbf{u}+\pi_N(\mathbf{v})} = x_{\mathbf{v}}$.

Notation 1. Let us denote by $\overline{X}_N$ the set of configurations in X_N whose image by ϕ_N does not contain any forbidden pattern for X (in other words, this pattern can be wrapped on an infinite cylinder without breaking the rules defining X).

Similarly, we call (M, N)-*cylindrical pattern* of X a pattern on $\mathbb{U}_{M,N}$ that can be wrapped on a finite cylinder $\mathbb{Z}/N\mathbb{Z} \times \{1, \ldots, M\}$. Let us prove a preliminary result on the entropy of square ice, which relates the entropy of X^s to the sequence $(h(\overline{X}^s_N))_N$.

LEMMA 1. *The subshift X^s has entropy equal to*

$$h(X^s) = \lim_N \frac{h(\overline{X}^s_N)}{N}.$$

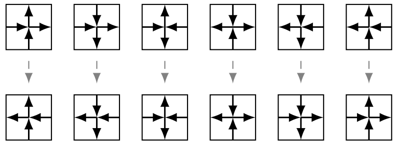
Remark 3. To prove this lemma, we use a technique that first appeared in a work of Friedland [**F97**], which relies on a symmetry of the alphabet and rules of the SFT.

Proof. (1) *Lower bound:* Since for all N, $\overline{X}^s_N \subset X^s_N$, then $h(\overline{X}^s_N) \leq h(X^s_N)$. We deduce by Proposition 1 that

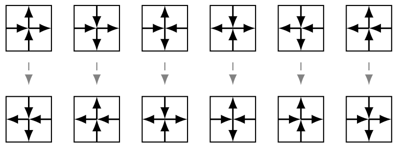
$$\limsup_N \frac{h(\overline{X}^s_N)}{N} \leq h(X^s).$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

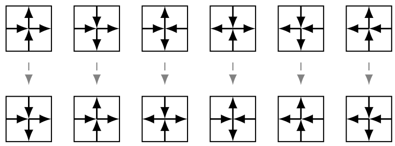
(2) *Upper bound:* Consider the transformation τ on the six-vertex model alphabet that consists in a horizontal symmetry of the symbols and then the inversion of all the arrows. The symmetry can be represented as follows:



The inversion is represented by



As a consequence, τ is



We define then a horizontal symmetry operation $\mathcal{T}_N$ (see Figure 4 for an illustration) on patterns p whose support is some $\mathbb{U}^{(2)}_{M,N}$, with $M \geq 1$.

For all such M and p, $\mathcal{T}_M(p)$ has also support $\mathbb{U}^{(2)}_{M,N}$ and for all $(i, j) \in \mathbb{U}^{(2)}_{M,N}$,

$$\mathcal{T}_N(p)_{i,j} = \tau(p_{N-i,j}).$$

We define also the applications ∂^r_N (respectively, ∂^l_N, ∂^t_N) that acts on patterns of the six-vertex model whose support is some $\mathbb{U}^{(2)}_{M,N}$, $M \geq 1$ and such that for all $M \geq 1$ and p on support $\mathbb{U}^{(2)}_{M,N}$, $\partial^r_N(p)$ (respectively, $\partial^l_N(p)$) is a length M (respectively, M, N) word and for all j between 1 and M (respectively, M), $\partial^r_N(p)_j$ (respectively, $\partial^l_N(p)_j$) is the east (respectively, west) arrow in the symbol $p_{N,j}$ (respectively, $p_{1,j}$). For instance, if p is the pattern on the left in Figure 4, then $\partial^r_N(p)$ (respectively, $\partial^l_N(p)$) is the word:

$$\leftarrow\leftarrow\rightarrow\leftarrow \quad \text{(respectively, } \rightarrow\rightarrow\rightarrow\rightarrow).$$

For the purpose of notation, we denote also by π_s the application that transforms patterns of the six-vertex model into patterns of X_s via the application of π_s letter by letter. Let us consider the transformation $\mathcal{T}^s_N \equiv \pi_s \circ \mathcal{T}_N \circ \pi_s^{-1}$ on patterns of X^s on some $\mathbb{U}^{(2)}_{M,N}$. We also denote by $\partial^{l,s}_N \equiv \partial^l_N \circ \pi_s^{-1}$, $\partial^{r,s}_N \equiv \partial^r_N \circ \pi_s^{-1}$. Let us prove some properties of these transformations. For any word $\mathbf{w}$ on the alphabet $\{\leftarrow, \rightarrow\}$ or $\{\uparrow, \downarrow\}$, we denote by $\overline{\mathbf{w}}$ the word obtained by exchanging the two letters in the word $\mathbf{w}$.

(a) *Preservation of global admissibility:* For any p globally admissible, $\mathcal{T}_N(p)$ is also locally admissible, and as a consequence globally admissible; indeed, it is sufficient to check that for all u, v in the alphabet, if uv is not a forbidden pattern in the

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

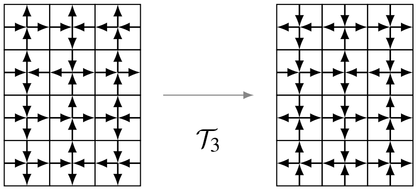


FIGURE 4. Illustration of the definition of $\mathcal{T}_3$: the pattern on the left (on support $\mathbb{U}^{(2)}_{3,4}$) is transformed into the pattern on the right via this transformation.

six-vertex model, then $\tau(v)\tau(u)$ is also not a forbidden pattern and that if $\substack{u\\v}$ is not forbidden, then $\substack{\tau(u)\\\tau(v)}$ is also not forbidden.

The first assertion is verified because uv is not forbidden if and only if the arrows of these symbols attached to their adjacent edge are pointing in the same direction, and this property is preserved when changing uv into $\tau(v)\tau(u)$. The second one is verified for a similar reason.

(b) *Gluing patterns:* Let us consider any $N, M \geq 1$ and p, p' two patterns of X^s on support $\mathbb{U}^{(2)}_{M,N}$, such that $\partial^{r,s}_N(p) = \partial^{r,s}_N(p')$ and $\partial^{l,s}_N(p) = \partial^{l,s}_N(p')$. Let us denote by pattern p'' on support $\mathbb{U}^{(2)}_{M,2N}$ such that the restriction of p'' on $\mathbb{U}^{(2)}_{M,N}$ is p and the restriction on $(0, N) + \mathbb{U}^{(2)}_{M,N}$ is $\mathcal{T}_N(p')$.

- *This pattern is admissible* (locally and thus globally). Indeed, this is sufficient to check that gluing the two patterns p and p' does not make forbidden patterns appear, and this comes from that for all letter u, $u\tau(u)$ is not forbidden. This can be checked directly, letter by letter.
- *Moreover, p'' is in $\mathcal{N}_M(\overline{X}_{2N})$.* Indeed, this pattern can be *wrapped on a cylinder*, and this comes from the fact that if u is a symbol of the six-vertex model, $\tau(u)u$ is not forbidden.

(3) *From the gluing property to an upper bound:* Given $\mathbf{w} = (\mathbf{w}^l, \mathbf{w}^r)$ as some pair of words on $\{\rightarrow, \leftarrow\}$, we denote by $\mathcal{N}^{\mathbf{w}}_{M,N}$ the number of patterns of X^s on support $\mathbb{U}^{(2)}_{M,N}$ such that $\partial^{l,s}_N = \mathbf{w}^l$ and $\partial^{r,s}_N = \mathbf{w}^r$. Since $\mathcal{T}_N$ is a bijection, denoting $\overline{\mathbf{w}} = (\overline{\mathbf{w}^l}, \overline{\mathbf{w}^r})$, we have

$$\mathcal{N}^{\mathbf{w}}_{M,N} = \mathcal{N}^{\overline{\mathbf{w}}}_{M,N}.$$

From the last point, for all $\mathbf{w}$,

$$\mathcal{N}_M(\overline{X}^s_{2N}) \geq \mathcal{N}^{\mathbf{w}}_{M,N}.\mathcal{N}^{\overline{\mathbf{w}}}_{M,N} = (\mathcal{N}^{\mathbf{w}}_{M,N})^2$$

$$(\mathcal{N}_M(\overline{X}^s_{2N}))^{1/2} \geq \mathcal{N}^{\mathbf{w}}_{M,N}.$$

By summing over all possible $\mathbf{w}$:

$$2^{2M} \cdot (\mathcal{N}_M(\overline{X}^s_{2N}))^{1/2} = \sum_{\mathbf{w}} (\mathcal{N}_M(\overline{X}^s_{2N}))^{1/2} \geq \sum_{\mathbf{w}} \mathcal{N}^{\mathbf{w}}_{M,N} = \mathcal{N}_{M,N}(X^s).$$

As a consequence, for all N,

$$2 + \frac{1}{2} h(\overline{X}^s_{2N}) \geq h(X^s_N).$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

This implies that

$$\liminf_N \frac{h(\overline{X}^s_{2N})}{2N} \geq \liminf_N \frac{h(X^s_N)}{N} = h(X^s).$$

For similar reasons,

$$\liminf_N \frac{h(\overline{X}^s_{2N+1})}{2N+1} \geq \liminf_N \frac{h(X^s_N)}{N} = h(X^s),$$

and thus,

$$\liminf_N \frac{h(\overline{X}^s_{N})}{N} \geq \liminf_N \frac{h(X^s_N)}{N} = h(X^s).$$

□

3. *Overview of the proof*

In the following, we provide a complete proof of the following theorem.

THEOREM 2. *The entropy of square ice is equal to*

$$h(X^s) = \frac{3}{2}\log_2\left(\frac{4}{3}\right).$$

The proof of Theorem 2 can be summarized as follows. Some of the terms will be defined in the text; however, this overview will provide us with a way in which to situate every argument in the overall strategy, which consists of:

(1) finding a formula for $h(\overline{X}^s_N)$ for all N (in practice for all N odd);

(2) then using Lemma 1 to compute $h(X^s)$.

- The first point is done using the *transfer matrix* method, which allows us to express $h(\overline{X}^s_N)$ with a formula involving a sequence of numbers defined implicitly through a system of nonlinear equations called *Bethe equations*. This method itself consists of several steps.
  (1) *Formulation with transfer matrices [§4]:* it is usual, when dealing with unidimensional subshifts of finite type, to express their entropy as the greatest eigenvalue of the adjacency matrix, which tells which couples of symbols (which are rows of symbols in the case of stripe subshifts) can be adjacent. In this text, we use the adjacency matrix V^*_N of a subshift which is isomorphic to $\overline{X}_N$, and see the matrix of a linear operator on $\Omega_N = \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$.
  (2) *Lieb path—transport of information through analyticity [§4]:* In quantum physics, *transfer matrices*, which are complexifications of the adjacency matrices in a local way (in the sense that the coefficient relative to a couple of rows is the product of some coefficients in $\mathbb{C}$ relative to the symbols in the two rows) are used to derive properties of the system. In this text, we will see the adjacency matrix as a particular value of an analytic path of such transfer matrices, $t \in \mathbb{R} \mapsto V_N(t)$ such that for all t, $V_N(t)$ is an irreducible non-negative and symmetric matrix, and such that $V_N(1) = V^*_N$—we will call such a path a *Lieb path* in the following. The analyticity is used here to gain some information on the whole path, including on $V_N(1)$, from information on a segment of the path. This part is contained in §4 and is a detailed exposition of notions defined in the article of Lieb [**L67**].

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

(3) *Coordinate Bethe ansatz [§5]:* We use the coordinate Bethe ansatz (due originally to Hans Bethe and exposed recently in [**DGHMT18**] and related in the present text), which consists in a clever guess on the form of potential eigenvectors, and actually provides some candidates, for the matrices $V_N(t)$. In practice, we apply this on each of the subspaces $\Omega_N^{(n)}$ of a decomposition of Ω_N:

$$\Omega_N = \bigoplus_{n=0}^{N} \Omega_N^{(n)}.$$

The candidate eigenvectors and eigenvalues depend each on a solution $(p_j)_{j=1\ldots n}$ of a nonlinear system of equations, with parameter t, called Bethe equations.

It is shown that the system of Bethe equations admits a unique solution for each n,N and t, which we denote by $(\mathbf{p}_j(t))_j$ for all $t \in (0, \sqrt{2})$, in a context where n, N are fixed, using convexity arguments on an auxiliary function. The analyticity of the Lieb path and the convexity of the auxiliary function ensure together that $t \mapsto (\mathbf{p}_j(t))_j$ is analytic. This part completes the proof of an argument left incomplete in [**YY66a**]. To identify the greatest eigenvalue of $V_N(t)$ for all t, we use the fact that $V_N(\sqrt{2})$ commutes with some Hamiltonian H_N that is completely diagonalized (following [**LSM61**]). The consequence of this fact is that $V_N(\sqrt{2})$ and H_N have a common base of eigenvectors. The candidate eigenvector provided by the Bethe ansatz on this point is proved not to be null and associated to the maximal eigenvalue of H_N on $\Omega_N^{(n)}$. By Perron–Frobenius theorem, the vector has positive coordinates, and by the same theorem (uniqueness part), it is an eigenvector of $V_N(\sqrt{2})$ and is associated with the maximal eigenvalue of this matrix on $\Omega_N^{(n)}$. By continuity, this is true also for t in a neighborhood of $\sqrt{2}$, and by analyticity, this identity is true for all $t \in (0, \sqrt{2})$.

- The second point is derived in two steps:

(1) *Asymptotic condensation of Bethe roots [§6]:* The sequences $(\mathbf{p}_j(t))_j$ are transformed into sequences $(\boldsymbol{\alpha}_j(t))_j$ through an analytic bijection. The values of these second sequences are in $\mathbb{R}$ and are called *Bethe roots*.

We first prove that the sequences of Bethe roots are condensed according to a density function ρ_t over $\mathbb{R}$, relative to any continuous decreasing and integrable function $f : (0, +\infty) \to (0, +\infty)$, which means that the Cesàro mean of the finite sequence $(f(\boldsymbol{\alpha}_j(t)))_j$ converges towards $\int \rho_t(x) f(x)\, dx$. This part involves rigorous proofs, some simplifications and adaptations of arguments that appeared in [**K18**]. The density ρ_t is defined as the solution of a Fredholm integral equation which can be thought of as the asymptotic version of Bethe equations. This equation is solved through Fourier analysis, following a computation done in [**YY66b**].

(2) *Computation of integrals [§7]:* The condensation property proved in the last point implies that the formula obtained for $\frac{1}{N} h(\overline{X}_N^s)$ converges to an integral involving ρ_1. The formula obtained for ρ_1 allows the computation of this integral, via loop integrals techniques. This part is a detailed version of computations exposed in [**L67**].

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

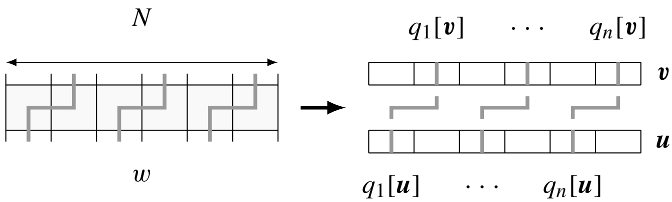


FIGURE 5. Illustration for the definition of the notation $\boldsymbol{u}\mathcal{R}[w]\boldsymbol{v}$.

For the purpose of clarity, we will begin each section with a paragraph beginning with ▹ relating the section to this overview.

4. *A Lieb path for square ice*

▹ In this section, we define the matrices V_N^* [§4.1], the Lieb paths $t \mapsto V_N(t)$ that we will use in the following [§4.2] and prove that $V_N(1) = V_N^*$ [§4.3].

4.1. *The interlacing relation and the matrices V_N^*.*

▹ The definition of the matrix V_N^* relies on a relation between words on $\{0, 1\}$ having length N. In this section, we prove the properties of this relation which will be translated later into properties of the matrices $V_N(t)$, and in particular V_N^*.

In the following, for a square matrix M, we will denote by $M[u, v]$ its entry on (u, v). Moreover, we denote by $\{0, 1\}_N^*$ the set of length N words on $\{0, 1\}$.

Notation 2. Consider $\mathbf{u}$, $\mathbf{v}$ two words in $\{0, 1\}_N^*$, and w some $(N, 1)$-cylindrical pattern of the subshift X. We say that the pattern w *connects* $\mathbf{u}$ to $\mathbf{v}$, and we denote this by $\mathbf{u}\mathcal{R}[w]\mathbf{v}$, when for all $k \in [\![1, N]\!]$, $\mathbf{u}_k = 1$ (respectively, $\mathbf{v}_k = 1$) if and only if w has an incoming (respectively, outgoing) curve on the bottom (respectively, top) of its kth symbol. This notation is illustrated in Figure 5.

Definition 3. Let us denote by $\mathcal{R} \subset \{0, 1\}^N \times \{0, 1\}^N$ the relation defined by $\boldsymbol{u}\mathcal{R}\boldsymbol{v}$ if and only if there exists an $(N, 1)$-cylindrical pattern w of the discrete curves shift X^s such that $\boldsymbol{u}\mathcal{R}[w]\boldsymbol{v}$.

Notation 3. Let $N \geq 1$ be an integer and $t > 0$. Let us denote by Ω_N the space $\mathbb{C}^2 \bigotimes \cdots \bigotimes \mathbb{C}^2$, the tensor product of N copies of $\mathbb{C}^2$, whose canonical basis elements are denoted indifferently by $\boldsymbol{\epsilon} = |\boldsymbol{\epsilon}_1 \cdots \boldsymbol{\epsilon}_N\rangle$ or the words $\boldsymbol{\epsilon}_1 \cdots \boldsymbol{\epsilon}_N$, for $(\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_N) \in \{0, 1\}^N$, according to quantum mechanics notation, to distinguish them from the coordinate definition of vectors of Ω_N. For the definition of the matrices, we order the elements of this basis with the lexicographic order.

Definition 4. Let us define $V_N^* \in \mathcal{M}_{2^N}(\mathbb{C})$ the matrix such that for all $\boldsymbol{\epsilon}, \boldsymbol{\eta} \in 0, 1_N^*$, $V_N[\boldsymbol{\epsilon}, \boldsymbol{\eta}]$ is the number of w such that $\boldsymbol{\epsilon}\mathcal{R}[w]\boldsymbol{\eta}$.

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

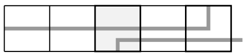


FIGURE 6. Illustration of (impossible) crossing situation, which would imply non-authorized symbols.

Notation 4. For all $\mathbf{u} \in \{0, 1\}_N^*$, we denote by $|\mathbf{u}|_1$ the number of $k \in [\![1, N]\!]$ such that $\mathbf{u}_k = 1$. If $|\mathbf{u}|_1 = n$, we denote by $q_1[\mathbf{u}] < \cdots < q_n[\mathbf{u}]$ the integers such that $\mathbf{u}_k = 1$ if and only if $k = q_i[\mathbf{u}]$ for some $i \in [\![1, n]\!]$.

Let us also notice that $\mathbf{u}\mathcal{R}\mathbf{v}$ implies that the number of 1 symbols in $\mathbf{u}$ is equal to the number of 1 symbols in $\mathbf{v}$.

Definition 5. We say that two words $\mathbf{u}, \mathbf{v}$ in $\{0, 1\}_N^*$ such that $|\mathbf{u}|_1 = |\mathbf{v}|_1 \equiv n$ are *interlaced* when one of the two following conditions is satisfied:

$$q_1[\boldsymbol{u}] \leq q_1[\boldsymbol{v}] \leq q_2[\boldsymbol{u}] \leq \cdots \leq q_n[\boldsymbol{u}] \leq q_n[\boldsymbol{v}],$$

$$q_1[\boldsymbol{v}] \leq q_1[\boldsymbol{u}] \leq q_2[\boldsymbol{v}] \leq \cdots \leq q_n[\boldsymbol{v}] \leq q_n[\boldsymbol{u}].$$

PROPOSITION 2. *For two length N words $\boldsymbol{u}, \boldsymbol{v}$, we have $\boldsymbol{u}\mathcal{R}\boldsymbol{v}$ if and only if $|\boldsymbol{u}|_1 = |\boldsymbol{v}|_1 \equiv n$ and $\boldsymbol{u}, \boldsymbol{v}$ are interlaced.*

Proof. ($\Rightarrow$): assume that $\boldsymbol{u}\mathcal{R}[w]\boldsymbol{v}$ for some w.

First, since w is an $(N, 1)$-cylindrical pattern, each of the curves that cross its bottom side also crosses its top side, which implies that $|\boldsymbol{u}|_1 = |\boldsymbol{v}|_1$.

We assume that $q_1[\boldsymbol{u}] \leq q_1[\boldsymbol{v}]$ (the other case is processed similarly).

(1) *The position $q_1[\boldsymbol{u}]$ is connected to $q_1[\boldsymbol{v}]$ or $q_1[\boldsymbol{u}] = q_1[\boldsymbol{v}]$*: Let us assume that $q_1[\boldsymbol{u}] \neq q_1[\boldsymbol{v}]$ and $q_1[\boldsymbol{u}]$ is not connected to $q_1[\boldsymbol{v}]$. Then because $\boldsymbol{u}\mathcal{R}[w]\boldsymbol{v}$, another curve would have to connect another position $q_k[\boldsymbol{u}]$, $k \neq 1$ of $\boldsymbol{u}$ to $q_1[\boldsymbol{v}]$. Since $q_k[\boldsymbol{u}] > q_1[\boldsymbol{u}]$ (by definition), this curve would cross the left border of w. It would imply that in the $q_1[\boldsymbol{u}]$th symbol of w, two pieces of curves would appear: one horizontal, corresponding to the curve connecting the position $q_k[\boldsymbol{u}]$ to $q_1[\boldsymbol{v}]$, and the one that connects $q_1[\boldsymbol{u}]$ to another position in $\boldsymbol{u}$, which is not possible, by the definition of the alphabet of X^s. This is illustrated in Figure 6.

(2) $q_1[\boldsymbol{v}] \leq q_2[\boldsymbol{u}] \leq q_2[\boldsymbol{v}]$: In both cases, this derives from similar arguments.

(3) *Repetition:* We can then repeat these arguments to obtain:

$$q_1[\boldsymbol{u}] \leq q_1[\boldsymbol{v}] \leq q_2[\boldsymbol{u}] \leq \cdots \leq q_n[\boldsymbol{u}] \leq q_n[\boldsymbol{v}],$$

meaning that $\boldsymbol{u}$ and $\boldsymbol{v}$ are interlaced.

($\Leftarrow$): if $|\boldsymbol{u}|_1 = \boldsymbol{v}_1$ and $\boldsymbol{u}, \boldsymbol{v}$ are interlaced, then we define w by connecting $q_i[\boldsymbol{u}]$ to $q_i[\boldsymbol{v}]$ for all $i \in [\![1, n]\!]$. We thus have directly $\boldsymbol{u}\mathcal{R}[w]\boldsymbol{v}$. □

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

PROPOSITION 3. *When* $\boldsymbol{u}\mathcal{R}\boldsymbol{v}$ *and* $\boldsymbol{u} \neq \boldsymbol{v}$, *there exists a unique w such that* $\boldsymbol{u}\mathcal{R}[w]\boldsymbol{v}$. *When* $\boldsymbol{u} = \boldsymbol{v}$, *there are exactly two possibilities, either the word w that connects* $q_i[\boldsymbol{u}]$ *to itself for all i, or the one connecting* $q_i[\boldsymbol{u}]$ *to* $q_{i+1}[\boldsymbol{u}]$ *for all i.*

Proof. Consider words $\boldsymbol{u} \neq \boldsymbol{v}$ and w, such that $\boldsymbol{u}\mathcal{R}[w]\boldsymbol{v}$. Because of Proposition 2, $\boldsymbol{u}$ and $\boldsymbol{v}$ are interlaced. Let us assume that we have (the other case is processed similarly)

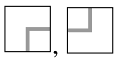
$$q_1[\boldsymbol{u}] \leq q_1[\boldsymbol{v}] \leq q_2[\boldsymbol{u}] \leq \cdots \leq q_n[\boldsymbol{u}] \leq q_n[\boldsymbol{v}].$$

Because the words are different, there is some j such that the position $q_j[\boldsymbol{u}]$ is connected to $q_j[\boldsymbol{v}]$. This forces that the position $q_{j+1}[\boldsymbol{u}]$ is connected to $q_{j+1}[\boldsymbol{v}]$ if $j < n$. If n, $q_1[\boldsymbol{u}]$ is connected to $q_1[\boldsymbol{v}]$. By repeating this, we obtain that for all i, $q_i[\boldsymbol{u}]$ is connected to $q_i[\boldsymbol{v}]$. This determines w, which implies that there is a unique w such that $\boldsymbol{u}\mathcal{R}[w]\boldsymbol{v}$.

When $\boldsymbol{u} = \boldsymbol{v}$, it is clear that there is a unique w connecting position $q_i[\boldsymbol{u}]$ to $q_i[\boldsymbol{v}]$ for all i. Any other w connecting $\boldsymbol{u}$ to $\boldsymbol{v}$ connects $q_i[\boldsymbol{u}]$ to $q_j[\boldsymbol{v}]$ for some $j \neq i$. This j is forced to be $i+1$ (for similar arguments as in the first point of the proof of Proposition 2) and for similar arguments as above, this forces that $q_i[\boldsymbol{u}]$ to $q_{i+1}[\boldsymbol{v}]$ for all $i \leq n$ and connects $q_n[\boldsymbol{u}]$ to $q_1[\boldsymbol{v}]$. □

4.2. *The Lieb path* $t \mapsto V_N(t)$. ▹ In this section, we define the matrices $V_N(t)$. This definition is similar to that of V_N^*, and relies in particular on the interlacing relation and on an additional parameter t. We prove here properties of matrices $V_N(t)$, symmetry and irreducibility, which derive from properties of the relation $\mathcal{R}$. These properties are essential to apply later the Perron–Frobenius theorem, which we recall here.

Notation 5. For all N and $(N, 1)$-cylindrical pattern w, let us denote by $|w|$ the number of symbols



in this pattern. For instance, for the word w in Figure 5, $|w| = 6$.

Definition 6. For all $t \geq 0$, let us define $V_N(t) \in \mathcal{M}_{2^N}(\mathbb{C})$ the matrix such that for all $\boldsymbol{\epsilon}, \boldsymbol{\eta} \in \{0, 1\}_N^*$,

$$V_N(t)[\boldsymbol{\epsilon}, \boldsymbol{\eta}] = \sum_{\boldsymbol{\epsilon}\mathcal{R}[w]\boldsymbol{\eta}} t^{|w|}.$$

It is immediate that V_N^* is equal to $V_N(1)$.

For all N and $n \leq N$, let us denote by $\Omega_N^{(n)} \subset \Omega_N$ the vector space generated by the $\boldsymbol{\epsilon} = |\boldsymbol{\epsilon}_1 \cdots \boldsymbol{\epsilon}_N\rangle$ such that $|\boldsymbol{\epsilon}|_1 = n$.

PROPOSITION 4. *For all N and* $n \leq N$, *the matrix* $V_N(t)$ *stabilizes the vector subspaces* $\Omega_N^{(n)}$:

$$V_N(t).\Omega_N^{(n)} \subset \Omega_N^{(n)}.$$

Proof. This is a direct consequence of Proposition 2, since if $V_N(t)[\boldsymbol{\epsilon}, \boldsymbol{\eta}] \neq 0$ for $\boldsymbol{\epsilon}, \boldsymbol{\eta}$, two elements of the canonical basis of Ω_N, then $|\boldsymbol{\epsilon}|_1 = |\boldsymbol{\eta}|_1$. □

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Let us recall that a non-negative matrix A is called *irreducible* when there exists some $k \geq 1$ such that all the coefficients of A^k are positive. Let us also recall the Perron–Frobenius theorem for symmetric, non-negative and irreducible matrices.

THEOREM 2. (Perron–Frobenius) *Let A be a symmetric, non-negative and irreducible matrix. Then A has a positive eigenvalue λ such that any other eigenvalue μ of A satisfies $|\mu| \leq \lambda$. Moreover, there exists some eigenvector u for the eigenvalue λ with positive coordinates such that if v is another eigenvector (not necessarily for λ) with positive coordinates, then $v = \alpha.u$ for some $\alpha > 0$.*

Let us prove the uniqueness of the positive eigenvector up to a multiplicative constant.

Proof. Let us denote by $u \in \Omega_N$ the Perron–Frobenius eigenvector and $v \in \Omega_N$ another vector whose coordinates are all positive, associated to the eigenvalue μ. Then

$$\mu u^t.v = (Au)^t.v = u^t Av = \lambda u^t.v.$$

Thus, since $u^t.v > 0$, then $\mu = \lambda$, and by (usual version of) Perron–Frobenius, there exists some $\alpha \in \mathbb{R}$ such that $v = \alpha.u$. Since v has positive coordinates, $\alpha > 0$. □

LEMMA 2. *The matrix $V_N(t)$ is symmetric, non-negative (all its coefficients are non-negative numbers) when $t \geq 0$ and for all $n \leq N$, its restriction to $\Omega_N^{(n)}$ is irreducible whenever $t > 0$.*

Proof. The non-negativity of the matrix is immediate when $t \geq 0$ is immediate. Let us prove the other properties.

Symmetry: since the interlacing relation is symmetric, for all $\boldsymbol{\epsilon}, \boldsymbol{\eta} \in \{0, 1\}_N^*$, we have that $V_N(t)[\boldsymbol{\epsilon}, \boldsymbol{\eta}] > 0$ if and only if $V_N(t)[\boldsymbol{\eta}, \boldsymbol{\epsilon}] > 0$. When this is the case, and $\boldsymbol{\epsilon} \neq \boldsymbol{\eta}$ (the case $\boldsymbol{\epsilon} = \boldsymbol{\eta}$ is trivial), there exists a unique (Proposition 3) w connecting $\boldsymbol{\epsilon}$ to $\boldsymbol{\eta}$. The coefficient of this word is exactly $t^{2(n-|\{k:\boldsymbol{\epsilon}_k=\boldsymbol{\eta}_k=1\}|)}$, where $n = |\{k : \boldsymbol{\epsilon}_k = 1\}| = |\{k : \boldsymbol{\eta}_k = 1\}|$, and this coefficient is indifferent to the exchange of $\boldsymbol{\epsilon}$ and $\boldsymbol{\eta}$. This implies that $V_N(t)[\boldsymbol{\epsilon}, \boldsymbol{\eta}] = V_N(t)[\boldsymbol{\eta}, \boldsymbol{\epsilon}]$. As a consequence, $V_N(t)$ is symmetric.

Irreducibility: Let $\boldsymbol{\epsilon}, \boldsymbol{\eta}$ be two elements of the canonical basis of Ω_N such that $|\boldsymbol{\epsilon}|_1 = |\boldsymbol{\eta}|_1 = n$. We shall prove that $V_N^{N^n}(t)[\boldsymbol{\epsilon}, \boldsymbol{\eta}] > 0$.

(1) *Interlacing case:* If we have $\boldsymbol{\epsilon}\mathcal{R}\boldsymbol{\eta}$, then $V_N(t)[\boldsymbol{\epsilon}, \boldsymbol{\eta}] > 0$. Since $V_N(t)[\boldsymbol{\eta}, \boldsymbol{\eta}] > 0$ and $V_N(t)$ is non-negative, for all $k \geq 1$,

$$V_N(t)^k[\boldsymbol{\epsilon}, \boldsymbol{\eta}] \geq V_N(t)[\boldsymbol{\epsilon}, \boldsymbol{\eta}](V_N(t)[\boldsymbol{\eta}, \boldsymbol{\eta}])^{k-1} > 0.$$

In particular, $V_N(t)^{N^n}[\boldsymbol{\epsilon}, \boldsymbol{\eta}] > 0$.

(2) *Non-interlacing case:*

- *Decreasing the interlacing degree:* If we do not have the relation $\boldsymbol{\epsilon}\mathcal{R}\boldsymbol{\eta}$, let us denote by $\omega(\boldsymbol{\epsilon}, \boldsymbol{\eta})$ the following quantity (interlacing degree):

$$\omega(\boldsymbol{\epsilon}, \boldsymbol{\eta}) = \max_i \#\{j : q_j[\boldsymbol{\epsilon}] \in [\![q_i[\boldsymbol{\eta}], q_{i+1}[\boldsymbol{\eta}][\![\}.$$

Since the relation $\boldsymbol{\epsilon}\mathcal{R}\boldsymbol{\eta}$ does not hold, $\omega(\boldsymbol{\epsilon}, \boldsymbol{\eta}) \geq 2$. Let us also denote by $\lambda(\boldsymbol{\epsilon}, \boldsymbol{\eta})$ the number of integers i such that i realizes the maximum in the definition of $\omega(\boldsymbol{\epsilon}, \boldsymbol{\eta})$.

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Let us see that there exists some $\boldsymbol{\epsilon}'$ such that $\boldsymbol{\epsilon}\mathcal{R}\boldsymbol{\epsilon}'$ and $\lambda(\boldsymbol{\epsilon}', \boldsymbol{\eta}) < \lambda(\boldsymbol{\epsilon}, \boldsymbol{\eta})$ if $\lambda(\boldsymbol{\epsilon}, \boldsymbol{\eta}) \geq 2$ and otherwise $\omega(\boldsymbol{\epsilon}', \boldsymbol{\eta}) < \omega(\boldsymbol{\epsilon}, \boldsymbol{\eta})$.

Since $|\boldsymbol{\epsilon}|_1 = |\boldsymbol{\eta}|_1$, there exist i and i' such that $\#\{j : q_j[\boldsymbol{\epsilon}] \in [\![q_i[\boldsymbol{\eta}], q_{i+1}[\boldsymbol{\eta}]]\!]\}$ is maximal and $\#\{j : q_j[\boldsymbol{\epsilon}] \in [\![q_{i'}[\boldsymbol{\eta}], q_{i'+1}[\boldsymbol{\eta}]]\!]\}$ is equal to 0. Let us assume that $i < i'$ (the other case is processed similarly). We can also assume that there is no i'' such that $i < i'' < i'$ such that $\#\{j : q_j[\boldsymbol{\epsilon}] \in [\![q_{i''}[\boldsymbol{\eta}], q_{i''+1}[\boldsymbol{\eta}]]\!]\}$ is equal to 0.

Let us denote j_0 and j_1 such that $q_{j_0}[\boldsymbol{\epsilon}]$ and $q_{j_1}[\boldsymbol{\epsilon}]$ are respectively the maxima of the sets $[\![q_i[\boldsymbol{\eta}], q_{i+1}[\boldsymbol{\eta}]]\!]$ and $[\![q_{i'}[\boldsymbol{\eta}], q_{i'+1}[\boldsymbol{\eta}]]\!]$. There is a word w which connects $q_j[\boldsymbol{\epsilon}]$ to $q_{j+1}[\boldsymbol{\epsilon}]$ for all $j \in [\![j_0, j_1 - 1]\!]$, and fixes $q_j[\boldsymbol{\epsilon}]$ for all other j. The words w thus connects ϵ to ϵ' which satisfies the above properties.

- *A sequence with decreasing interlacing degree:* As a consequence, since $\omega(\boldsymbol{\epsilon}, \boldsymbol{\eta}) \leq N$, one can construct a finite sequence of words $\boldsymbol{\epsilon}^{(k)}$, $k = 1 \ldots m$ such that $m \leq N^n$, $\boldsymbol{\epsilon}^{(1)} = \boldsymbol{\epsilon}$, $\boldsymbol{\epsilon}^{(m)}$ and $\boldsymbol{\eta}$ are interlaced, and for all $k < m$, $\boldsymbol{\epsilon}^{(k)}\mathcal{R}\boldsymbol{\epsilon}^{(k+1)}$. This means that for all $k < m$, $V_N[\boldsymbol{\epsilon}^{(k)}, \boldsymbol{\epsilon}^{(k+1)}] > 0$ and $V_N[\boldsymbol{\epsilon}^{(m)}, \boldsymbol{\eta}] > 0$. As a consequence, $V_N(t)^{N^n}[\boldsymbol{\epsilon}, \boldsymbol{\eta}] > 0$.

This implies that $V_N(t)$ is irreducible on $\Omega_N^{(n)}$ for all $n \leq N$. □

4.3. *Relation between $h(X^s)$ and the matrices $V_N(1)$.* ▷ We know that the entropy of X^s can be obtained out of the sequence $h(\overline{X}_N^s)$. In this section, we prove that the $h(\overline{X}_N^s)$ is related to the eigenvalues of $V_N(1)$, which enables us to use linear algebra to compute the entropy of X^s.

Notation 6. For all N and $n \leq N$, let us denote by $\overline{X}_{n,N}^s$ the subset (which is also a subshift) of $\overline{X}_N^s$ which consists in the set of configurations of $\overline{X}_N^s$ such that the number of curves that cross each of its rows is n, and $\overline{X}_{n,N}$ the subset of $\overline{X}_N$ such that the number of arrows pointing south in the south part of the symbols in any raw is n.

Notation 7. Let us denote, for all N and $n \leq N$, by $\lambda_{n,N}(t)$ the greatest eigenvalue of $V_N(t)$ on $\Omega_N^{(n)}$.

PROPOSITION 5. *For all N and $n \leq N$, $h(\overline{X}_{n,N}^s) = \log_2(\lambda_{n,N}(1))$.*

Proof. Correspondence between elements of $\overline{X}_{n,N}^s$ and trajectories under action of $V_N(1)$: Since for all N, $n \leq N$ and $\boldsymbol{\epsilon}, \boldsymbol{\eta}$ in the canonical basis of $\Omega_N^{(n)}$, $V_N(1)[\boldsymbol{\epsilon}, \boldsymbol{\eta}]$ is the number of ways to connect $\boldsymbol{\epsilon}$ to $\boldsymbol{\eta}$ by an $(N, 1)$-cylindrical pattern, and that there is a natural invertible map from the set of (M, N)-cylindrical patterns to the sequences $(w_i)_{i=1\ldots M}$ of $(N, 1)$-cylindrical patterns such that there exists some $(\boldsymbol{\epsilon}_i)_{i=1\ldots M+1}$ such that for all i, $|\boldsymbol{\epsilon}_i| = n$ and for all $i \leq M$, $\boldsymbol{\epsilon}_i\mathcal{R}[w_i]\boldsymbol{\epsilon}_{i+1}$,

$$\|(V_N(1)^M_{\Omega_N^{(n)}})\|_1 = \mathcal{N}_M(\overline{X}_{n,N}^s).$$

Gelfand's formula: It is known (Gelfand's formula) that $\|(V_N(1)^M_{\Omega_N^{(n)}})\|_1^{1/M} \to \lambda_{n,N}(1)$.

As a consequence of the first point, $h(\overline{X}_{n,N}^s) = \log_2(\lambda_{n,N}(1))$. □

PROPOSITION 6. *For all N, $h(X^s) = \lim_N (1/N) \max_{n\leq N} h(\overline{X}_{n,N}^s)$.*

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Proof. We have the decomposition

$$\overline{X}^s_N = \bigcup_{n=0}^{N} \overline{X}^s_{n,N}.$$

Moreover, these subshifts are disjoint. As a consequence,

$$h(\overline{X}^s_N) = \max_{n \leq N} h(\overline{X}^s_{n,N}).$$

From this, we deduce the statement. □

LEMMA 3. *For all $N \geq 1$ and $n \leq N$, $h(\overline{X}^s_{n,N}) = h(\overline{X}^s_{N-n,N})$.*

Proof. For the purpose of notation, we also denote by π_s the application from $\overline{X}_{n,N}$ to $\overline{X}^s_{n,N}$ that consists in an application of π_s letter by letter. This map is invertible. Let us consider the application $\overline{\mathcal{T}}_{n,N}$ from $\overline{X}_{n,N}$ to $\overline{X}_{N-n,N}$ that inverts all the arrows. This map is an isomorphism, and thus the map $\pi_s \circ \overline{X}_{n,N} \circ \pi_s^{-1}$ is also an isomorphism from $\overline{X}^s_{n,N}$ to $\overline{X}^s_{N-n,N}$. As a consequence, the two subshifts have the same entropy:

$$h(\overline{X}^s_{n,N}) = h(\overline{X}^s_{N-n,N}).$$ □

The following corollary is a straightforward consequence of Lemma 3.

COROLLARY 1. *The entropy of X^s is given by the following formula:*

$$h(X^s) = \lim_N \frac{1}{2N} \max_{n \leq N} \log_2(\lambda_{n,2N}(1)).$$

LEMMA 4. *We deduce that*

$$h(X^s) = \lim_N \frac{1}{2N} \max_{2n+1 \leq N} \log_2(\lambda_{2n+1,2N}(1)).$$

Proof. Let us fix some integer N and for all n between 1 and $N/2+1$, and consider the application that the set of patterns of $\overline{X}^s_{n,N}$ on $\mathbb{U}^{(1)}_M$ associates a pattern of $\overline{X}^s_{n-1,N}$ on $\mathbb{U}^{(1)}_M$ by suppressing the curve that crosses the leftmost symbol in the bottom row of the pattern crossed by a curve [see the schema in Figure 7].

For each pattern of $\overline{X}^s_{n-1,N}$, the number of patterns in its preimage by this transformation is bounded from above by N^M. As a consequence, for all M:

$$\mathcal{N}_M(\overline{X}^s_{n-1,N}).N^M \geq \mathcal{N}_M(\overline{X}^s_{n,N}),$$

and thus

$$h(\overline{X}^s_{n-1,N}) + \log_2(N) \geq h(\overline{X}^s_{n,N}).$$

As a consequence,

$$h(X^s) = \lim_N \frac{1}{2N} \max(\max_{2n+1 \leq N} h(\overline{X}^s_{2n+1,2N}), \max_{2n \leq N} h(\overline{X}^s_{2n,2N}))$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

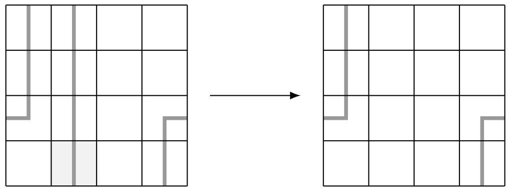


FIGURE 7. Illustration of the curve suppressing operation; the leftmost position of the bottom row crossed by a curve is colored gray on the left pattern.

$$\leq \lim_N \frac{1}{2N} \max\Big(\max_{2n+1\leq N} h(\overline{X}^s_{2n+1,2N}), \max_{2n\leq N} h(\overline{X}^s_{2n-1,2N}) + \log_2(N)\Big)$$
$$\leq \lim_N \frac{1}{2N} \max_{2n+1\leq N} h(\overline{X}^s_{2n+1,2N}).$$

Moreover, it is straightforward that

$$h(X^s) \geq \lim_N \frac{1}{2N} \max_{2n+1\leq N} h(\overline{X}^s_{2n+1,2N}),$$

thus we have the following equality:

$$h(X^s) = \lim_N \frac{1}{2N} \max_{2n+1\leq N} \log_2(\lambda_{2n+1,2N}(1)). \qquad \square$$

5. *Coordinate Bethe ansatz*

▹ Let us remember that we proved in the last section that the entropy of X^s can be computed out of the eigenvalues of the matrices $V_{2N}(1)$. Ideally, we would like to diagonalize these matrices, which is in fact very difficult. The purpose of the (coordinate) Bethe ansatz [§5.2] is to provide instead candidate eigenvectors for the matrix $V_{2N}(t)$ for all t on all $\Omega^{(n)}_{2N}$, whose formulation relies on some solution of the system of Bethe equations $(E_j)[t, n, N]$, $j \leq n$ (see §5.2). We prove the existence unicity and analyticity relative to the parameter t of the solutions of the system in §5.3. For the statement of the ansatz, we need to introduce some auxiliary functions [§5.1] which are involved in its formulation, and prove some properties they satisfy which will be useful in particular to prove the existence and analyticity of the solutions of the system of Bethe equations $(E_j)[t, n, N]$, $j \leq n$. We will prove that the candidate eigenvalue corresponding to the candidate eigenvector is the maximal eigenvalue of $V_{2N}(t)$ on $\Omega^{(2n+1)}_{2N}$ for $2n + 1 \leq N$, for t close to $\sqrt{2}$, in §5.5. This relies on the diagonalization of the Hamiltonian mentioned in the overview [§5.4]. The analyticity of the solutions to the system of Bethe equations implies that this is true for all $t \in (0, \sqrt{2})$.

5.1. *Auxiliary functions.*

▹ The purpose of the present section is to introduce the functions Θ and κ (respectively, Notation 8 and Notation 9) which will be used in the statement of the ansatz and then through the whole article. The reader may skip the details of computations done in this section. We provide them because they were not properly

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

proven and may ultimately help to understand under which conditions it is possible to use an argument similar to Bethe ansatz in order to compute entropy for other multidimensional SFT.

5.1.1. *Notation.* Let us denote by $\mu : (-1, 1) \to (0, \pi)$ the inverse of the function $\cos : (0, \pi) \to (-1, 1)$. For all $t \in (0, \sqrt{2})$, we will set $\Delta_t = (2 - t^2)/2$, $\mu_t = \mu(-\Delta_t)$ and $I_t = (-(\pi - \mu_t), (\pi - \mu_t))$.

Notation 8. Let us denote by Θ the unique analytic function $(t, x, y) \mapsto \Theta_t(x, y)$ from the set $\{(t, x, y) : x, y \in I_t\}$ to $\mathbb{R}$ such that $\Theta_{\sqrt{2}}(0, 0) = 0$ and for all t, x, y,

$$\exp(-i\Theta_t(x, y)) = \exp(i(x - y)) \cdot \frac{e^{-ix} + e^{iy} - 2\Delta_t}{e^{-iy} + e^{ix} - 2\Delta_t}.$$

By a unicity argument, one can see that for all t, x, y, $\Theta_t(x, y) = -\Theta_t(y, x)$. As a consequence, for all x, $\Theta_t(x, x) = 0$. For the same reason, $\Theta_t(x, -y) = -\Theta_t(-x, y)$ and $\Theta_t(-x, -y) = -\Theta_t(x, y)$. Moreover, Θ_t and all its derivatives can be extended by continuity on $I_t^2 \backslash \{(x, x) : x \in \partial I_t\}$. For the purpose of notation, we will also denote by Θ_t the extended function. We will use the following.

COMPUTATION 1. *For all $y \neq (\pi - \mu_t)$, $\Theta_t((\pi - \mu_t), y) = 2\mu_t - \pi$.*

Proof. From the definition of μ_t, $\Delta_t = -\cos(\mu_t) = -(e^{i\mu_t} + e^{-i\mu_t})/2$. As a consequence, from the definition of Θ_t,

$$\exp(-i\Theta_t((\pi - \mu_t), y)) = e^{i(\pi-\mu_t-y)} \cdot \frac{e^{iy} - e^{i\mu_t} - 2\Delta_t}{e^{-iy} - e^{-i\mu_t} - 2\Delta_t} = e^{i(\pi-\mu_t-y)} \cdot \frac{e^{iy} + e^{-i\mu_t}}{e^{-iy} + e^{i\mu_t}}.$$

As a consequence,

$$\exp(-i\Theta_t((\pi - \mu_t), y)) = e^{i(\pi-\mu_t-y)} \cdot \frac{e^{iy}}{e^{i\mu_t}} \frac{1 + e^{-i(\mu_t+y)}}{e^{-i(y+\mu_t)} + 1} = e^{i(\pi-2\mu_t)}.$$

This yields the statement as a consequence. □

Notation 9. Let us denote by κ the unique analytic map $(t, \alpha) \mapsto \kappa_t(\alpha)$ from $(0, \sqrt{2}) \times \mathbb{R}$ to $\mathbb{R}$ such that $\kappa_{\sqrt{2}/2}(0) = 0$ and for all t, α,

$$e^{i\kappa_t(\alpha)} = \frac{e^{i\mu_t} - e^{\alpha}}{e^{i\mu_t+\alpha} - 1}.$$

With the argument of uniqueness, we have that for all t, α, $\kappa_t(-\alpha) = -\kappa_t(\alpha)$, and as a consequence, $\kappa_t(0) = 0$. We also set for all t, α, β,

$$\theta_t(\alpha, \beta) = \Theta_t(\kappa_t(\alpha), \kappa_t(\beta)).$$

5.1.2. *Properties of the auxiliary functions.* ▹ In this section, we prove some properties of the functions Θ and κ (computation of derivatives, invertibility, and a relation between Θ and κ), which will be used later.

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Computation of the derivative κ_t'.

COMPUTATION 2. *Let us fix some* $t \in (0, \sqrt{2})$. *For all* $\alpha \in \mathbb{R}$,

$$\kappa_t'(\alpha) = \frac{\sin(\mu_t)}{\cosh(\alpha) - \cos(\mu_t)}.$$

Proof. Computation of $\cos(\kappa_t(\alpha))$ *and* $\sin(\kappa_t(\alpha))$: By definition of κ (for the first equality, we multiply both numerator and denominator by $(e^{-i\mu_t+\alpha} - 1)$),

$$e^{i\kappa_t(\alpha)} = \frac{(e^{-i\mu_t+\alpha} - 1)(e^{i\mu_t} - e^{\alpha})}{|e^{i\mu_t+\alpha} - 1|^2} = \frac{e^{\alpha} - e^{2\alpha}e^{-i\mu_t} - e^{i\mu_t} + e^{\alpha}}{(\cos(\mu_t)e^{\alpha} - 1)^2 + (\sin(\mu_t)e^{\alpha})^2}.$$

Thus by taking the real part,

$$\cos(\kappa_t(\alpha)) = \frac{2e^{\alpha} - (e^{2\alpha} + 1)\cos(\mu_t)}{\cos^2(\mu_t)e^{2\alpha} - 2\cos(\mu_t)e^{\alpha} + 1 + (1 - \cos^2(\mu_t))e^{2\alpha}},$$

$$\cos(\kappa_t(\alpha)) = \frac{2e^{\alpha} - (e^{2\alpha} + 1)\cos(\mu_t)}{e^{2\alpha} - 2\cos(\mu_t)e^{\alpha} + 1} = \frac{1 - \cos(\mu_t)\cosh(\alpha)}{\cosh(\alpha) - \cos(\mu_t)},$$

where we factorized by $2e^{\alpha}$ for the second equality. As a consequence,

$$\cos(\kappa_t(\alpha)) = \frac{\sin^2(\mu_t) + \cos^2(\mu_t) - \cos(\mu_t)\cosh(\alpha)}{\cosh(\alpha) - \cos(\mu_t)}$$

$$= \frac{\sin^2(\mu_t)}{\cosh(\alpha) - \cos(\mu_t)} - \cos(\mu_t). \tag{3}$$

A similar computation gives

$$\sin(\kappa_t(\alpha)) = \frac{\sin(\mu_t)\sinh(\alpha)}{\cosh(\alpha) - \cos(\mu_t)}. \tag{4}$$

Deriving the expression $\cos(\kappa_t(\alpha))$: By deriving equation (3), for all α,

$$-\kappa_t'(\alpha)\sin(\kappa_t(\alpha)) = -\frac{\sin^2(\mu_t)\sinh(\alpha)}{(\cosh(\alpha) - \cos(\mu_t))^2} = -\frac{\sin(\kappa_t(\alpha))^2}{\sinh(\alpha)},$$

where we used equation (4) for the second equality.

Thus, for all α but in a discrete subset of $\mathbb{R}$,

$$\kappa_t'(\alpha) = \frac{\sin(\mu_t)}{\cosh(\alpha) - \cos(\mu_t)}.$$

This identity is thus verified on all $\mathbb{R}$ by continuity. □

Domain and invertibility. Let us remember that $I_t = (-(\pi - \mu_t), (\pi - \mu_t))$.

PROPOSITION 7. *For all* t, $\kappa_t(\mathbb{R}) \subset I_t$. *Moreover,* κ_t, *considered as a function from* $\mathbb{R}$ *to* I_t, *is bijective.*

Proof. Injectivity: Since $\mu_t \in (0, \pi)$, then $\sin(\mu_t) > 0$ and we have the inequality $\cosh(\alpha) \geq 1 > \cos(\mu_t)$. As a consequence of Computation 2, κ_t is strictly increasing, and thus injective.

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

The equality $\kappa_t(\alpha) = n\pi$ implies $\alpha = 0$: Assume that for some α, $\kappa_t(\alpha) = n\pi$ for some integer n. By definition of κ,

$$e^{in\pi} = \frac{e^{i\mu_t} - e^{\alpha}}{e^{i\mu_t+\alpha} - 1}.$$

If n was odd, then

$$e^{\alpha} - e^{i\mu_t} = e^{i\mu_t+\alpha} - 1,$$

$$e^{\alpha} + 1 = e^{i\mu_t} \cdot (e^{\alpha} + 1),$$

and thus $e^{i\mu_t} = 1$, which is impossible, since $\mu_t \in (0, \pi)$. Thus n is even, and then

$$-e^{\alpha} + e^{i\mu_t} = e^{i\mu_t+\alpha} - 1.$$

As a consequence, since $e^{i\mu_t} \neq -1$, we have $e^{\alpha} = 1$, and thus $\alpha = 0$.

Extension of the images: Since when α tends towards $+\infty$ (respectively, $-\infty$), $(e^{i\mu_t} - e^{\alpha})/(e^{i\mu_t+\alpha} - 1)$ tends towards $-e^{i\mu_t}$ (respectively, $e^{i\mu_t}$), $\kappa_t(\alpha)$ tends towards some $n\pi - \mu_t$ (respectively, $m\pi + \mu_t$). and from the last point, $n = 1$ (respectively, $m = -1$). Thus the image of κ_t is the set I_t.

Thus κ_t is an invertible map from $\mathbb{R}$ to I_t. □

A relation between θ_t and κ_t. The following equality originates in [YY66a]. We provide some details of a relatively simple way to compute it.

COMPUTATION 3. *For any numbers t, α, β,*

$$\frac{\partial \theta_t}{\partial \alpha}(\alpha, \beta) = -\frac{\partial \theta_t}{\partial \beta}(\alpha, \beta) = -\frac{\sin(2\mu_t)}{\cosh(\alpha - \beta) - \cos(2\mu_t)}.$$

Proof. Deriving the equation that defines Θ_t: Let us set, for all x, y,

$$G_t(x, y) = \frac{x(1 - 2\Delta_t y) + y}{x + y - 2\Delta_t}.$$

Then we have that for all x, y,

$$\frac{\partial G_t}{\partial x}(x, y) = \frac{(1 - 2\Delta_t y) \cdot (x + y - 2\Delta_t) - (x(1 - 2\Delta_t y) + y)}{(x + y - 2\Delta_t)^2}.$$

Then,

$$\frac{\partial G_t}{\partial x}(x, y) = -2\Delta_t \frac{1 + y^2 - 2\Delta_t y}{(x + y - 2\Delta_t)^2}. \tag{5}$$

For all t, α, let us set $\alpha_t \equiv \kappa_t(\alpha)$. By definition of Θ_t, for all α, β,

$$\exp(-i\Theta_t(\alpha_t, \beta_t)) = G(e^{i\alpha_t}, e^{-i\beta_t}). \tag{6}$$

Thus we have, by deriving the equality in equation (6),

$$-i\frac{d}{d\alpha}(\Theta_t(\alpha_t, \beta_t)) \exp(-i\Theta_t(\alpha_t, \beta_t)) = i\kappa_t'(\alpha)e^{i\alpha_t}\frac{\partial}{\partial x}G_t(e^{i\alpha_t}, e^{-i\beta_t}).$$

Then,

$$\frac{d}{d\alpha}(\Theta_t(\alpha_t,\beta_t)) = -\kappa_t'(\alpha)e^{i\alpha_t}\frac{(\partial/\partial x)G_t(e^{i\alpha_t},e^{-i\beta_t})}{G_t(e^{i\alpha_t},e^{-i\beta_t})},$$

and thus, using equation (5),

$$\frac{d}{d\alpha}\Theta_t(\alpha_t,\beta_t) = \frac{(\kappa_t'(\alpha)2\Delta e^{i\alpha_t})(1+e^{-2i\beta_t}-2\Delta e^{-i\beta_t})}{(e^{i\alpha_t}+e^{-i\beta_t}-2\Delta_t)(e^{i\alpha_t}+e^{-i\beta_t}-2\Delta_t.e^{i(\alpha_t-\beta_t)})}.$$

Factoring by $e^{i(\alpha_t-\beta_t)}$,

$$\frac{d}{d\alpha}\Theta_t(\alpha_t,\beta_t) = 2\Delta_t\kappa_t'(\alpha)\cdot\frac{e^{i\beta_t}+e^{-i\beta_t}-2\Delta_t}{(e^{i\alpha_t}+e^{-i\beta_t}-2\Delta_t)(e^{i\beta_t}+e^{-i\alpha_t}-2\Delta_t)}. \tag{7}$$

Simplification of the term $e^{i\alpha_t}+e^{-i\beta_t}-2\Delta_t$ in equation (7): Let us denote the function F defined on α,β by

$$F_t(\alpha,\beta) = e^{i\alpha_t}+e^{-i\beta_t}-2\Delta_t. \tag{8}$$

By definition of κ_t and $-2\Delta_t = e^{-i\mu_t}+e^{i\mu_t}$, we have

$$F_t(\alpha,\beta) = \frac{e^{i\mu_t}-e^{\alpha}}{e^{i\mu_t+\alpha}-1}+\frac{e^{i\mu_t+\beta}-1}{e^{i\mu_t}-e^{\beta}}+e^{i\mu_t}+e^{-i\mu_t}.$$

Thus $F_t(\alpha,\beta)$ is equal to

$$\frac{(e^{i\mu_t}-e^{\alpha})(e^{i\mu_t}-e^{\beta})+(e^{i\mu_t+\alpha}-1)(e^{i\mu_t+\beta}-1)+(e^{i\mu_t}+e^{-i\mu_t})\cdot(e^{i\mu_t+\alpha}-1)(e^{i\mu_t}-e^{\beta})}{(e^{i\mu_t+\alpha}-1)(e^{i\mu_t}-e^{\beta})}.$$

Finally,

$$F_t(\alpha,\beta) = \frac{e^{3i\mu_t+\alpha}+e^{\beta-i\mu_t}-e^{i\mu_t}\cdot(e^{\alpha}+e^{\beta})}{(e^{i\mu_t+\alpha}-1)(e^{i\mu_t}-e^{\beta})}. \tag{9}$$

Simplification of the derivative of Θ_t: For all α,β, we have, using equations (7) and (8),

$$\frac{1}{\kappa_t'(\alpha)}\frac{d}{d\alpha}\Theta_t(\alpha,\beta) = 2\Delta_t\frac{F_t(\beta,\beta)}{F_t(\alpha,\beta)\cdot F_t(\beta,\alpha)}.$$

As a consequence of equation (9),

$$\begin{aligned}&\frac{1}{\kappa_t'(\alpha)}\frac{d}{d\alpha}\Theta_t(\alpha_t,\beta_t)\\ &\quad= \frac{-(e^{i\mu_t+\alpha}-1)(e^{i\mu_t}-e^{\alpha})(e^{-i\mu_t}+e^{i\mu_t})(e^{3i\mu_t+\beta}+e^{\beta-i\mu_t}-2e^{i\mu_t}.e^{\beta})}{(e^{3i\mu_t+\alpha}+e^{\beta-i\mu_t}-e^{i\mu_t}.(e^{\alpha}+e^{\beta}))(e^{3i\mu_t+\beta}+e^{\alpha-i\mu_t}-e^{i\mu_t}.(e^{\beta}+e^{\alpha}))},\end{aligned}$$

and thus

$$\begin{aligned}&\frac{1}{\kappa_t'(\alpha)}\frac{d}{d\alpha}\Theta_t(\alpha_t,\beta_t)\\ &\quad= -\frac{(e^{i\mu_t+\alpha}-1)(e^{i\mu_t}-e^{\alpha})e^{\beta}\cdot(e^{2i\mu_t}-1)\cdot(e^{2i\mu_t}-e^{-2i\mu_t})}{e^{2i\mu_t}(e^{2i\mu_t+\alpha}+e^{\beta-2i\mu_t}-(e^{\alpha}+e^{\beta}))(e^{2i\mu_t+\beta}+e^{\alpha-2i\mu_t}-(e^{\beta}+e^{\alpha}))}.\end{aligned}$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Since in the denominator of the fraction is the square of the modulus of some number, we rewrite it.

$$\frac{1}{\kappa_t'(\alpha)}\frac{d}{d\alpha}\Theta_t(\alpha_t,\beta_t) = -\frac{(e^{i\mu_t+\alpha}-1)(e^{i\mu_t}-e^{\alpha})e^{\beta}\cdot(e^{2i\mu_t}-1)\cdot(e^{2i\mu_t}-e^{-2i\mu_t})}{e^{2i\mu_t}((e^{\alpha}+e^{\beta})^2(\cos(2\mu_t)-1)^2+(e^{\alpha}-e^{\beta})^2\sin^2(2\mu_t))}.$$

We rewrite also the other terms, by splitting the $e^{2i\mu_t}$ in the denominator in two parts, one makes $\sin(\mu_t)$ appear and the other one the square modulus in the following formula:

$$\frac{1}{\kappa_t'(\alpha)}\frac{d}{d\alpha}\Theta_t(\alpha_t,\beta_t) = -4|e^{i\mu_t+\alpha}-1|^2\frac{e^{\beta}\cdot\sin(\mu_t)\cdot\sin(2\mu_t)}{(e^{\alpha}+e^{\beta})^2(\cos(2\mu_t)-1)^2+(e^{\alpha}-e^{\beta})^2\sin^2(2\mu_t)}.$$

By writing $\sin^2(2\mu_t)=1-\cos^2(2\mu_t)$ and then factoring by $1-\cos(2\mu_t)$,

$$\frac{1}{\kappa_t'(\alpha)}\frac{d}{d\alpha}\Theta_t(\alpha_t,\beta_t) = -4\frac{|e^{i\mu_t+\alpha}-1|^2}{1-\cos(2\mu_t)}\cdot\frac{e^{\beta}\cdot\sin(\mu_t)\cdot\sin(2\mu_t)}{(e^{\alpha}+e^{\beta})^2(1-\cos(2\mu_t))+(e^{\alpha}-e^{\beta})^2(1+\cos(2\mu_t))}.$$

Simplifying the denominator and factoring it by $4e^{\alpha+\beta}$, we obtain

$$\frac{1}{\kappa_t'(\alpha)}\frac{d}{d\alpha}\Theta_t(\alpha_t,\beta_t) = -\frac{|e^{i\mu_t+\alpha}-1|^2}{e^{\alpha}(1-\cos(2\mu_t))}\cdot\frac{\sin(\mu_t)\cdot\sin(2\mu_t)}{\cosh(\alpha-\beta)-\cos(2\mu_t)}.$$

We have left to see that

$$\frac{\sin(\mu_t)\kappa_t'(\alpha)\cdot|e^{i\mu_t+\alpha}-1|^2}{e^{\alpha}(1-\cos(2\mu_t))}=1.$$

This derives directly from $1-\cos(2\mu_t)=2\sin^2(\mu_t)$ and the value of $\kappa_t'(\alpha)$ given by Computation 2.

We thus have the stated formula of $\partial\theta_t/\partial\alpha(\alpha,\beta)=d/d\alpha(\Theta_t(\alpha_t,\beta_t))$.

The other equality: We obtain the value of $\partial\theta_t/\partial\beta(\alpha,\beta)=d/d\beta(\Theta_t(\alpha_t,\beta_t))$ through the equality $\Theta_t(x,y)=-\Theta_t(y,x)$ for all x,y (§5.1.1). □

LEMMA 5. *For all t,α,β,*

$$\theta_t(\alpha+\beta,\alpha)=\theta_t(\beta,0).$$

Proof. Let us fix some $\alpha\in\mathbb{R}$. By Computation 3, the derivative of the function $\beta\mapsto\theta_t(\alpha+\beta,\alpha)$ is equal to the derivative of the function $\beta\mapsto\theta_t(\beta,0)$. As a consequence, these two functions differ by a constant. Since they have the same value 0 in $\beta=0$ (§5.1.1), they are equal. □

5.2. *Statement of the ansatz.* ▹ In this section, we state of the coordinate Bethe ansatz in Theorem 3 (let us remember that this provides candidate eigenvectors and eigenvalues for $V_N(t)$ for all N and t).

Notation 10. For all $(p_1, \ldots, p_n) \in I_t^n$, let us denote by $\psi_{\mu_t,n,N}(p_1, \ldots, p_n)$ the vector in Ω_N such that for all $\boldsymbol{\epsilon} \in \{0, 1\}_N^*$,

$$\psi_{\mu_t,n,N}(p_1, \ldots, p_n)[\boldsymbol{\epsilon}] = \sum_{\sigma \in \Sigma_n} C_\sigma(t)[p_1, \ldots, p_n] \prod_{k=1}^{n} e^{ip_{\sigma(k)} q_k[\boldsymbol{\epsilon}]},$$

where, denoting by $\epsilon(\sigma)$ the signature of σ,

$$C_\sigma(t)[p_1, \ldots, p_n] = \epsilon(\sigma) \prod_{1 \leq k < l \leq n} (1 + e^{i(p_{\sigma(k)} + p_{\sigma(l)})} - 2\Delta_t e^{ip_{\sigma(k)}}).$$

Definition 7. We say that $p_1, \ldots, p_n \in I_t$ satisfy the system of *Bethe equations* when for all j,

$$(E_j)[t, n, N] : Np_j = 2\pi j - (n+1)\pi - \sum_{k=1}^{n} \Theta_t(p_j, p_k).$$

THEOREM 3. *For all N and $n \leq N/2$, and $p_1, \ldots, p_n \in I_t$ distinct which satisfy the system of Bethe equations, we have*

$$V_N(t) \cdot \psi_{n,N}(p_1, \ldots, p_n) = \Lambda_{n,N}(t)[p_1, \ldots, p_n] \cdot \psi_{n,N}(p_1, \ldots, p_n),$$

where $\Lambda_{n,N}(t)[p_1, \ldots, p_n]$ is equal to

$$\prod_{k=1}^{n} L_t(e^{ip_k}) + \prod_{k=1}^{n} M_t(e^{ip_k})$$

when all the p_k are distinct from 0. *Else, it is equal to*

$$\left(2 + t^2(N-1) + \sum_{k \neq l} \frac{\partial \Theta_t}{\partial x}(0, p_k)\right) \prod_{k=1}^{n} M_t(e^{ip_k})$$

for l such that $p_l = 0$.

In [**DGHMT18**] (Theorem 2.2), the equations (BE) are implied by the equations $(E_j)[t, n, N]$ in Theorem 3 by taking the exponential of the members of (BE). To make the connection easier with [**DGHMT18**], here is a list of correspondences between the notation: in [**DGHMT18**], the notation t corresponds to c, and it is fixed in the formulation of the theorem. Thus, Δ_t corresponds to Δ, $\mathcal{I}_t$ to $\mathcal{D}_\Delta$, $V_N(t)$ to V, $\psi_{t,n,N}(p_1, \ldots, p_n)$ to ψ, L_t and M_t to L and M, Θ_t to Θ, $\Lambda_{n,N}(t)[p_1, \ldots, p_n]$ to Λ, $C_\sigma(t)[p_1, \ldots, p_n]$ to A_σ and the sequence $(x_k)_k$ to the sequence $(q_k[\boldsymbol{\epsilon}])_k$ for some $\boldsymbol{\epsilon}$.

5.3. *Existence of solutions of Bethe equations and analyticity.* ▹ As a matter of fact, the Bethe ansatz presented in §5.2 provides a candidate eigenvector on the condition that there exists a solution to the system of Bethe equations. We prove that this is the case for all N and t and that this solution is unique. Furthermore, we will need that this solution is analytic relative to t for all N. These statements are encompassed in the following theorem, whose proof is a rigorous and complete version of an argument in [**YY66a**].

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

THEOREM 4. *There exists a unique sequence of analytic functions* $\boldsymbol{p}_j : (0, \sqrt{2}) \mapsto (-\pi, \pi)$ *such that for all* $t \in (0, \sqrt{2})$, $\boldsymbol{p}_j(t) \in I_t$ *and we have the system of Bethe equations:*

$$(E_j)[t, n, N] : N p_j(t) = 2\pi j - (n+1)\pi - \sum_{k=1}^{n} \Theta_t(p_j(t), p_k(t)).$$

Moreover, for all t and j, $\boldsymbol{p}_{n-j+1}(t) = -\boldsymbol{p}_j(t)$*; for all t, the* $\boldsymbol{p}_j(t)$ *are all distinct.*

Idea of the proof: Following Yang and Yang [YY66a], we use an auxiliary multivariate function ζ_t whose derivative is zero exactly when the equations $(E_j)[t, n, N]$ are verified. We prove that up to a monotonous change of variable, this function is convex, which implies that it admits a unique local (and thus global) minimum (this relies on the properties of θ_t and κ_t). Since we rule out the possibility that the minimum is on the border of the domain, this function admits a point where its derivative is zero, and thus the system of equations $(E_j)[t, n, N]$ admits a unique solution. To prove the analyticity, we then define a function of t that verifies an analytic differential equation (and thus is analytic), whose value in some point coincides with the minimum of ζ_t. Since the differential equation ensures that ζ_t' is null on the values of this function, this means that for all t, its value in t is the minimum of ζ_t.

Proof. The solutions are critical points of an auxiliary function ζ_t: Let us set, for all $t, p_1, \ldots, p_n$ (in the third sum, both k and j are arguments of the sum),

$$\zeta_t(p_1, \ldots, p_n) = N \sum_{j=1}^{n} \int_0^{\kappa_t^{-1}(p_j)} \kappa_t(x)\, dx + \pi(n+1-2j) \sum_{j=1}^{n} \kappa_t^{-1}(p_j) + \sum_{k<j} \int_0^{\kappa_t^{-1}(p_j) - \kappa_t^{-1}(p_k)} \theta_t(x, 0)\, dx.$$

The interest of this function lies in the fact that for all j (here the argument in each of the sums is k),

$$\begin{aligned}
\frac{\partial \zeta_t}{\partial p_j}(p_1, \ldots, p_n) &= (\kappa_t^{-1})'(p_j).\Big(N p_j - 2\pi j + (n+1)\pi \\
&\quad - \sum_{k<j} \theta_t(\kappa_t^{-1}(p_j) - \kappa_t^{-1}(p_k), 0)\Big) + \sum_{k>j} \theta_t(\kappa_t^{-1}(p_k) - \kappa_t^{-1}(p_j), 0). \\
&= (\kappa_t^{-1})'(p_j).\Big(N p_j - 2\pi j + (n+1)\pi - \sum_{k<j} \theta_t(\kappa_t^{-1}(p_k), \kappa_t^{-1}(p_j))\Big) \\
&\quad + \sum_{k>j} \theta_t(\kappa_t^{-1}(p_j), \kappa_t^{-1}(p_k)).
\end{aligned}$$

$$\frac{\partial \zeta_t}{\partial p_j}(p_1, \ldots, p_n) = (\kappa_t^{-1})'(p_j).\Big(N p_j - 2\pi j + (n+1)\pi + \sum_{k} \Theta_t(p_j, p_k)\Big), \tag{10}$$

since for all x, y, $\Theta_t(x, y) = -\Theta_t(y, x)$ and for all x, $\Theta_t(x, x) = 0$ (§5.1.1).

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Hence, the system of Bethe equations is verified for the sequence $(p_j)_j$ if and only for all j, $\partial\zeta_t/\partial p_j(p_1, \ldots, p_n) = 0$.

Uniqueness of the local minimum of ζ_t *using convexity:* Let us set $\tilde{\zeta}_t : \mathbb{R}^n \to \mathbb{R}$ such that for all $\alpha_1, \ldots, \alpha_n$:

$$\tilde{\zeta}_t(\alpha_1, \ldots, \alpha_n) = \zeta_t(\kappa_t(\alpha_1), \ldots, \kappa_t(\alpha_n)).$$

From equation (10), we have that for all sequence $(\alpha_k)_k$ and all j,

$$\frac{\partial\tilde{\zeta}_t}{\partial p_j}(\alpha_1, \ldots, \alpha_n) = N\kappa_t(\alpha_j) - 2\pi j + (n+1)\pi + \sum_k \theta_t(\alpha_j, \alpha_k).$$

As a consequence, using Computation 3, for all $k \neq j$,

$$\frac{\partial^2\tilde{\zeta}_t}{\partial p_k \partial p_j}(\alpha_1, \ldots, \alpha_n) = \frac{\partial\theta_t}{\partial\beta}(\alpha_j, \alpha_k) = \frac{\sin(2\mu_t)}{\cosh(\alpha_j - \alpha_k) - \cos(2\mu_t)}. \tag{11}$$

Moreover, for all j,

$$\frac{\partial^2\tilde{\zeta}_t}{\partial^2 p_j}(\alpha_1, \ldots, \alpha_n, t) = N\kappa_t'(\alpha_j) + \sum_{k\neq j}\frac{\partial\theta_t}{\partial\alpha}(\alpha_j, \alpha_k) = N\kappa_t'(\alpha_j) - \sum_{k\neq j}\frac{\partial\theta_t}{\partial\beta}(\alpha_j, \alpha_k). \tag{12}$$

Let us denote by $\tilde{H}_t(\alpha_1, \ldots, \alpha_n)$ the Hessian matrix of $\tilde{\zeta}_t$. For any $(x_1, \ldots, x_n) \in \mathbb{R}^n$, we have, from equations (11) and (12),

$$\begin{aligned}
&||(x_1, \ldots, x_n)\cdot\tilde{H}_t(\alpha_1, \ldots, \alpha_n)\cdot(x_1, \ldots, x_n)^t \\
&\quad = N\sum_j \kappa_t'(\alpha_j){x_j}^2 - \sum_{j\neq k}\bigg(\frac{\partial\theta_t}{\partial\beta}(\alpha_j, \alpha_k)x_j(x_j - x_k).\bigg) \\
&\quad = N\sum_j \kappa_t'(\alpha_j){x_j}^2 - \sum_{j<k}\bigg(\frac{\partial\theta_t}{\partial\beta}(\alpha_j, \alpha_k)x_j(x_j - x_k)\bigg) \\
&\qquad - \sum_{j<k}\bigg(\frac{\partial\theta_t}{\partial\beta}(\alpha_k, \alpha_j)x_k(x_k - x_j)\bigg) \\
&\quad = N\sum_j \kappa_t'(\alpha_j){x_j}^2 - \sum_{j<k}\bigg(\frac{\partial\theta_t}{\partial\beta}(\alpha_j, \alpha_k)(x_j - x_k)^2\bigg) > 0.
\end{aligned}$$

As a consequence, $\tilde{\zeta}_t$ is a convex function. Thus, if it has a local minimum, it is unique. Since κ_t is increasing, this property is also true for ζ_t. *The function* ζ_t *has a minimum in* I_t^n: Let us consider $(C_l)_l$ as an increasing sequence of compact intervals such that $\bigcup_l C_l = I_t$.

Let us assume that ζ_t has no local minimum in I_t^n. As a consequence, for all j, the minimum $\mathbf{p}^{(l)}$ of ζ_t on $(C_l)^n$ is on its border. Without loss of generality, we can assume that there exists some $\mathbf{p}^{(\infty)} \in \overline{I_t}^n$ such that $\mathbf{p}^{(l)} \to \mathbf{p}^{(\infty)}$.

We can assume without loss of generality that there exists some $j_0 \in [\![1, n]\!]$ such that $j \leq j_0$ if and only if $\mathbf{p}_j^{(\infty)} = \pi - \mu_t$. The number j_0 is the number of j such that $\mathbf{p}_j^{(\infty)} = \pi - \mu_t$. We can assume furthermore that $j_0 \leq n/2$: if it is not the case, then we use a reasoning similar to the one that follows, replacing $\pi - \mu_t$ by $-(\pi - \mu_t)$.

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

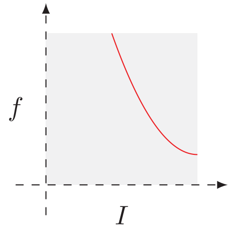


FIGURE 8. Illustration of the fact that the minimum of a convex continuously differentiable function on a real compact interval has non-positive derivative.

Thus, there exists l_0 such that for all l and $j \leq j_0$, $\mathbf{p}_j^{(l)} \geq 0$. Since $\tilde{\zeta}_t$ is convex and that $\mathbf{p}^{(l)}$ is a minimum for this function on the compact set $(C_l)^n$, then for all $j \leq j_0$,

$$\frac{\partial \zeta_t}{\partial p_j}(\mathbf{p}_1^{(l}, \ldots, \mathbf{p}_n^{(l)}) \leq 0. \tag{13$_j$}$$

This is a particular case of the fact that for a convex and continuously differentiable function $f : I \mapsto \mathbb{R}$, where I is a compact interval of $\mathbb{R}$, if its minimum on I occurs at maximal element of I, then f' is non-positive on this point, as illustrated in Figure 8.

Since Θ_t cannot be defined on $\{(x, x) : x \in \partial I_t\}$, to have an inequality that can be transformed by continuity into an inequality on $\mathbf{p}$, we sum the inequalities in equation $(13)_j$:

$$\sum_{j=1}^{j_0} \frac{\partial \zeta_t}{\partial p_j}(\mathbf{p}_1^{(l)}, \ldots, \mathbf{p}_n^{(l)}) \leq 0.$$

According to the first point of the proof (equation (10)), this inequality can be re-written:

$$N \sum_{j=1}^{j_0} \mathbf{p}_j - 2\pi \sum_{j=1}^{j_0} j + j_0(n+1)\pi + \sum_{j=1}^{j_0} \sum_{k} \Theta_t(\mathbf{p}_j^{(l)}, \mathbf{p}_k^{(l)}) \leq 0.$$

For all $j, j' \leq j_0$, the terms $\Theta_t(\mathbf{p}_j^{(l)}, \mathbf{p}_{j'}^{(l)})$ and $\Theta_t(\mathbf{p}_{j'}^{(l)}, \mathbf{p}_j^{(l)})$ cancel out in this sum. As a consequence,

$$N \sum_{j=1}^{j_0} \mathbf{p}_j - 2\pi \sum_{j=1}^{j_0} j + j_0(n+1)\pi + \sum_{j=1}^{j_0} \sum_{k>j_0} \Theta_t(\mathbf{p}_j^{(l)}, \mathbf{p}_k^{(l)}) \leq 0.$$

This time, the inequality can be extended by continuity and we obtain

$$N \sum_{j=1}^{j_0} \mathbf{p}_j - 2\pi \sum_{j=1}^{j_0} j + j_0(n+1)\pi + \sum_{j=1}^{j_0} \sum_{k>j_0} \Theta_t(\mathbf{p}_j^{(\infty)}, \mathbf{p}_k^{(\infty)}) \leq 0.$$

From Computation 1, we have

$$N j_0(\pi - \mu_t) - 2\pi \sum_{j=1}^{j_0} j + j_0(n+1)\pi + j_0(n - j_0) \cdot (2\mu_t - \pi) \leq 0.$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Thus,

$$j_0(n+1)\pi + j_0(n-j_0)\pi + (Nj_0 - 2j_0(n-j_0))(\pi - \mu_t) \leq 2\pi \sum_{j=1}^{j_0} j.$$

Since $\mu_t \leq \pi$ and that $2j_0(n-j_0) - Nj_0 = -2j_0^2 < 0$, this last inequality implies

$$j_0(n+1)\pi + j_0(n-j_0)\pi \leq 2\pi \sum_{j=1}^{j_0} j. \tag{14}$$

However, we have

$$\sum_{j=1}^{j_0} j \leq nj_0 - \sum_{j=1}^{j_0} j = nj_0 - \frac{j_0(j_0+1)}{2}. \tag{15}$$

As a consequence of equations (14) and (15),

$$\begin{aligned} j_0(n+1)\pi + j_0(n-j_0)\cdot\pi &\leq 2\pi n j_0 - j_0(j_0+1)\pi \\ (2n+1)j_0\pi - {j_0}^2\pi &\leq 2\pi n j_0 - {j_0}^2\pi - j_0\pi \\ j_0\pi &\leq -j_0\pi. \end{aligned}$$

Since this last inequality is impossible, this means that ζ_t has a minimum in I_t^n.

Characterization of the solutions with an analytic differential equation: Let us denote by $\mathbf{p}(t) = (\mathbf{p}_1(t), \ldots, \mathbf{p}_n(t))$, for all $t \in (0, \sqrt{2})$, the unique minimum of the function ζ_t in I_t^n. Let us denote by $t \mapsto \mathbf{s}(t) = (\mathbf{s}_1(t), \ldots, \mathbf{s}_n(t))$ the unique solution of the differential equation:

$$\mathbf{s}'(t) = -(H_t(\mathbf{s}_1(t), \ldots, \mathbf{s}_n(t)))^{-1} \cdot \left(\frac{\partial^2 \zeta_t}{\partial t \partial p_j}(\mathbf{s}_1(t), \ldots, \mathbf{s}_n(t))\right)_j, \tag{16}$$

such that $\mathbf{s}(t)$ is the minimum of the function ζ_t when $t = \sqrt{2}/2$, where H_t is the Hessian matrix of ζ_t (existence and uniqueness are provided by classical theorems on first-order nonlinear differential equations). Since this is an analytic differential equation, its solution $\mathbf{s}$ is analytic.

Let us rewrite equation (16):

$$H_t(\mathbf{s}_1(t), \ldots, \mathbf{s}_n(t)) \cdot \mathbf{s}'(t) = -\left(\frac{\partial^2 \zeta_t}{\partial t \partial p_j}(\mathbf{s}_1(t), \ldots, \mathbf{s}_n(t))\right)_j$$

$$\frac{\partial^2 \zeta_t}{\partial t \partial p_j}(\mathbf{s}_1(t), \ldots, \mathbf{s}_n(t)) + \sum_k \mathbf{s}_k'(t) \cdot \frac{\partial^2 \zeta_t}{\partial p_k \partial p_j}(\mathbf{s}_1(t), \ldots, \mathbf{s}_n(t)) = 0.$$

This means that for all j,

$$\frac{\partial \zeta_t}{\partial p_j}(\mathbf{s}_1(t), \ldots, \mathbf{s}_n(t))$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

is a constant. Since $\mathbf{s}(t)$ is the minimum of ζ_t when $t = \sqrt{2}/2$, this constant is zero. As a consequence, by uniqueness of the minimum of ζ_t for all t, $\mathbf{s}(t) = \mathbf{p}(t)$. This means that $t \mapsto \mathbf{p}(t)$ is analytic.

Antisymmetry of the solutions: For all t, j, since $\mathbf{p}(t)$ is the minimum of ζ_t, using equation (10) and the fact that κ_t is increasing, we can write successively the following equations:

$$N\mathbf{p}_{n-j+1} - 2\pi(n-j+1) + (n+1)\pi + \sum_k \Theta_k(\mathbf{p}_{n-j+1}, \mathbf{p}_{n-k+1}) = 0;$$

$$N\mathbf{p}_{n-j+1} + 2\pi j - (n+1)\pi + \sum_k \Theta_k(\mathbf{p}_{n-j+1}, \mathbf{p}_{n-k+1}) = 0;$$

$$-N\mathbf{p}_{n-j+1} - 2\pi j + (n+1)\pi - \sum_k \Theta_k(\mathbf{p}_{n-j+1}, \mathbf{p}_{n-k+1}) = 0;$$

$$-N\mathbf{p}_{n-j+1} - 2\pi j + (n+1)\pi + \sum_k \Theta_k(-\mathbf{p}_{n-j+1}, -\mathbf{p}_{n-k+1}) = 0.$$

This means that the sequence $(-\mathbf{p}_{n-j+1}(t))_j$ is a minimum for ζ_t, and as a consequence, for all j, $\mathbf{p}_{n-j+1} = -\mathbf{p}_j$.

The numbers $\mathbf{p}_{j(t)}$, $j \leq n$, are all distinct: Let us consider the function

$$\chi_t : \alpha \mapsto N\kappa_t(\alpha) + \sum_{k=1}^{n} \theta_t(\alpha, \boldsymbol{\alpha}_k(t)),$$

where for all k, $\boldsymbol{\alpha}_k(t)$ is equal to $\kappa_t^{-1}(\mathbf{p}_k(t))$. For all j, this function has value $\pi(2j-(n+1))$ in $\boldsymbol{\alpha}_j(t)$ (by the Bethe equations). The finite sequence $(\pi(2j-(n+1)))_j$ is increasing, thus it is sufficient to prove that the function χ_t is increasing. Its derivative is

$$\chi_t' : \alpha \mapsto N\kappa_t'(\alpha) + \sum_{k=1}^{n} \frac{\partial\theta_t}{\partial\alpha}(\alpha, \boldsymbol{\alpha}_k(t)).$$

Since $t \in (0, \sqrt{2})$, $\sin(\mu_t) < 0$, and thus this function is positive. As a consequence, χ_t is increasing. □

5.4. *Diagonalization of some Heisenberg Hamiltonian.* ▹ Now that we have proved that the system of Bethe equations has a unique solution, the Bethe ansatz effectively provides a candidate eigenvector and eigenvalue for $V_N(t)$ for each of the sub-spaces $\Omega_N^{(n)}$. In the following, we will focus on $t = \sqrt{2}$ and show that the candidate eigenvalue is indeed the maximal eigenvalue of $V_N(t)$ on the corresponding sub-space, for all t sufficiently close to $\sqrt{2}$ (let us remember that the analyticity of the solutions of Bethe equations will imply that this is also the case away from $\sqrt{2}$). For this purpose, we will need first to introduce a Hamiltonian H_N which we diagonalize completely, following Lieb, Schultz and Mattis [**LSM61**]. The term 'Hamiltonian' is only borrowed from the article [**LSM61**]. In this paper, it is sufficient to see H_N as a matrix acting on Ω_N.

5.4.1. *Bosonic creation and annihilation operators.* ▹ The Hamiltonian H_N is expressed by elementary operators which are defined in this section. We also prove the usefulness of

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

their properties as Bosonic creation and annihilation operators. The Hamiltonian itself will be defined in the next section.

Let us recall that $\Omega_N = \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$. In this section, for the purpose of notation, we identify $\{1, \ldots, N\}$ with $\mathbb{Z}/N\mathbb{Z}$.

Notation 11. Let us denote by a and a^* the matrices in $\mathcal{M}_2(\mathbb{C})$ equal to

$$a := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad a^* := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

For all $j \in \mathbb{Z}/N\mathbb{Z}$, we denote by a_j (*creation* operator at position *j*) and a_j^* (*annihilation* operator at position *j*) the matrices in $\mathcal{M}_{2^N}(\mathbb{C})$ equal to

$$a_j := \mathrm{id} \otimes \cdots \otimes a \otimes \cdots \otimes \mathrm{id}, \quad a_j^* := \mathrm{id} \otimes \cdots \otimes a^* \otimes \cdots \otimes \mathrm{id}.$$

where id is the identity, and a acts on the *j*th copy of $\mathbb{C}^2$.

In other words, the image of a vector $|\boldsymbol{\epsilon}_1 \cdots \boldsymbol{\epsilon}_N\rangle$ in the basis of Ω_N by a_j (respectively, a_j^*) is as follows:

- if $\boldsymbol{\epsilon}_j = 0$ (respectively, $\boldsymbol{\epsilon}_j = 1$), then the image vector is $\mathbf{0}$;
- if $\boldsymbol{\epsilon}_j = 1$ (respectively, $\boldsymbol{\epsilon}_j = 0$), then the image vector is $|\boldsymbol{\eta}_1 \cdots \boldsymbol{\eta}_N\rangle$ such that $\boldsymbol{\eta}_j = 0$ (respectively, $\boldsymbol{\eta}_j = 1$) and for all $k \neq j$, $\boldsymbol{\eta}_k = \boldsymbol{\epsilon}_k$.

Remark 4. The term creation (respectively, annihilation) refers to the fact that for two elements $\boldsymbol{\epsilon}$, $\boldsymbol{\eta}$ of the basis of Ω_N, $a_j[\boldsymbol{\epsilon}, \boldsymbol{\eta}] \neq 0$ (respectively, $a_j^*[\boldsymbol{\epsilon}, \boldsymbol{\eta}] \neq 0$) implies that $|\boldsymbol{\eta}|_1 = |\boldsymbol{\epsilon}|_1 + 1$ (respectively, $|\boldsymbol{\eta}|_1 = |\boldsymbol{\epsilon}|_1 - 1$). If we think of 1 symbols as particles, this operator acts by creating (respectively, annihilating) a particle.

LEMMA 6. *The matrices a_j and a_j^* satisfy the following properties, for all j and $k \neq j$:*

(1) $a_j a_j^* + a_j^* a_j = \mathrm{id}$;

(2) $a_j^2 = {a_j^*}^2 = 0$;

(3) a_j, a_j^* *commute both with* a_k *and* a_k^*.

Proof. (1) By straightforward computation, we get

$$aa^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$a^*a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus $aa^* + a^*a$ is the identity of $\mathbb{C}^2$. As a consequence, for all *j*,

$$a_j a_j^* + a_j^* a_j = \mathrm{id} \otimes \cdots \otimes \mathrm{id},$$

which is the identity of Ω_N.

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

(2) The second set of equalities comes directly from

$$a^2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$(a^*)^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

(3) The last set derives from the fact that any operator on $\mathbb{C}^2$ commutes with the identity. □

5.4.2. *Definition and properties of the Heisenberg Hamiltonian*

Notation 12. Let us denote by H_N the matrix in $\mathcal{M}_{2^N}(\mathbb{C})$ defined as

$$H_N = \sum_{j \in \mathbb{Z}/N\mathbb{Z}} (a_j^* a_{j+1} + a_j a_{j+1}^*).$$

LEMMA 7. *This matrix H_N is non-negative, symmetric and for all n, its restriction to $\Omega_N^{(n)}$ is irreducible.*

The proof of Lemma 7 is similar to that of Lemma 2, following the interpretation of the action of H_N described in Remark 5.

Remark 5. For all j, $a_j^* a_{j+1} + a_j a_{j+1}^*$ acts on a vector $\boldsymbol{\epsilon}$ on the basis of Ω_N by exchanging the symbols in positions j and $j+1$ if they are different. If they are not, the image of $\boldsymbol{\epsilon}$ by this matrix is *0*. As a consequence, for two vectors $\boldsymbol{\epsilon}$ and $\boldsymbol{\eta}$ on the basis of Ω_N, $H_N[\boldsymbol{\epsilon}, \boldsymbol{\eta}] \neq 0$ if and only if $\boldsymbol{\eta}$ is obtained from $\boldsymbol{\epsilon}$ by exchanging a 1 symbol of $\boldsymbol{\epsilon}$ with a 0 in its neighborhood. The Hamiltonian H_N thus corresponds to H in [**DGHMT18**] for $\Delta = 0$.

5.4.3. *Fermionic creation and annihilation operators.* ▹ The reason for defining the Hamiltonian, as in the last section, comes from the way it is obtained in the first place. We will not provide details on this here, and only rewrite the definition using other operators, called Fermionic creation and annihilation operators. This rewriting, exposed in the present section, will help diagonalize the matrix H_N.

Notation 13. Let us denote by σ the matrix of $\mathcal{M}_2(\mathbb{C})$ defined as

$$\sigma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Let us denote, for all $j \in \mathbb{Z}/N\mathbb{Z}$, by c_j and c_j^* the matrices

$$c_j = \sigma \otimes \cdots \sigma \otimes a \otimes \mathrm{id} \otimes \cdots \otimes \mathrm{id}, \quad c_j^* = \sigma \otimes \cdots \sigma \otimes a^* \otimes \mathrm{id} \otimes \cdots \otimes \mathrm{id}.$$

Let us recall that two matrices P, Q anti-commute when $PQ = -QP$.

LEMMA 8. *These operators verify the following properties for all j and $k \neq j$:*

(1) $c_j c_j^* + c_j^* c_j = \mathrm{id}$;
(2) c_j^* *and* c_j *anti-commute with both* c_k^* *and* c_k;
(3) $a_{j+1}^* a_j = -c_{j+1}^* c_j$ *and* $a_j^* a_{j+1} = -c_j^* c_{j+1}$.

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Proof. (1) Since $\sigma^2 = \mathrm{id}$, for all j,

$$c_j c_j^* + c_j^* c_j = a_j a_j^* + a_j^* a_j.$$

From Lemma 6, we know that this operator is equal to identity.

(2) We can assume without loss of generality that $j < k$. Let us prove that c_j anti-commutes with c_k (the other cases are similar):

$$c_j c_k = \mathrm{id} \otimes \cdots \otimes a\sigma \otimes \sigma \otimes \cdots \otimes \sigma \otimes \sigma a \otimes \mathrm{id} \otimes \cdots \otimes \mathrm{id},$$
$$c_j c_k = \mathrm{id} \otimes \cdots \otimes \sigma a \otimes \sigma \otimes \cdots \otimes \sigma \otimes a\sigma \otimes \mathrm{id} \otimes \cdots \otimes \mathrm{id}.$$

Hence it is sufficient to see

$$\sigma a = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix},$$

$$a\sigma = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = -\sigma a.$$

(3) Let us prove the first equality (the other one is similar),

$$c_{j+1}^* c_j = \mathrm{id} \otimes \cdots \otimes \mathrm{id} \otimes \sigma a \otimes a^* \otimes \mathrm{id} \otimes \cdots \otimes \mathrm{id}.$$

We have just seen in the last point that $\sigma a = -a$. As a consequence, $c_{j+1}^* c_j = -a_{j+1}^* a_j$. □

5.4.4. *Action of a symmetric orthogonal matrix.* ▹ In this section, we consider some operators derived from the Fermionic operators introduced in the last section by the action of a symmetric and orthogonal matrix. We also prove some of their properties. Using these new operators, we define a family of vectors that we prove in the next section to be a basis of eigenvectors for H_N.

Let us denote by c^* the vector $(c_1^*, \ldots, c_N^*)$ and c^t is the transpose of the vector $(c_1, \ldots, c_N)$. Let us consider a symmetric and orthogonal matrix $U = (u_{i,j})_{i,j}$ in $\mathcal{M}_N(\mathbb{R})$ and denote by b and b^* the matrices

$$b = U \cdot c^t = (b_1, \ldots, b_N), \quad b^* = c^* \cdot U^t = (b_1^*, \ldots, b_N^*).$$

Notation 14. For all $\alpha \in \{0,1\}^N$, we set

$$\psi_\alpha = (b_1^*)^{\alpha_1} \ldots (b_N^*)^{\alpha_N} \cdot \boldsymbol{\nu}_N,$$

where $\boldsymbol{\nu}_N = |0, \ldots, 0\rangle$.

LEMMA 9. *For all j and $k \neq j$:*

(1) *b_j and b_j^* anti-commute with both b_k and b_k^* and $b_j b_j^* + b_j^* b_j = \mathrm{id}$;*
(2) *for all $\alpha \in \{0,1\}^N$, $\psi_\alpha \neq 0$;*
(3) *for all j and α, we have:*
   (i) *$b_j^* b_j \psi_\alpha = 0$ if $\alpha_j = 0$;*
   (ii) *$b_j^* b_j \psi_\alpha = \psi_\alpha$ if $\alpha_j = 1$.*

Proof. (1) *Anti-commutation relations:* Let us prove that b_j and b_k^* anti-commute (the other statements of the first point have a similar proof). We rewrite the definition of b_j and b_k^*:

$$b_j = \sum_i u_{i,j} c_i \quad \text{and} \quad b_k^* = \sum_i u_{k,i} c_i^* = \sum_i u_{i,k} c_i^*.$$

Thus,

$$b_j b_k^* = \sum_i \sum_{l \neq i} u_{i,j} u_{l,k} c_i c_l^* + \sum_i u_{i,j} u_{i,k} c_i c_i^*.$$

From Lemma 8,

$$b_j b_k^* = -\sum_l \sum_{i \neq l} u_{i,j} u_{l,k} c_l^* c_i + \sum_i u_{i,j} u_{i,k} (\mathrm{id} - c_i^* c_i).$$

Since the matrix U is orthogonal,

$$b_j b_k^* = -\sum_l \sum_{i \neq l} u_{i,j} u_{l,k} c_l^* c_i - \sum_i u_{i,j} u_{i,k} c_i^* c_i = -b_k^* b_j.$$

Let us notice that this step is the reason why we use the operators c_i instead of the operators a_i.

(2) For all k, $b_k^* = \sum_l u_{k,l} a_l^*$. As a consequence, for a sequence $k_1, \ldots, k_s$,

$$b_{k_1}^* \ldots b_{k_s}^* \cdot \boldsymbol{v}_N = \sum_{l_1} \cdots \sum_{l_s} \left(\prod_{j=1}^{s} u_{k_j, l_j} \right) \left(\prod_{j=1}^{s} a_{l_j}^* \right) . \boldsymbol{v}_N.$$

Since $(a^*)^2 = 0$, the sum can be considered on the integers $l_1, \ldots, l_s$ such that they are two by two distinct. The operator $a_{l_1}^* \ldots a_{l_s}^*$ acts on $\boldsymbol{v}_N$ by changing the 0 on positions $l_1, \ldots, l_s$ into symbols 1. The coefficient of the image of $\boldsymbol{v}_N$ by this operator in the vector $b_{k_1}^* \ldots b_{k_s}^* \cdot \boldsymbol{v}_N$ is thus:

$$\sum_{\sigma \in \Sigma_s} \prod_{j=1}^{s} u_{k_j, l_{\sigma(j)}}.$$

If this coefficient was equal to zero for all σ, it would mean that all $s \times s$ sub-matrices of U have determinant equal to zero, which is impossible since U is orthogonal, and thus invertible (this derives from iterating Laplace expansion of the determinant of U). As a consequence, none of the vectors ψ_α is equal to zero.

(3) When $\alpha_j = 0$, from the fact that when $j \neq k$, b_j and b_k^* anti-commute, we get that

$$b_j \psi_\alpha = (-1)^{|\alpha|_1} (b_1^*)^{\alpha_1} \ldots (b_N^*)^{\alpha_N} b_j \boldsymbol{v}_N,$$

and $b_j \boldsymbol{v}_N = \mathbf{0}$, since for all j, $a_j \boldsymbol{v}_N = \mathbf{0}$. As a consequence, $b_j^* b_j \boldsymbol{v}_N = \mathbf{0}$. When $\alpha_j = 1$, by the anti-commutation relations,

$$b_j^* b_j \psi_\alpha = (b_1^*)^{\alpha_1} \cdots (b_{j-1}^*)^{\alpha_{j-1}} b_j^* b_j b_j^* (b_{j+1}^*)^{\alpha_{j+1}} \cdots (b_N^*)^{\alpha_N} \boldsymbol{v}_N,$$

since the coefficients -1 introduced by anti-commutation are canceled out by the fact that we use it for b_j and b^*. From the first point,

$$b_j^* b_j \psi_\alpha = (b_1^*)^{\alpha_1} \cdots (b_{j-1}^*)^{\alpha_{j-1}} b_j^* (\mathrm{id} - b_j^* b_j)(b_{j+1}^*)^{\alpha_{j+1}} \cdots (b_N^*)^{\alpha_N}.$$

$$\begin{aligned} b_j^* b_j \psi_\alpha &= (b_1^*)^{\alpha_1} \cdots (b_{j-1}^*)^{\alpha_{j-1}} b_j^* (\mathrm{id} - b_j^* b_j)(b_{j+1}^*)^{\alpha_{j+1}} \cdots (b_N^*)^{\alpha_N} \\ &= \psi_\alpha - (b_1^*)^{\alpha_1} \cdots (b_N^*)^{\alpha_N} b_j \boldsymbol{\nu}_N \\ &= \psi_\alpha. \end{aligned}$$

□

5.4.5. *Diagonalization of the Hamiltonian.* ▹ In this section, we diagonalize H_N using the last section. Since we will only use its eigenvalues, we formulate the following.

THEOREM 5. *The eigenvalues of H_N are exactly the numbers*

$$2 \sum_{\alpha_j = 1} \cos\left(\frac{2\pi j}{N}\right),$$

for $\alpha \in \{0, 1\}^N$.

Proof. (1) *Rewriting H_N*: From Lemma 8, we can write H_N as

$$H_N = \sum_j c_j^* c_{j+1} + c_{j+1}^* c_j.$$

The Hamiltonian H_N can be then rewritten as $H_N = c^* M c^t$, where M is the matrix defined by blocks

$$M = \frac{1}{2} \begin{pmatrix} \mathbf{0} & \mathrm{id} & & \mathrm{id} \\ \mathrm{id} & \ddots & \ddots & \\ & \ddots & \ddots & \mathrm{id} \\ \mathrm{id} & & \mathrm{id} & \mathbf{0} \end{pmatrix},$$

where id is the identity matrix on $\mathbb{C}^2$, and $\mathbf{0}$ is the null matrix. Let us denote by M' the matrix of $\mathcal{M}_{2^N}(\mathbb{R})$ obtained from M by replacing $\mathbf{0}$, id by 0, 1.

(2) *Diagonalization of M*: The matrix M' is symmetric and thus can be diagonalized in $\mathcal{M}_{2^N}(\mathbb{R})$ in an orthogonal basis. It is rather straightforward to see that the vectors ψ^k, $k \in \{0, \ldots, N-1\}$ form an orthonormal family of eigenvectors of M' for the eigenvalue $\lambda_k = \cos(2\pi k/N)$, where for all $j \in \{1, \ldots, N\}$,

$$\psi_j^k = \sqrt{\frac{2}{N}} \left(\sin\left(\frac{2\pi kj}{N}\right), \cos\left(\frac{2\pi kj}{N}\right)\right).$$

This comes from the equalities

$$\begin{aligned} \cos(x-y) + \cos(x+y) &= 2\cos(x)\cos(y), \\ \sin(x-y) + \sin(x+y) &= 2\cos(x)\sin(y), \end{aligned}$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

applied to $x = k(j-1)$ and $y = k(j+1)$. This family of vectors is linearly independent, since the Vandermonde matrix with coefficients $e^{2\pi kj/N}$ is invertible. As a consequence, one can write

$$U'M'U'^t = D',$$

where D' is the diagonal matrix whose diagonal coefficients are the numbers λ_k, and U' is the orthogonal matrix given by the vectors ψ^k. Replacing any coefficient of these matrices by the product of this coefficient with the identity, one gets an orthogonal matrix U and a diagonal one D such that

$$UMU^t = D.$$

(3) *Some eigenvectors of* H_N: Let us consider the vectors ψ_α constructed in §5.4.4 for the matrix U of the last point, which is symmetric and orthogonal. From the expression of H_N, we get that

$$H_N\psi_\alpha = \left(2 \sum_{j:\alpha_j=1} \cos\left(\frac{2\pi j}{N}\right)\right) \cdot \psi_\alpha.$$

Since ψ_α is non-zero, this is an eigenvector of H_N.

(4) *The family* (ψ_α) *is a basis of* Ω_N: From cardinality of this family (the number of possible α, equal to 2^N), it is sufficient to prove that this family is linearly independent. For this purpose, let us assume that there exists a sequence $(x_\alpha)_{\alpha\in\{0,1\}}$ such that

$$\sum_{\alpha\in\{0,1\}^N} x_\alpha \cdot \psi_\alpha = \mathbf{0}.$$

We apply first $b_1^*b_1 \cdots .b_N^*b_N$ and get that $x_{(1,\ldots,1)}\psi_{(1,\ldots,1)} = 0$, and thus $x_{(1,\ldots,1)} = 0$ (by Lemma 9, the vector $\psi_{(1,\ldots,1)}$ is not equal to zero). Then we apply successively the operators $\prod_{j\neq k} b_j^*b_j$ for all k, and obtain that for all $\alpha \in \{0,1\}_N$ such that $|\alpha|_1 = N-1$, $x_\alpha = 0$. By repeating this argument, we obtain that all the coefficient x_α are null. As a consequence, $(\psi_\alpha)_\alpha$ is a base of eigenvectors for H_H, and the eigenvalues obtained in the last point cover all the eigenvalues of H_N. □

5.5. *Identification.* ▹ In this section, we use the diagonalization sub-spaces $\Omega_N^{(2n+1)}$ for $t \in (0, \sqrt{2})$ and $2n+1 \leq N/2$. of H_N and a commutation relation between H_N and $V_N(\sqrt{2})$ to prove that the candidate eigenvalue is the maximal eigenvalue of $V_N(t)$ for each of the The proofs for the following two lemmas can be found in [**DGHMT18**] (respectively Lemma 5.1 and Theorem 2.3). In Lemma 10, our notation H_N corresponds to their notation H for $\Delta = 0$, and $V_N(\sqrt{2})$ corresponds to V for $\Delta = 0$. In Lemma 11, the equations $(E_j)[\sqrt{2}, n, N]$ correspond to their (BE), ψ to ψ for $\Delta = 0$.

LEMMA 10. *For all* $N \geq 1$*, the Hamiltonian* H_N *and* $V_N(\sqrt{2})$ *commute*

$$H_N \cdot V_N(\sqrt{2}) = V_N(\sqrt{2}) \cdot H_N.$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

LEMMA 11. *For all N and* $n \leq N$, *let us denote by* $(\boldsymbol{p}_j)_j$ *the solution of the system of equations* $(E_j)[\sqrt{2}, n, N]$, *then denoting* $\psi \equiv \psi_{\sqrt{2},n,N}(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)$,

$$H_N \cdot \psi = \left(2\sum_{k=1}^{n} \cos(\boldsymbol{p}_k)\right) \cdot \psi.$$

THEOREM 6. *For all N and* $2n+1 \leq N/2$, *and* $t \in (0, \sqrt{2})$,

$$\lambda_{2n+1,N}(t) = \Lambda_{2n+1,N}(t)[\boldsymbol{p}_1(t), \ldots, \boldsymbol{p}_{2n+1}(t)].$$

Proof. (1) *The Bethe vector is* $\neq 0$ *for t in a neighborhood of* $\sqrt{2}$:

- *Limit of the Bethe vector in* $\boldsymbol{\sqrt{2}}$: Let us denote by $(\mathbf{p}_j(t))_j$ the solution of the system of equations $(E_j)[t, 2n+1, N]$.

  Let us recall [Theorem 3] that for all t, and $\boldsymbol{\epsilon}$ in the canonical basis of Ω_N,

$$\psi_{t,2n+1,N}(\mathbf{p}_1(t), \ldots, \mathbf{p}_{2n+1}(t))[\boldsymbol{\epsilon}] = \sum_{\sigma \in \Sigma_{2n+1}} C_\sigma(t)[\mathbf{p}(t)] \prod_{k=1}^{2n+1} e^{i\mathbf{p}_{\sigma(k)(t)} \cdot q_k[\boldsymbol{\epsilon}]}.$$

  This expression admits a limit when $t \to \sqrt{2}$, given by

$$\sum_{\sigma \in \Sigma_{2n+1}} C_\sigma(\sqrt{2})[\mathbf{p}(\sqrt{2})] \prod_{k=1}^{2n+1} e^{i\mathbf{p}_{\sigma(k)(\sqrt{2})} \cdot q_k[\boldsymbol{\epsilon}]},$$

  where $(\mathbf{p}_k(\sqrt{2}))_k$ is solution of the system of equations $(E_k)[\sqrt{2}, 2n+1, N]$.

- *The expression* $\epsilon(\sigma)C_\sigma(\sqrt{2})[\mathbf{p}(\sqrt{2})]$ *is independent from* σ: Indeed, from the definition of $C_\sigma(t)$ we have

$$\prod_{1 \leq k < l \leq 2n+1} (1 + e^{i(\mathbf{p}_{\sigma(k)}(\sqrt{2}) + \mathbf{p}_{\sigma(l)}(\sqrt{2}))})$$
$$= \prod_{1 \leq \sigma^{-1}(k) < \sigma^{-1}(l) \leq 2n+1} (1 + e^{i(\mathbf{p}_k(\sqrt{2}) + \mathbf{p}_l(\sqrt{2}))})$$
$$= \prod_{1 \leq k < l \leq 2n+1} (1 + e^{i(\mathbf{p}_k(\sqrt{2}) + \mathbf{p}_l(\sqrt{2}))}).$$

  Indeed, for all $l \neq k$, one of the conditions $\sigma^{-1}(k) < \sigma^{-1}(l)$ or $\sigma^{-1}(l) < \sigma^{-1}(k)$ is verified, exclusively. This means that $(1 + e^{i(\mathbf{p}_k(\sqrt{2}) + \mathbf{p}_l(\sqrt{2}))})$ appears exactly once in the product for each l, k such that $l \neq k$.

- *The terms* $(1 + e^{i(\mathbf{p}_k(\sqrt{2}) + \mathbf{p}_l(\sqrt{2}))})$, $k \neq l$, *are all non-zero:* Indeed, none of the $\mathbf{p}_k(\sqrt{2}) + \mathbf{p}_l(\sqrt{2})$ can be equal to $\pm\pi$. This comes from the fact that the system of Bethe equations $(E_k)[\sqrt{2}, 2n+1, N]$ has a unique simple solution given by

$$\mathbf{p}_k(\sqrt{2}) = \frac{\pi}{N}\left(2k - \frac{(2n+1+1)}{2}\right) = \frac{2\pi}{N}(k - (n+1)).$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

These numbers are bounded respectively above and below by

$$\mathbf{p}_1(\sqrt{2}) = -2\pi \frac{n-1}{N}, \quad \mathbf{p}_n(\sqrt{2}) = 2\pi \frac{n-1}{N}.$$

Since $2n+1 \leq N/2$, these numbers are in $[-\pi/2, \pi/2]$, and the possible sums of two different values of these numbers is in $]-\pi, \pi[$.

- *The limit of Bethe vectors is non-zero:* As a consequence of the last points, we have that the $\boldsymbol{\epsilon}$ coordinate of the limit of Bethe vectors when $t \to \sqrt{2}$ is, up to a non-zero constant,

$$\sum_{\sigma \in \Sigma_{2n+1}} \epsilon(\sigma) \cdot \prod_{k=1}^{2n+1} e^{i\mathbf{p}_{\sigma(k)(\sqrt{2})} \cdot q_k[\boldsymbol{\epsilon}]},$$

which is the determinant of the matrix $(e^{i\mathbf{p}_{\sigma(k)(\sqrt{2})} \cdot Here, q_l[\boldsymbol{\epsilon}]})_{k,l}$, which is a submatrix of the matrix $(e^{is_k \cdot s'_l})_{k,l \in [\![1,N]\!]}$, where (s_k) is a sequence of distinct numbers in $]-\pi/2, \pi/2[$ such that for all $k \leq 2n+1$,

$$s_k = \mathbf{p}_{\sigma(k)(\sqrt{2})},$$

and $(s'_l)_l$ is a sequence of distinct integers such that for all $l \leq n$,

$$s'_l = q_l[\boldsymbol{\epsilon}].$$

If the determinant is non-zero, then the sum above is non-zero. This is the case since this last matrix is obtained from the Vandermonde matrix $(e^{is_k . l})_{k,l \in [\![1,N]\!]}$, whose determinant is

$$\prod_{k<l} (e^{is_l} - e^{is_k}) \neq 0,$$

by a permutation of the columns.

(2) *From the Hamiltonian to the transfer matrix:*

- *Eigenvector of $V_N(\sqrt{2})$ and H_N*: Since the limit of Bethe vectors is not equal to zero, and it satisfies an equation which is the 'limit' of equations which make the Bethe vectors candidate eigenvectors, it is an eigenvector of the matrix $V_N(\sqrt{2})$. It is also an eigenvector of the Hamiltonian H_N, for the eigenvalue

$$2\bigg(\sum_{k=1}^{n-1} \cos\bigg(\frac{2\pi k}{N}\bigg) + \sum_{k=N-n+1}^{N} \cos\bigg(\frac{2\pi k}{N}\bigg)\bigg). \tag{17}$$

This is a consequence of Lemma 11, since for all j, $N\mathbf{p}_j(\sqrt{2}) = 2\pi(j-(n+1))$, the eigenvalue is

$$\begin{aligned} 2\sum_{k=1}^{2n+1} \cos(\mathbf{p}_k(\sqrt{2})) &= 2\sum_{k=1}^{n} \cos(\mathbf{p}_k(\sqrt{2})) + 2\sum_{k=n+1}^{2n+1} \cos(N - \mathbf{p}_k(\sqrt{2})) \\ &= 2\bigg(\sum_{k=1}^{n-1} \cos\bigg(\frac{2\pi k}{N}\bigg) + \sum_{k=N-n+1}^{N} \cos\bigg(\frac{2\pi k}{N}\bigg)\bigg). \end{aligned}$$

- *Comparison with the other eigenvalues of* H: From Theorem 5, we know that the number given by the expression in equation (17) is the largest eigenvalue of H_N on $\Omega_N^{(2n+1)}$. Indeed, it is straightforward that ψ_α is in $\Omega_N^{(2n+1)}$ if and only if the number of k such that $\alpha_k = 1$ is $2n+1$. The sum in the statement of Theorem 5 is maximal among these sequences when

$$\alpha_1 = \cdots = \alpha_{n-1} = 1 = \alpha_{N-n+1} = \cdots = \alpha_N$$

  and the other α_k are equal to 0.
- *Identification:* As a consequence, from Perron–Frobenius theorem, the limit of Bethe vectors in $\sqrt{2}$ is positive, thus this is also true for t sufficiently close to $\sqrt{2}$. From the same theorem, the Bethe vector is associated to the maximal eigenvalue of $V_N(t)$. As a consequence, the Bethe value $\Lambda_{2n+1,N}(t)[\mathbf{p}_1(t), \cdots \mathbf{p}_{2n+1}(t)]$ is equal to the largest eigenvalue $\lambda_{2n+1,N}(t)$ of $V_N(t)$ on $\Omega_N^{(2n+1)}$ for these values of t. Since these two functions are analytic in t (by the implicit functions theorem on the characteristic polynomial, using the fact that the largest eigenvalue is simple), one can identify these two functions on the interval $(0, \sqrt{2})$. □

6. *Asymptotic properties of Bethe roots*

▷ At this point, we have an expression of the largest eigenvalue of $V_N(1)$ on each of the sub-spaces $\Omega_N^{(2n+1)}$ with $2n+1 \leq N/2$. To obtain the entropy of X^s, we need to understand how these expressions behave asymptotically, when n and N tend towards infinity—we can in fact assume that n/N tends towards some d. For this, we need to understand how Bethe roots behave asymptotically. This is what we will do in this section.

Let us fix some $d \in [0, 1/2]$, and $(N_k)_k$ and $(n_k)_k$ some sequences of integers such that for all k, $n_k \leq N_k/2 + 1$ and $n_k/N_k \to d$. In this section, we study the asymptotic behavior of the sequences $(\boldsymbol{\alpha}_j^{(k)}(t))_j$, where $t \in (0, \sqrt{2})$,

$$(\mathbf{p}_j^{(k)}(t))_j \equiv (\kappa_t(\boldsymbol{\alpha}_j^{(k)}(t)))_j$$

is solution of the system of Bethe equations $(E_j)[t, n_k, N_k]$, $j \leq n_k$, when k tends towards $+\infty$. For this purpose, we introduce in §6.1 the counting functions $\xi_t^{(k)}$ associated to the corresponding Bethe roots. The term 'counting function' refers to the fact that between two Bethe roots, the function increases by a constant, and thus 'counts' Bethe roots. In other words, these functions represent the distribution of Bethe roots in the real line. In §6.2, we prove that the sequence of functions $(\xi_t^{(k)})_k$ converges uniformly on any compact to a function $\boldsymbol{\xi}_{t,d}$. In §6.3, we then prove the following, which will be used in §7 to compute the entropy of square ice: for all function $f : (0, +\infty) \to (0, +\infty)$ which is continuous, decreasing and integrable,

$$\frac{1}{N_k} \sum_{j=\lceil n_k/2 \rceil + 1}^{n_k} f(\boldsymbol{\alpha}_j^{(k)}(t)) \to \int_0^{\boldsymbol{\xi}_{t,d}^{-1}(d)} f(\alpha)\boldsymbol{\xi}_{t,d}(\alpha)\, d\alpha.$$

6.1. *The counting functions associated to Bethe roots.*

In this section, we define the counting functions and prove some additional preliminary facts on the auxiliary functions

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

θ_t and κ_t that we will use in the following [§6.1.1]. We prove also that the number of Bethe roots vanishes as one get close to $\pm\infty$, with a speed that does not depend on k [§6.1.2].

6.1.1. *Definition*

Notation 15. For all $t \in (0, \sqrt{2})$ and all integer k, let us denote by $\xi_t^{(k)} : \mathbb{R} \to \mathbb{R}$ the *counting function* defined as follows:

$$\xi_t^{(k)} : \alpha \mapsto \left(\frac{1}{2\pi} \kappa_t(\alpha) + \frac{n_k + 1}{2N_k} + \frac{1}{2\pi N_k} \sum_j \theta_t(\alpha, \boldsymbol{\alpha}_j^{(k)}(t)) \right).$$

Fact 1. Let us notice some properties of these functions, that we will use in the following.

(1) By Bethe equations, for all j and k,

$$\xi_t^{(k)}(\boldsymbol{\alpha}_j^{(k)}(t)) = \frac{j}{N_k} \equiv \rho_j^{(k)}.$$

(2) For all k, t, the derivative of $\xi_t^{(k)}$ is the function

$$\alpha \mapsto \frac{1}{2\pi} \kappa_t'(\alpha) + \frac{1}{2\pi N_k} \sum_j \frac{\partial \theta_t}{\partial \alpha}(\alpha, \boldsymbol{\alpha}_j(t)) > 0.$$

Indeed, this comes directly from the fact that $\mu_t \in (\pi/2, \pi)$. As a consequence, the counting functions are increasing.

We will use also the following proposition.

PROPOSITION 8. *We have the following limits for the functions κ_t and θ_t on the border of their domains:*

$$\lim_{+\infty} \kappa_t = - \lim_{-\infty} \kappa_t = \pi - \mu_t$$

and that for all $\beta \in \mathbb{R}$,

$$\lim_{+\infty} \theta_t(\alpha, y) = - \lim_{-\infty} \theta_t(\alpha, y) = 2\mu_t - \pi.$$

Proof. Let us prove this property for κ_t, where the limits for θ_t are obtained applying the same reasoning. Let us recall that for all $\alpha \in \mathbb{R}$,

$$\kappa_t'(\alpha) = \frac{\sin(\mu_t)}{\cosh(\alpha) - \cos(\mu_t)}.$$

Since this function is positive, κ_t is increasing, and thus admits a limit in $\pm\infty$. Since κ_t' is integrable, these limits are finite. Since for all α,

$$e^{i\kappa_t(\alpha)} = \frac{e^{i\mu_t} - e^{\alpha}}{e^{i\mu_t + \alpha} - 1},$$

and the limit of this expression when α tends to $+\infty$ is $-e^{-i\mu_t}$, then there exists some $k \in \mathbb{Z}$ such that

$$\lim_{+\infty} \kappa_t = 2k\pi + \pi - \mu_t$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Since κ_t is a bijective map from $\mathbb{R}$ to I_t [Proposition 7], then $k = 0$. Thus we have

$$\lim_{+\infty} \kappa_t = \pi - \mu_t.$$

The limit in $-\infty$ is obtained by symmetry. □

Notation 16. For any compact interval $I \subset \mathbb{R}$, we set

$$\mathcal{V}_I(\epsilon, \eta) = \{z \in \mathbb{C} : |\text{Im}(z)| < \eta, d(\text{Re}(z), I) < \epsilon\}.$$

6.1.2. *Rarefaction of Bethe roots near infinities.* For all k, t and $M > 0$, we set

$$P_t^{(k)}(M) \equiv \{j \in [\![1, n_k]\!] : \boldsymbol{\alpha}_j^{(k)}(t) \notin [-M, M]\}.$$

THEOREM 7. *For all $t \in (0, \sqrt{2})$, $\epsilon > 0$, there exists some $M > 0$ and k_0 such that for all $k \geq k_0$,*

$$\frac{1}{N_k}|P_t^{(k)}(M)| \leq \epsilon.$$

Idea of the proof: To prove this statement, we introduce a quantity q_t which represents roughly the density of Bethe roots near infinities, defined as a lim sup on pairs of integer and interval. It is sufficient to prove that $q_t = 0$ to prove the statement. We extract a sequence of integers $(\nu(k_l))_l$ and $(I_l)_l = ([-M_l, M_l])_l$ that realizes this lim sup. For these sequences, we bind from above and below the smallest (respectively, greatest) integer such that the corresponding Bethe root is greater than M_l (respectively, smaller than $-M_l$). Using Bethe equations and properties of κ_t and θ_t (boundedness and monotonicity), we prove a inequality relating these two bounds. Taking the limit $l \to +\infty$, we obtain an inequality that forces $q_t = 0$.

Proof. In this proof, we assume, to simplify the computations, that for all k, n_k is even, and we set $n_k = 2m_k$. However, similar arguments are valid for any sequence $(n_k)_k$. Moreover, if $d = 0$, the statement is trivial, and as a consequence, we assume in the remainder of the proof that $d > 0$. It is sufficient to prove then that for all $\epsilon > 0$, there exists some M and k_0 such that for all $k \geq k_0$,

$$\frac{1}{n_k}|P_t^{(k)}(M)| \leq \epsilon.$$

Formulation with superior limits: If $\limsup_m \boldsymbol{\alpha}_{n_k}^{(k)}$ is finite, then the Bethe roots are bounded independently from k (from below this comes from the asymmetry of $\boldsymbol{\alpha}^{(k)}$), and thus the statement is verified.

Let us thus assume that $\limsup_k \boldsymbol{\alpha}_{n_k}^{(k)} = +\infty$, meaning that there exists some $\nu : \mathbb{N} \to \mathbb{N}$ such that

$$\boldsymbol{\alpha}_{n_{\nu(k)}}^{(\nu(k))} \to +\infty.$$

Let us denote, for all k, t and $M > 0$, by $q_t^{(k)}(M)$ the proportion of positive Bethe roots $\boldsymbol{\alpha}_j^{(k)}$ that are greater than M. Since for all k, $\boldsymbol{\alpha}^{(k)}$ is an anti-symmetric and increasing sequence,

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

$\boldsymbol{\alpha}_j^{(k)} > 0$ implies that $j \geq m_k + 1$, and we define this proportion as

$$q_t^{(k)}(M) = \frac{1}{m_k}|\{j \in [\![m_k + 1, 2m_k]\!] : \boldsymbol{\alpha}_j^{(k)} \geq M\}|. \tag{18}$$

We also denote by $q_t(M) = \limsup_k q_t^{(\nu(k))}(M)$ and

$$q_t = \limsup_M q_t(M) \leq 1.$$

By construction, there exists an increasing sequence $(M_l)_l$ of real numbers and a sequence $(k_l)_l$ of integers such that for all $\epsilon > 0$, there exists some l_0 and for all $l \geq l_0$,

$$q_t - \epsilon < q_t(M_l) - \frac{\epsilon}{2} < q_t^{(\nu(k_l))}(M_l) < q_t(M_l) + \frac{\epsilon}{2} < q_t + \epsilon. \tag{19}$$

The proof of the statement reduces to prove that $q_t = 0$.

Bounds for the cutting integers sequence:

(1) *Lower bound:* As a consequence of equation (18) and inequalities in equation (19), for ϵ and l_0 as in the previous point,

$$(q_t + \epsilon)m_{\nu(k_l)} \geq |\{j \in [\![2m_{\nu(k_l)} + 1, 2m_{\nu(k_l)}]\!] : \boldsymbol{\alpha}_j^{(\nu(k_l))} \geq M_l\}|.$$

Moreover,

$$\begin{aligned} |\{j \in [\![m_{\nu(k_l)} + 1, 2m_{\nu(k_l)}]\!] : \boldsymbol{\alpha}_j^{(\nu(k_l))} < M_l\}| &= m_{\nu(k_l)} \\ &\quad - |\{j \in [\![m_{\nu(k_l)} + 1, 2m_{\nu(k_l)}]\!] : \boldsymbol{\alpha}_j^{(\nu(k_l))} \geq M_l\}| \\ &\geq m_{\nu(k_l)} \cdot (1 - q_t - \epsilon). \end{aligned}$$

Thus the ‘cutting integer’, denoted by σ_l, defined as the greatest j such that the associated Bethe root satisfies the inequality $\boldsymbol{\alpha}_j^{(\nu(k_l))} < M_l$, is bounded from below by

$$m_{\nu(k_l)} + m_{\nu(k_l)} \cdot \max(0, 1 - \epsilon - q_t) \geq \max(0, 2m_{\nu(k_l)}(1 - \epsilon - q_t)).$$

Since it is an integer, it is also greater than the integer $\underline{a}_l$ defined by

$$\underline{a}_l = \max(0, \lfloor 2m_{\nu(k_l)}.(1 - \epsilon - q_t)\rfloor).$$

(2) *Upper bound:* Let us also set $\overline{a}_l = \lfloor 2m_{\nu(k_l)}.(1 + \epsilon - q_t)\rfloor + 1$. For a similar reason, the cutting integer σ_l is smaller than $\overline{a}_l$. See a schema in Figure 9.

(3) *Another similar bound:* Moreover, since $l \geq l_0$, by definition of the sequence $(q_t^{(k)})$ and by the inequalities in equation (19),

$$q_t{}^{(\nu(k_l))}(M_{l_0}) \geq q_t{}^{(\nu(k_l))}(M_l) > q_t(M_l) - \frac{\epsilon}{2}.$$

As a consequence of a reasoning similar to the first point,

$$|\{j \in [\![m_{\nu(k_l)} + 1, 2m_{\nu(k_l)}]\!] : \boldsymbol{\alpha}_j^{(\nu(k_l))} < M_{l_0}\}| \geq m_{\nu(k_l)} \cdot (1 - q_t - \epsilon),$$

and thus for all $j \leq \underline{a}_l$, $\boldsymbol{\alpha}_j^{\nu(k_l)} < M_{l_0}$.

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

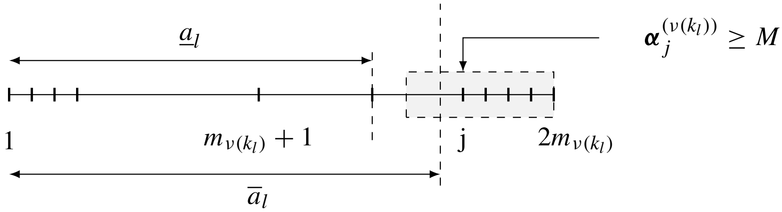


FIGURE 9. Illustration of the definition and lower bound of the cutting integer.

Inequality involving $\underline{a}_l$ *and* $\overline{a}_l$ *through Bethe equations:* By summing values of the counting function,

$$\sum_{k=\overline{a}_l}^{2m_{\nu(k_l)}} \xi_t^{(\nu(k_l))}(\boldsymbol{\alpha}_k^{(\nu(k_l))}) = \frac{1}{2\pi}\sum_{k=\overline{a}_l}^{2m_{\nu(k_l)}} \kappa_t(\boldsymbol{\alpha}_k^{(\nu(k_l))}) + \frac{2m_{\nu(k_l)}+1}{2N_{\nu(k_l)}}(2m_{\nu(k_l)} + 1 - \overline{a}_l)$$
$$+ \frac{1}{2\pi N_{\nu(k_l)}} \sum_{k=\overline{a}_l}^{2m_{\nu(k_l)}} \sum_{k'=1}^{2m_{\nu(k_l)}} \theta_t(\boldsymbol{\alpha}_k^{(\nu(k_l))}, \boldsymbol{\alpha}_{k'}^{(\nu(k_l))}).$$

By Bethe equations, we also have

$$\sum_{k=\overline{a}_l}^{2m_{\nu(k_l)}} \xi_t^{(\nu(k_l))}(\boldsymbol{\alpha}_k^{(\nu(k_l))}) = \frac{1}{N_{\nu(k_l)}}\sum_{k=\overline{a}_l}^{2m_{\nu(k_l)}} k = \frac{(2m_{\nu(k_l)} + \overline{a}_l)(2m_{\nu(k_l)} - \overline{a}_l + 1)}{2N_{\nu(k_l)}}.$$

As a direct consequence, and since θ_t is increasing in its first variable and $\theta_t(\alpha, \alpha) = 0$ for all α,

$$\frac{(2m_{\nu(k_l)} - \overline{a}_l + 1)(\overline{a}_l - 1)}{2N_{\nu(k_l)}}$$
$$= \frac{1}{2\pi}\sum_{k=\overline{a}_l}^{2m_{\nu(k_l)}} \kappa_t(\boldsymbol{\alpha}_k^{(\nu(k_l))}) + \frac{1}{2\pi N_{\nu(k_l)}} \sum_{k=\overline{a}_l}^{2m_{\nu(k_l)}} \sum_{k'=1}^{2m_{\nu(k_l)}} \theta_t(\boldsymbol{\alpha}_k^{(\nu(k_l))}, \boldsymbol{\alpha}_{k'}^{(\nu(k_l))})$$
$$\geq \frac{1}{2\pi}\sum_{k=\overline{a}_l}^{2m_{\nu(k_l)}} \kappa_t(\boldsymbol{\alpha}_k^{(\nu(k_l))}) + \frac{1}{2\pi N_{\nu(k_l)}} \sum_{k=\overline{a}_l}^{2m_{\nu(k_l)}} \sum_{k'<k} \theta_t(\boldsymbol{\alpha}_k^{(\nu(k_l))}, \boldsymbol{\alpha}_{k'}^{(\nu(k_l))}).$$

Using again the fact that θ_t is increasing in its first variable, we have

$$\theta_t(\boldsymbol{\alpha}_k^{(\nu(k_l))}, \boldsymbol{\alpha}_{k'}^{(\nu(k_l))}) \geq \theta_t(M_l, M_{l_0})$$

when $k \geq \overline{a}_l$ and $k' \leq \underline{a}_l$ (this is a consequence of the third bound proved in the last point). The terms corresponding to other pairs (k, k') are bounded from below by 0. Using these facts and the fact that κ_t is increasing,

$$\frac{(2m_{\nu(k_l)} - \overline{a}_l + 1)(\overline{a}_l - 1)}{2N_{\nu(k_l)}} \geq (2m_{\nu(k_l)} - \overline{a}_l + 1)\frac{1}{2\pi}\kappa_t(M_l)$$
$$+ (2m_{\nu(k_l)} - \overline{a}_l + 1)\frac{\underline{a}_l}{2\pi N_{\nu(k_l)}}\theta_t(M_l, M_{l_0}).$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Simplifying by $(2m_{\nu(k_l)} - \overline{a}_l + 1)$ (which is possible because it is positive),

$$\frac{(\overline{a}_l - 1)}{2N_{\nu(k_l)}} \geq \frac{1}{2\pi}\kappa_t(M_l) + \frac{\underline{a}_l}{2\pi N_{\nu(k_l)}}\theta_t(M_l, M_{l_0}).$$

We take the limit when $l \to +\infty$, and obtain, using the definitions of $\overline{a}_l$ and $\underline{a}_l$,

$$\frac{d}{2}(1 - \epsilon - q_t) \geq \frac{\pi - \mu_t}{2\pi} + \frac{d}{2}\frac{2\mu_t - \pi}{\pi}(1 - q_t).$$

Taking the limit when $\epsilon \to 0$,

$$\frac{d}{2}(1 - q_t) \geq \frac{\pi - \mu_t}{2\pi} + \frac{d}{2}\frac{2\mu_t - \pi}{\pi}(1 - q_t).$$

This inequality can be rewritten as

$$(1 - q_t)\left(\frac{d}{2} - \frac{d}{2}\frac{2\mu_t - \pi}{\pi}\right) \geq \frac{\pi - \mu_t}{2\pi}.$$

Finally, $1 - q_t \geq 1/2d \geq 1$, and thus since by definition q_t is non-negative, $q_t = 0$. □

6.2. *Convergence of the sequence of counting functions* $(\xi_t^{(k)})_k$. ▹ In this section, we prove that the sequence of functions $(\xi_t^{(k)})_k$ converges uniformly on any compact to a function $\boldsymbol{\xi}_{t,d}$. After some recalls on complex analysis [§6.2.1], we prove that if a subsequence of this sequence of functions converges on any compact of their domain towards a function, then this function verifies a Fredholm integral equation [§6.2.2], which is solved through Fourier analysis, and the solution is proved to be unique, in §6.2.3, by solving a similar equation verified by the derivative of this function. We prove in §6.2.4 that this fact implies that the sequence of counting functions converges to $\boldsymbol{\xi}_{t,d}$.

For all t, there exists $\tau_t > 0$ such that for all k, the functions κ_t, Θ_t and $\xi_t^{(k)}$ can be extended analytically on the set $\mathcal{I}_{\tau_t} := \{z \in \mathbb{C} : |\mathrm{Im}(z)| < \tau_t\} \subset \mathbb{C}$. For the purpose of notation, the extended functions are denoted by the same symbols as their restriction on $\mathbb{R}$.

6.2.1. *Some complex analysis background.* Let us recall some results of complex analysis that we will use in the remainder of this section. Let U be an open subset of $\mathbb{C}$.

Definition 8. We say that a sequence $(f_m)_m$ of functions $U \to \mathbb{C}$ is *locally bounded* when for all $z \in U$, the sequence $(|f_m(z)|)_m$ is bounded.

THEOREM 8. (Montel) *Let $(f_m)_m$ be a locally bounded sequence of holomorphic functions $U \to \mathbb{C}$. There exists a subsequence of $(f_m)_m$ which converges uniformly on any compact subset of U.*

LEMMA 12. *Let $(f_m)_m$ be a locally bounded sequence of continuous functions $U \to \mathbb{C}$ and $f : U \to \mathbb{C}$ such that any subsequence of $(f_m)_m$ which converges uniformly on any compact subset of U converges towards f. Then $(f_m)_m$ converges uniformly on any compact towards f.*

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Proof. Let us assume that $(f_m)_m$ does not converge towards f. Then there exists some $\epsilon > 0$, compact $K \subset U$ and a non-decreasing function $\nu : \mathbb{N} \to \mathbb{N}$ such that for all m,

$$\|(f_{\nu(m)} - f)_K\|_\infty \geq \epsilon.$$

From Montel theorem, one can extract a subsequence of $(f_{\nu(m)})_m$ which converges towards f uniformly on any compact of U, and in particular on the compact K. This is in contradiction to the above inequality, and we deduce that $(f_m)_m$ converges towards f. □

THEOREM 9. (Cauchy formula) *Let us assume that U is simply connected and let $f : U \to \mathbb{C}$ be a holomorphic function and γ a loop included in U that is homeomorphic to a circle positively oriented. Then for all z in the interior domain of the loop,*

$$f(z) = \frac{1}{2\pi i} \oint_\gamma \frac{f(s)}{s - z}\, ds.$$

Let us also recall a sufficient condition for a holomorphic function to be biholomorphic.

THEOREM 10. *Let U be an open and simply connected set and $f : U \to \mathbb{C}$ be a holomorphic function. Let $V \subset U$ be an open set and γ a loop included in U that is homeomorphic to a positively oriented circle, and such that V is included in the interior domain of γ. We assume that:*

(1) *for all $z \in V$ and $s \in \gamma$, $f(z) \neq f(s)$;*
(2) *and for all $z \in V$, $f'(z) \neq 0$.*

Then f is a biholomorphism from V onto its image, meaning that there exists some holomorphic function $g : f(V) \to U$ such that for all $z \in f(V)$, $f(g(z)) = z$ and for all $z \in U$, $g(f(z)) = z$. Moreover, for all $z \in f(V)$,

$$g(z) = \frac{1}{2\pi i} \int_\gamma s \frac{f'(s)}{f(s) - z} ds.$$

6.2.2. *The limits of subsequences of $(\xi_t^{(k)})_k$ satisfy a Fredholm integral equation.* In this section, we prove the following theorem.

THEOREM 11. *Let $\nu : \mathbb{N} \to \mathbb{N}$ be a non-decreasing function, and assume that $(\xi_t^{(\nu(m))})_m$ converges uniformly on any compact of $\mathcal{I}_{\tau_t}$ towards a function ξ_t. Then this function satisfies the following equation for all $\alpha \in \mathcal{I}_\tau$:*

$$\xi_t'(\alpha) = \frac{1}{2\pi}\kappa_t'(\alpha) + \int_{\mathbb{R}} \frac{\partial \theta_t}{\partial \alpha}(\alpha, \beta)\xi_t'(\beta) d\beta.$$

Moreover, $\xi_t(0) = d/2$.

Proof. Convergence of the derivative of the counting functions: Since any compact of $\mathcal{I}_{\tau_t}$ can be included in the interior domain of a rectangle loop, through derivation of the Cauchy formula, the derivative of $\xi_t^{(\nu(m))}$ converges also uniformly on any compact, towards ξ_t'. Since $|(\xi_t^{(m)})'|$ is bounded by a constant that does not depend on m, and that $s \mapsto |\theta_t(\alpha, s)|$ is integrable on $\mathbb{R}$ for all α, then $s \mapsto \theta_t(\alpha, s)\xi_t'(s)$ is integrable on $\mathbb{R}$.

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Some notation: Let us fix some $\epsilon > 0$, and $\alpha_0 \in \mathbb{R}$. In the following, we consider some *irrational* number (and as a consequence, not the image of a Bethe root) $M > 1$ such that:

(1) $M \in \xi_t(\overset{\circ}{\mathbb{R}})$;
(2) $|P_t^{(k)}(M)| \leq \epsilon/2(2\mu_t - \pi)$ for all k greater than some k_0 (this is possible to impose in virtue of Theorem 7);
(3) and $\alpha_0 \in \xi_t^{-1}([-M, M])$.

Since $\xi_t(\mathbb{R})$ is an interval (this function is increasing on $\mathbb{R}$), one can take M arbitrarily close to the supremum of this interval. When M tends towards this supremum, $\xi_t^{-1}(M)$ tends to $+\infty$: if it did not, then this would contradict the fact that this is the supremum (again by monotonicity). One can assume that M is such that

$$\frac{1}{2\pi}\left| \int_{(\xi_t^{-1}([-M,M]))^c} \theta_t(\alpha, \beta)\xi_t'(\beta)\, d\beta \right| \leq \frac{\epsilon}{4}. \tag{20}$$

Let us also set $J_t = \xi_t^{-1}([-M, M])$.

The derivative of ξ_t relative to the axis $i\mathbb{R}$ is non-zero when close enough to $\mathbb{R}$: Indeed, for all $\alpha, \lambda \in \mathbb{R}$,

$$\xi_t^{(k)}(\alpha + i\lambda) = \frac{1}{2\pi}\kappa_t(\alpha + i\lambda) + \frac{n_k + 1}{2N_k} + \frac{1}{2\pi N_k}\sum_j \theta_t(\alpha + i\lambda, \boldsymbol{\alpha}_j^{(k)}(t)).$$

As a direct consequence, the derivative of the function $\lambda \mapsto -i\xi_t^{(k)}(\alpha + i\lambda)$ in 0 is

$$\frac{1}{2\pi}\kappa_t'(\alpha) + \frac{1}{2\pi N_k}\sum_j \theta_t(\alpha, \boldsymbol{\alpha}_j^{(k)}(t)) = (\xi_t^{(k)})'(\alpha) \geq \frac{1}{2\pi}\kappa_t'(\alpha) > 0. \tag{21}$$

Thus for all α, the derivative of the function $\lambda \mapsto -i\xi_t(\alpha + i\lambda)$ in 0 is greater than

$$\frac{1}{2\pi}\kappa_t'(\alpha).$$

Moreover, since the second derivative of $\lambda \mapsto -i\xi_t^{(k)}(\alpha + i\lambda)$ is a bounded function of α, with a bound that is independent from k, through Taylor integral formula, there exists a constant $p_t > 0$ such that for all $\lambda \in \mathbb{R}$ and $\alpha \in \mathbb{R}$,

$$|\xi_t(\alpha + i\lambda) - i\xi_t'(\alpha).\lambda - \xi_t(\alpha)| \leq p_t\lambda^2,$$

which implies

$$|\mathrm{Im}(\xi_t(\alpha + i\lambda)) - \xi_t'(\alpha).\lambda| \leq p_t\lambda^2.$$

By virtue of equation (21),

$$\mathrm{Im}(\xi_t(\alpha + i\lambda)) \geq \xi_t'(\alpha)\lambda - p_t\lambda^2 \geq \frac{1}{2\pi}\kappa_t'(\alpha)\lambda - p_t\lambda^2.$$

The derivative of ξ_t relative to the axis $i\mathbb{R}$ in $\alpha \in \mathbb{R}$ is greater than $(1/2\pi)\kappa_t'(\alpha)$.

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

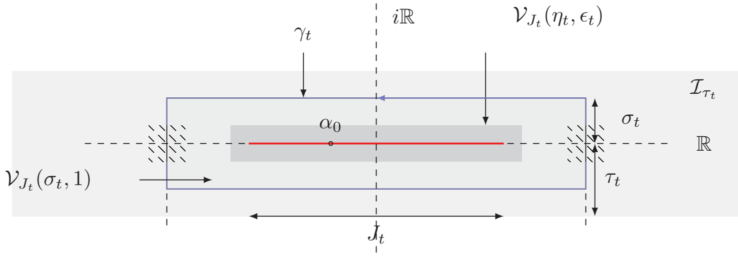


FIGURE 10. Illustration of the proof that ξ_t is a biholomorphism on a neighborhood of J_t.

The restriction of ξ_t on some $\mathcal{V}_{J_t}(\eta_t,\epsilon_t)$ is a biholomorphism onto its image: Since M is defined so that

$$M \in \xi_t(\overset{\circ}{\mathbb{R}}),$$

then $J_t = \xi_t^{-1}([-M, M])$ is compact. This means, as a consequence of the last point, that there exists some positive number $\sigma_t < \tau_t$ such that for all $z \in \mathcal{V}_{J_t}(\sigma_t, 1)\backslash\mathbb{R}$, then $\xi_t(z) \notin \mathbb{R}$.

Let us consider the loop γ_t given by $\partial\mathcal{V}_{J_t}(\sigma_t, 1)$ (see the illustration in Figure 10).

Let us prove that there exist some $\epsilon_t > 0$ and $\eta_t > 0$ such that the values taken by the function ξ_t on $\mathcal{V}_{J_t}(\eta_t,\epsilon_t)$ are distinct from any value taken by the same function on the loop γ_t. This is done in two steps, as follows.

(1) First, we consider open neighborhoods (illustrated by dashed squares in Figure 10) for the two points of $\gamma_t \cap \mathbb{R}$ such that the values taken by ξ_t on these sets are distant by more than a positive constant from the values taken on J_t. This is possible since ξ_t is strictly increasing on $\mathbb{R}$.

(2) On the part of γ_t that is not included in these two open sets, the function ξ_t takes non-real values, and the set of values taken is compact, by continuity. As a consequence, the set of values taken on the loop γ_t is included into a compact that does not intersect the set of values taken on J_t. Thus one can separate these two sets of values with open sets, meaning that there exist some $\epsilon_t > 0$ and $\eta_t > 0$ such that the set of values taken by ξ_t on $\mathcal{V}_{J_t}(\eta_t,\epsilon_t)$ does not intersect the set of values taken by this function on γ_t.

By virtue of Theorem 10, this means that ξ_t is a biholomorphism from $\mathcal{V}_{J_t}(\eta_t,\epsilon_t)$ onto its image on this set. As a consequence, it is also an open function, and its image on $\mathcal{V}_{J_t}(\eta_t,\epsilon_t)$ contains the image of J_t, where $[-M, M]$ by definition.

Asymptotic biholomorphism property for $\xi_t^{(\nu(k))}$: It can be derived from the last point that there exists some $k_1 \geq k_0$ such that for all $k \geq k_1$, the values of $\xi_t^{(\nu(k))}$ on γ_t are distinct from the values of $\xi_t^{(k)}$ on $\mathcal{V}_{J_t}(\eta_t,\epsilon_t)$, and as a consequence, for the same reason as the last point, $\xi_t^{(\nu(k))}$ is a biholomorphism from $\mathcal{V}_{J_t}(\eta_t,\epsilon_t)$ onto its image on this set. Moreover, since $\xi_t^{(\nu(k))}$ converges uniformly to ξ_t on $\overline{\mathcal{V}_{J_t}(\eta_t,\epsilon_t)}$, it converges also uniformly on $\mathcal{V}_{J_t}(\eta_t,\epsilon_t)$. Furthermore, $\xi_t(\mathcal{V}_{J_t}(\eta_t,\epsilon_t))$ contains $\mathcal{V}_{[-M,M]}(\eta_t',\epsilon_t')$, and thus there exists some $\eta_t',\epsilon_t' > 0$ and some $k_2 \geq k_1$ such that for all $k \geq k_2$, $\xi_t^{(\nu(k))}(\mathcal{V}_{J_t}(\eta_t,\epsilon_t))$ contains $\mathcal{V}_{[-M,M]}(\eta_t',\epsilon_t')$.

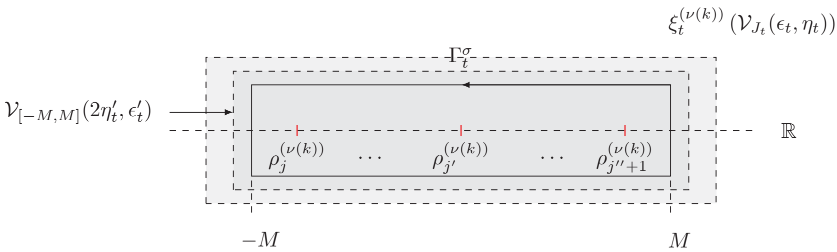


FIGURE 11. Illustration for the definition of the loop Γ_t^σ.

Loop integral expression of the counting functions and approximation of $\xi_t^{(\nu(k))}$: We deduce that for all $k \geq k_2$ and $\sigma < \eta_t'$ positive, the loop

$$\Gamma_t^\sigma \equiv (\{-M, M\} \times [\![-\sigma, \sigma]\!]) \bigcup ([-M, M] \times \{-\sigma, \sigma\})$$

is included into $\xi_t^{(\nu(k))}(\mathcal{V}_{J_t}(\eta_t, \epsilon_t))$. See Figure 11 for an illustration.

We then have, since $\alpha_0 \in J_t$, the following equation for all k, t, σ:

$$\frac{1}{2\pi N_{\nu(k)}} \sum_{j \in P_t^{(\nu(k))}(M)} \theta_t(\alpha_0, \boldsymbol{\alpha}_j^{(\nu(k))}(t)) \\ = \frac{1}{2\pi} \oint_{\Gamma_t^\sigma} \theta_t(\alpha_0, (\xi_t^{(\nu(k))})^{-1}(s)) \frac{e^{2i\pi N_{\nu(k)} s}}{(e^{2i\pi N_{\nu(k)} s} - 1)}\, ds.$$

Indeed, there are no poles for $\xi_t^{(\nu(k))}$ on Γ_t^σ since M is irrational. The poles of the function inside the domain delimited by Γ_t^σ are exactly the numbers $\rho_j^{(\nu(k))}$. By the residues theorem, and since for all j, $\xi_t^{(\nu(k))}(\boldsymbol{\alpha}_j(t)) = \rho_j^{(\nu(k))}$,

$$\oint_{\Gamma_t^\sigma} \theta_t(\alpha_0, (\xi_t^{(\nu(k))})^{-1}(s)) \frac{e^{2i\pi N_{\nu(k)} s}}{(e^{2i\pi N_{\nu(k)} s} - 1)}\, ds \\ = 2\pi i \sum_{j \in P_t^{(\nu(k))}(M)} \frac{1}{2i\pi N_{\nu(k)}} \theta_t(\alpha_0, \boldsymbol{\alpha}_j^{(\nu(k))}(t)). \tag{22}$$

Triangular inequality: We have the following triangular inequality:

$$\left| \xi_t^{(\nu(k))}(\alpha_0) - \frac{1}{2\pi}\kappa_t(\alpha_0) - \frac{d}{2} - \frac{1}{2\pi} \int_{-\infty}^{\infty} \theta_t(\alpha_0, \beta)\xi_t'(\beta)\, d\beta \right| \\ \leq \left| \frac{n_{\nu(k)} + 1}{2N_{\nu(k)}} - \frac{d}{2} \right| + \left| \xi_t^{(\nu(k))}(\alpha_0) - \frac{1}{2\pi}\kappa_t(\alpha_0) - \frac{n_{\nu(k)} + 1}{2N_{\nu(k)}} \right. \\ \left. - \frac{1}{2\pi} \oint_{\Gamma_t^\sigma} \theta_t(\alpha_0, (\xi_t^{(\nu(k))})^{-1}(s)) \frac{e^{2i\pi N_{\nu(k)} s}}{(e^{2i\pi N_{\nu(k)} s} - 1)}\, ds \right| \\ + \left| \int_{[-M,M]} \theta_t(\alpha_0, \xi_t^{-1}(\beta - i\sigma))\, d\beta - \int_{[-M,M]} \theta_t(\alpha_0, \xi_t^{-1}(\beta))\, d\beta \right|$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

$$+\left|\int_{[-M,M]} \theta_t(\alpha_0, \xi_t^{-1}(\beta - i\sigma))\, d\beta - \int_{\xi_t^{-1}([-M,M])} \theta_t(\alpha_0,\beta)\xi_t'(\beta)\, d\beta\right|$$
$$+\left|\frac{1}{2\pi}\int_{-\sigma}^{\sigma} \theta_t(\alpha_0, (\xi_t^{(\nu(k))})^{-1}(\pm M + i\lambda)))\frac{e^{2i\pi N_{\nu(k)}(\pm M+i\lambda)}}{(e^{2i\pi N_{\nu(k)}(\pm M+i\lambda)}-1)}\, d\lambda\right|$$
$$+\left|\frac{1}{2\pi}\int_{-M}^{M} \theta_t(\alpha_0, (\xi_t^{(\nu(k))})^{-1}(\beta + i\sigma_1))\frac{e^{2i\pi N_{\nu(k)}(\beta+i\sigma_1)}}{(e^{2i\pi N_{\nu(k)}(\pm(\beta+i\sigma_1)}-1)}\, d\beta\right|$$
$$+\left|\frac{1}{2\pi}\int_{-M}^{M} \theta_t(\alpha_0, (\xi_t^{(\nu(k))})^{-1}(\beta - i\sigma_1))\left(\frac{e^{2i\pi N_{\nu(k)}(\beta-i\sigma_1)}}{(e^{2i\pi N_{\nu(k)}(\pm(\beta-i\sigma_1)}-1)}-1\right) d\beta\right|$$
$$+\frac{1}{2\pi}\left|\int_{(\xi_t^{-1}(-[M,M]))^c} \theta_t(\alpha_0,\beta)\xi_t'(\beta)\, d\beta\right|. \tag{23}$$

We deduce from the last point (equation (22)) that for all $k \geq k_2$ and all $\sigma < \eta_t'$,

$$\left|\xi_t^{(\nu(k))}(\alpha_0) - \frac{1}{2\pi}\kappa_t(\alpha_0) - \frac{n_{\nu(k)}+1}{2N_{\nu(k)}}\right.$$
$$\left. - \frac{1}{2\pi}\oint_{\Gamma_t^\sigma} \theta_t(\alpha_0, (\xi_t^{(\nu(k))})^{-1}(s))\frac{e^{2i\pi N_{\nu(k)}s}}{(e^{2i\pi N_{\nu(k)}s}-1)}\, ds\right|$$
$$\leq \sum_{j\notin P_t^{(\nu(k))}(M)} |\theta_t(\alpha_0, \boldsymbol{\alpha}_j^{(\nu(k))}(t))|$$
$$\leq (2\mu_t - \pi)|\{j \in [\![1, n_{\nu(k)}]\!] : \boldsymbol{\alpha}_k^{(\nu(k))}(t) \notin [-M, M]\}| \leq \frac{\epsilon}{2}, \tag{24}$$

using the notation from the second point of this proof. Let us also note $k_3 \geq k_2$ some integer such that for all $k \geq k_3$,

$$\left|\frac{n_{\nu(k)}+1}{2N_{\nu(k)}} - \frac{d}{2}\right| \leq \frac{\epsilon}{8}. \tag{25}$$

We then evaluate convergence of the various other terms involved in the triangular inequality above.

(1) *Convergence of the bottom part of the loop integral to an integral on a real segment when* $\sigma \to 0$: By continuity of ξ_t^{-1}, there exists some $\sigma_0 > 0$ such that for all $k \geq k_3$, $\sigma \leq \sigma_0$,

$$\left|\int_{[-M,M]} \theta_t(\alpha_0, \xi_t^{-1}(\beta - i\sigma))\, d\beta - \int_{[-M,M]} \theta_t(\alpha_0, \xi_t^{-1}(\beta))d\beta\right| \leq \frac{\epsilon}{64}. \tag{26}$$

By change of variable in the second integral,

$$\left|\int_{[-M,M]} \theta_t(\alpha_0, \xi_t^{-1}(\beta - i\sigma))\, d\beta - \int_{\xi_t^{-1}([-M,M])} \theta_t(\alpha_0,\beta)\xi_t'(\beta)\, d\beta\right| \leq \frac{\epsilon}{64}. \tag{27}$$

(2) *Bounding the lateral parts of the loop integral for* $\sigma \to 0$: There exists some $\sigma_1 > 0$ such that $\sigma_1 \leq \sigma_0$ such that for all $\sigma \leq \sigma_1$, $k \geq k_3$,

$$\left|\frac{1}{2\pi}\int_{-\sigma}^{\sigma} \theta_t(\alpha_0, (\xi_t^{(\nu(k))})^{-1}(\pm M + i\lambda)))\frac{e^{2i\pi N_{\nu(k)}(\pm M+i\lambda)}}{(e^{2i\pi N_{\nu(k)}(\pm M+i\lambda)}-1)}\, d\lambda\right| \leq \frac{\epsilon}{64}. \tag{28}$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

(3) *Convergence of the top and bottom parts of the loop integral when* $k \to +\infty$: Then there exists some $k_4 \geq k_3$ such that for all $k \geq k_4$,

$$\left| \frac{1}{2\pi} \int_{-M}^{M} \theta_t(\alpha_0, (\xi_t^{(\nu(k))})^{-1}(\beta + i\sigma_1)) \frac{e^{2i\pi N_{\nu(k)}(\beta+i\sigma_1)}}{(e^{2i\pi N_{\nu(k)}(\pm(\beta+i\sigma_1)} - 1)} \, d\beta \right| \leq \frac{\epsilon}{64}. \quad (29)$$

$$\left| \frac{1}{2\pi} \int_{-M}^{M} \theta_t(\alpha_0, (\xi_t^{(\nu(k))})^{-1}(\beta - i\sigma_1)) \left(\frac{e^{2i\pi N_{\nu(k)}(\beta-i\sigma_1)}}{(e^{2i\pi N_{\nu(k)}(\pm(\beta-i\sigma_1)} - 1)} - 1 \right) d\beta \right| \leq \frac{\epsilon}{64}. \quad (30)$$

Using equations 23–30 together with equation (20), we have that for all $k \geq k_4$,

$$\left| \xi_t^{(\nu(k))}(\alpha_0) - \frac{1}{2\pi}\kappa_t(\alpha_0) - \frac{d}{2} - \frac{1}{2\pi}\int_{-\infty}^{\infty} \theta_t(\alpha_0, \beta)\xi_t'(\beta) \, d\beta \right| \leq \frac{\epsilon}{2} + \frac{\epsilon}{4} + \frac{\epsilon}{8} + 5\frac{\epsilon}{64} \leq \epsilon.$$

Integral equations: As a consequence, since this is true for all $\epsilon > 0$ and α_0, we have the following equality for all $\alpha \in \mathbb{R}$:

$$\xi_t(\alpha) = \frac{1}{2\pi}\kappa_t(\alpha) + \frac{d}{2} + \frac{1}{2\pi}\int_{-\infty}^{\infty} \theta_t(\alpha, \beta)\xi_t'(\beta) \, d\beta.$$

Moreover, this equality is verified for any α_0, and differentiating it relatively to α:

$$\xi_t'(\alpha) = \frac{1}{2\pi}\kappa_t'(\alpha) + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\partial\theta_t}{\partial\alpha}(\alpha, \beta)\xi_t'(\beta) \, d\beta.$$

Value of $\xi_t(0)$: Since $\xi_t^{(k)}$ is increasing for all k, we have directly

$$\frac{\lfloor n_k/2 \rfloor}{N_k} = \xi_t^{(k)}(\boldsymbol{\alpha}^{(k)}_{\lfloor n_k/2 \rfloor}(t)) \leq \xi_t^{(k)}(0) \leq \xi_t^{(k)}(\boldsymbol{\alpha}^{(k)}_{\lceil n_k/2 \rceil + 1}(t)) = \frac{\lceil n_k/2 \rceil + 2}{N_k}.$$

As a consequence, since we assumed at the very beginning of §6 that $n_k/N_k \to d$, $\xi_t(0) = d/2$. □

6.2.3. *Solution of the Fredholm equation.* In this section, we prove that the integral equation on ξ_t in the statement of Theorem 11 is unique and compute its solution.

PROPOSITION 9. *Let* $t \in (0, \sqrt{2})$ *and* ρ *a continuous function in* $L^1(\mathbb{R}, \mathbb{R})$ *such that for all* $\alpha \in \mathbb{R}$,

$$\rho(\alpha) = \frac{1}{2\pi}\kappa_t'(\alpha) + \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{\partial\theta_t}{\partial\alpha}(\alpha, \beta)\rho(\beta) \, d\beta.$$

Then for all α,

$$\rho(\alpha) = \frac{1}{4\mu_t \cosh(\pi\alpha/2\mu_t)}.$$

Proof. The proof consists essentially in the application of Fourier transform techniques. We will set, for convenience, for all α and μ,

$$\Xi_\mu(\alpha) = \frac{\sin(\mu)}{\cosh(\alpha) - \cos(\mu)}.$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Application of Fourier transform: Let us denote by $\hat{\rho}$ the Fourier transform of ρ: for all ω,

$$\hat{\rho}(\omega) = \int_{-\infty}^{+\infty} \rho(\alpha) e^{i\omega\alpha}\, d\alpha,$$

which exists since ρ is $L^1(\mathbb{R})$. Additionally, we denote by $\hat{\Xi}_\mu$ the Fourier transform of Ξ_μ. Thus, since

$$\int_{-\infty}^{+\infty} \frac{\partial\theta_t}{\partial\alpha}(\alpha,\beta)\rho(\beta)\, d\beta = -\int_{-\infty}^{+\infty} \Xi_\mu(\alpha-\beta)\rho(\beta)\, d\beta,$$

this defines a convolution product, which is transformed into a simple product through the Fourier transform, so that for all ω,

$$\hat{\rho}(\omega) = \frac{1}{2\pi}\hat{\Xi}_{\mu_t}(\omega) - \frac{1}{2\pi}\hat{\Xi}_{2\mu_t}(\omega)\hat{\rho}(\omega).$$

$$2\pi\hat{\rho}(\omega) = \frac{\hat{\Xi}_{\mu_t}(\omega)}{1+(1/2\pi)\hat{\Xi}_{2\mu_t}(\omega)}. \tag{31}$$

Computation of $\hat{\Xi}_\mu$:

- *Singularities of this function:* The singularities of the function Ξ_μ are exactly the numbers $i(\mu+2k\pi)$ for $k \geq 0$ and $i(-\mu+2k\pi)$ for $k \geq 1$, since for $\alpha \in \mathbb{C}$, $\cosh(\alpha) = \cos(\mu)$ if and only if

$$\cos(i\alpha) = \cos(\mu),$$

  and this implies that $\alpha = i(\pm\mu+2k\pi)$ for some k.
- *Computation of the residues:* For all k, the residue of Ξ_μ in $i(\mu+2k\pi)$ is

$$\text{Res}(\Xi_\mu, i\mu+2k\pi) = \frac{e^{i\gamma.i(\mu+2k\pi)}}{i} = \frac{1}{i}e^{-\gamma(\mu+2k\pi)}.$$

  As well,

$$\text{Res}(\Xi_\mu, -i\mu+2k\pi) = \frac{e^{i\gamma.i(-\mu+2k\pi)}}{i} = -\frac{1}{i}e^{-\gamma(-\mu+2k\pi)}.$$

  We have, for all γ,

$$\int_{-\infty}^{+\infty} \Xi_\mu(\alpha)e^{i\alpha\gamma}\, d\alpha = 2\pi\frac{\sinh[(\pi-\mu)\gamma]}{\sinh(\pi\gamma)}.$$

- *Residue theorem:* Let us set, for all integer n, the loop $\Gamma_n = [-n,n]+i[0,n]$. The residues Ξ_μ inside the domain delimited by this loop are the $i(\mu+2k\pi)$ with $k \geq 0$, and the $i(-\mu+2k\pi)$ with $k \geq 1$. For all n,

$$\int_{\Gamma_n} \Xi_\mu(\alpha)e^{i\alpha\gamma}\, d\alpha = \int_{\Gamma_n} \frac{\sinh(i\mu)}{i(\cosh(\alpha)-\cosh(i\mu))}e^{i\alpha\gamma}\, d\alpha.$$

  By the residue theorem,

$$\int_{\Gamma_N} \Xi_\mu(\alpha)e^{i\alpha\gamma}\, d\alpha = 2\pi i\Bigg(\sum_{k\geq 0}\text{Res}(\Xi_\mu, i(\mu+2k\pi)) - \sum_{k\geq 1}\text{Res}(\Xi_\mu, i(-\mu+2k\pi))\Bigg).$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

- *Asymptotic behavior:* Since only the contribution on $[-n, n]$ of the integral is non-zero asymptotically, and by convergence of the integral and the sums,

$$\int_{-\infty}^{+\infty} \Xi_\mu(\alpha) e^{i\alpha\gamma}\, d\alpha = 2\pi e^{-\gamma\mu} + 2\pi \sum_{k=1}^{+\infty} (-e^{\gamma\mu} + e^{-\gamma\mu}) e^{-2\gamma k\pi}$$

$$= 2\pi e^{-\gamma\mu} + 2\pi(-e^{\gamma\mu} + e^{-\gamma\mu})\left(\frac{1}{1 - e^{-2\gamma\pi}} - 1\right)$$

$$= 2\pi e^{-\gamma\mu} + 2\pi(-e^{\gamma\mu} + e^{-\gamma\mu}) \frac{e^{-\gamma\pi}}{e^{\gamma\pi} - e^{-\gamma\pi}}$$

$$= 2\pi \frac{e^{-\gamma(-\pi+\mu)} - e^{\gamma(-\pi-\mu)} - e^{\gamma(\mu-\pi)} + e^{\gamma(-\pi-\mu)}}{e^{\gamma\pi} - e^{-\gamma\pi}}$$

$$\hat{\Xi}_\mu(\gamma) = 2\pi \frac{\sinh(\gamma(\pi - \mu))}{\sinh(\gamma\pi)}. \tag{32}$$

Computation of $\hat{\rho}$: Using equations (31) and (32) with $\mu = \mu_t$, for all ω,

$$2\pi\hat{\rho}(\omega) = \frac{2\pi \sinh(\omega(\pi - \mu_t))}{\sinh(\pi\omega) + \sinh(\omega(\pi - 2\mu_t))}$$

$$= \frac{4\pi \sinh(\omega(\pi - \mu_t))}{e^{\omega\pi}.(1 + e^{-2\mu_t\omega}) - e^{-\omega\pi}.(1 + e^{2\mu_t\omega})}$$

$$= \frac{4\pi \sinh(\omega(\pi - \mu_t))}{e^{\omega(\pi-\mu_t)}.(e^{\mu_t\omega} + e^{-\mu_t\omega}) - e^{-\omega(\pi-\mu_t)}.(e^{-\mu_t\omega} + e^{\mu_t\omega})}$$

$$= \frac{\pi}{\cosh(\mu_t\omega)}.$$

Inverse transform: We thus have for all α,

$$2\pi\rho(\alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\pi}{\cosh(\mu_t\omega)} e^{-i\omega\alpha}\, d\omega = \frac{1}{\mu_t} \int_{-\infty}^{\infty} \frac{1}{2\cosh(u)} e^{-i\frac{u}{\mu_t}\alpha}\, du,$$

where we used the variable change $u = \mu_t\omega$. Using equation (32) for $\mu = \pi/2$,

$$\int_{-\infty}^{+\infty} \frac{1}{\cosh(\alpha)} e^{i\alpha\gamma}\, d\alpha = 2\pi \frac{\sinh(\pi\gamma/2)}{\sinh(\pi\gamma)} = \frac{\pi}{\cosh(\pi\gamma/2)}.$$

Thus we have

$$2\pi\rho(\alpha) = \frac{1}{2\mu_t} \frac{\pi}{\cosh(\pi\alpha/2\mu_t)} = \frac{\pi}{2\mu_t} \frac{1}{\cosh(\pi\alpha/2\mu_t)}$$

Finally,

$$\rho(\alpha) = \frac{1}{4\mu_t} \frac{1}{\cosh(\pi\alpha/2\mu_t)}.$$ □

6.2.4. *Convergence of* $\xi_t^{(k)}$

THEOREM 12. *There exists a function* $\boldsymbol{\xi}_{t,d} : \mathbb{R} \to \mathbb{R}$ *such that* $\xi_t^{(k)}$ *converges uniformly on any compact towards* $\boldsymbol{\xi}_{t,d}$. *Moreover, this function satisfies the following equation*

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

for all α*:*

$$\boldsymbol{\xi}'_{t,d}(\alpha) = \frac{1}{2\pi}\kappa'_t(\alpha) + \frac{1}{2\pi}\int_{\mathbb{R}} \frac{\partial\theta_t}{\partial\alpha}(\alpha,\beta)\boldsymbol{\xi}'_{t,d}(\beta)\, d\beta,$$

and $\boldsymbol{\xi}_{t,d}(0) = d/2$.

Proof. Consider any subsequence of $(\xi_t^{(k)})_k$ which converges uniformly on any compact of $\mathcal{I}_{\tau_t}$ to a function ξ_t. Using the Cauchy formula, the derivative of $(\xi_t^{(k)})$ converges uniformly on any compact to ξ'_t. Since the functions $\xi_t^{(k)}$ are uniformly bounded by a constant which is independent from k, and that for all k, $(\xi_t^{(k)})$, ξ'_t is positive and ξ_t is bounded, and thus ξ_t is in $L^1(\mathbb{R},\mathbb{R})$. From Theorem 11, we get that ξ'_t verifies a Fredholm equation, which has a unique solution in $L^1(\mathbb{R},\mathbb{R})$ [Proposition 9]. From Theorem 11, ξ_t, as a function on $\mathbb{R}$, is the unique primitive function of this one which has value $d/4$ on 0. Since this function is analytic, it determines its values on the whole stripe $\mathcal{I}_{\tau_t}$. By virtue of Lemma 12, $(\xi_t^{(k)})_k$ converge towards this function. □

PROPOSITION 10. *The limit of the function* $\boldsymbol{\xi}_{t,d}$ *in* $+\infty$ *is* $d/2 + \frac{1}{4}$*, and the limit in* $-\infty$ *is* $d/2 - 1/4$.

Proof. For all α,

$$\boldsymbol{\xi}_{t,d}(\alpha) = \frac{d}{2} + \frac{1}{4\mu_t}\int_0^{\alpha} \frac{1}{\cosh(\pi x/2\mu_t)}\, dx = \frac{d}{2} + \frac{1}{2\pi}\int_0^{2\mu_t\alpha/\pi} \frac{1}{\cosh(x)}\, dx.$$

This converges in $+\infty$ to

$$\frac{d}{2} + \frac{1}{\pi}\int_0^{+\infty} \frac{e^x}{e^{2x}+1}\, dx = \frac{d}{2} + \frac{1}{\pi}\int_0^{+\infty} (\arctan(\exp))'(x)\, dx = \frac{d}{2} + \frac{1}{2} - \frac{1}{\pi}\frac{\pi}{4} = \frac{d}{2} + \frac{1}{4}.$$

For the same reason, the limit in $-\infty$ is $d/2 - 1/4$. □

Remark 6. As a consequence, this limit is $> d$ when $d < 1/2$ and equal to d when $d = 1/2$.

6.3. *Condensation of Bethe roots relative to some functions.* In this section, we prove that if f is a continuous function $(0,+\infty) \to (0,+\infty)$, decreasing and integrable, then the scaled sum of the values of f on the Bethe roots converges to an integral involving f and $\boldsymbol{\xi}_{t,d}$ [Theorem 13]. Let us set, for all t, m and $M > 0$,

$$Q_t^{(k)}(M) := \left\{ j \in [\![1, n_k]\!] : \boldsymbol{\xi}_{t,d}^{-1}\left(\frac{j}{N_k}\right) \notin [-M, M] \right\},$$

and for two finite sets S, T, we set $S\Delta T = S\backslash T \cup T\backslash S$. For a compact set $K \subset \mathbb{R}$, we denote by $\delta(K) \equiv \max_{x,y\in K} |x-y|$ its diameter. For I a bounded interval of $\mathbb{R}$, we denote by $l(I)$ its length. When

$$J = \bigcup_j I_j$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

with (I_j) a sequence of bounded and disjoint intervals, the length of J is

$$l(J) = \sum_j l(I_j).$$

THEOREM 13. *Let $f : (0, +\infty) \to (0, +\infty)$ be a continuous, decreasing and integrable function. Then,*

$$\frac{1}{N_k} \sum_{j=\lceil n_k/2 \rceil+1}^{n_k} f(\boldsymbol{\alpha}_j^{(k)}(t)) \to \int_0^{\boldsymbol{\xi}_{t,d}^{-1}(d)} f(\alpha)\boldsymbol{\xi}'_{t,d}(\alpha)\, d\alpha,$$

where we set $\boldsymbol{\xi}_{t,1/2}^{-1}(1/2) = +\infty$.

Remark 7. This is another version of a statement proved in [**K18**] for bounded continuous and Lipschitz functions, which is not sufficient for the proof of Theorem 2.

Proof. In all the proof, the indexes j in the sums are in $[\![\lceil n_k/2 \rceil + 1, n_k]\!]$.

Setting: Let $\epsilon > 0$ and $t \in (0, \sqrt{2})$. Let us fix some M such that the following conditions are satisfied:

(1) for all k greater than some k_0,

$$\frac{1}{N_k} |P_t^{(k)}(M)| \leq \frac{\epsilon}{2\|f_{[M,+\infty)}\|_\infty + 1}; \tag{33}$$

(2) the following equation is satisfied:

$$\left| \int_{[M,+\infty)} f(\alpha)\boldsymbol{\xi}'_{t,d}(\alpha)\, d\alpha \right| \leq \frac{\epsilon}{2}; \tag{34}$$

(3) and if $d < 1/2$,

$$M > \boldsymbol{\xi}_{t,d}^{-1}(d), \tag{35}$$

which is possible to impose by virtue of Proposition 10.

Using the rarefaction of Bethe roots: We have the following, by definition:

$$\begin{aligned}\frac{1}{N_k} \sum_{j=\lceil n_k/2 \rceil+1}^{n_k} f(\boldsymbol{\alpha}_j^{(m)}(t)) &= \frac{1}{N_k} \sum_{j=\lceil n_k/2 \rceil+1}^{n_k} f\left((\xi_t^{(k)})^{-1} \left(\frac{j}{N_k} \right) \right) \\ &= \frac{1}{N_k} \sum_{j \notin P_t^{(k)}(M)} f\left((\xi_t^{(k)})^{-1} \left(\frac{j}{N_k} \right) \right) \\ &\quad + \frac{1}{N_k} \sum_{j \in P_t^{(k)}(M)} f\left((\xi_t^{(k)})^{-1} \left(\frac{j}{N_k} \right) \right).\end{aligned} \tag{36}$$

As a consequence of the inequality of equation (33) and then equation (34),

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

$$\begin{aligned}\left|\frac{1}{N_k}\sum_{j=\lceil n_k/2\rceil+1}^{n_k} f(\boldsymbol{\alpha}_j^{(k)}) - \frac{1}{N_k}\sum_{j\notin P_t^{(k)}(M)} f\left((\xi_t^{(k)})^{-1}\left(\frac{j}{N_k}\right)\right)\right| &\leq \frac{1}{N_k}|P_t^{(k)}(M)|.\|f_{[M,+\infty)}\|_\infty \\ &\leq \frac{\epsilon}{2\|f_{[M,+\infty)}\|_\infty+1}\|f_{[M,+\infty)}\|_\infty \\ &\leq \frac{\epsilon}{2}, \end{aligned} \tag{37}$$

since by definition, if $j \in P_t^{(k)}(M)$ and $j \geq \lceil n_k/2\rceil + 1$, then

$$(\xi_t^{(k)})^{-1}\left(\frac{j}{N_k}\right) \geq M.$$

On the asymptotic cardinality of $(P_t^{(k)}(M))^c\Delta(Q_t^{(k)}(M))^c$:

$$\frac{1}{N_k}|(P_t^{(k)}(M))^c\Delta(Q_t^{(k)}(M))^c| \underset{k\to+\infty}{\to} 0.$$

Indeed, $(P_t^{(k)}(M))^c\Delta(Q_t^{(k)}(M))^c$ is equal to the set

$$\left\{j \in [\![1, n_k]\!] : \frac{j}{N_k} \in (\xi_t{}^{(k)}([-M, M]))\Delta(\boldsymbol{\xi}_{t,d}([-M, M]))\right\},$$

thus its cardinality is smaller than

$$\delta(N_k((\xi_t{}^{(k)}([-M, M]))\Delta(\boldsymbol{\xi}_{t,d}([-M, M])))) + 1,$$

which is equal to

$$N_k\delta((\xi_t{}^{(k)}([-M, M]))\Delta(\boldsymbol{\xi}_{t,d}([-M, M]))) + 1.$$

As a consequence,

$$\frac{1}{N_k}|(P_t^{(k)}(M))^c\Delta(Q_t^{(k)}(M))^c| \leq \delta((\xi_t{}^{(k)}([-M, M]))\Delta(\boldsymbol{\xi}_{t,d}([-M, M]))) + \frac{1}{N_k}.$$

Since $\xi_t^{(k)}$ converges to $\boldsymbol{\xi}_{t,d}$ on any compact, and in particular $[-M, M]$, the diameter on the right of this inequality converges to 0 when k tends towards $+\infty$.

Replacing $P_t^{(k)}(M)$ *by* $Q_t^{(k)}(M)$ *in the sums:* Since f is decreasing and positive, for all $j \in [\![\lceil n_k/2\rceil + 1, n_k]\!]$,

$$\frac{1}{N_k}\left|f\left((\xi_t^{(k)})^{-1}\left(\frac{j}{N_k}\right)\right)\right| \leq \int_{[(j-1)/N_k, j/N_k]} f((\xi_t^{(k)})^{-1}(x))\,dx.$$

As a consequence, the difference

$$\frac{1}{N_k}\left|\sum_{j\notin P_t^{(k)}(M)} f\left((\xi_t^{(k)})^{-1}\left(\frac{j}{N_k}\right)\right) - \sum_{j\notin Q_t^{(k)}(M)} f\left((\xi_t^{(k)})^{-1}\left(\frac{j}{N_k}\right)\right)\right|$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

is smaller than

$$\int_{J_k} |f((\xi_t^{(k)})^{-1}(x))|\, dx = \int_{\xi_t^{(k)}(I_k)} f(x)(\xi_t^{(k)})'(x)\, dx,$$

where J_k is the union of the intervals

$$\left[\frac{j-1}{N_k}, \frac{j}{N_k}\right],$$

for $j \in (P_t^{(k)}(M))^c \Delta (Q_t^{(k)}(M))^c$. We also used the fact that f is a positive function. Since the functions $\xi_t^{(k)}$ are uniformly bounded independently of k, there exists a constant $C_t > 0$ such that for all k,

$$\int_{\xi_t^{(k)}(J_k)} f(x)(\xi_t^{(k)})'(x)\, dx \leq C_t \int_{\xi_t^{(k)}(J_k)} f(x)\, dx.$$

Since f is decreasing,

$$\int_{\xi_t^{(k)}(J_k)} f(x) \leq \int_{[0, l(\xi_t^{(k)}(J_k))]} f(x).$$

From the fact that $\xi_t^{(k)}$ is increasing,

$$l(\xi_t^{(k)}(J_k)) = \int_{J_k} (\xi_t^{(k)})'(\alpha)\, d\alpha.$$

Since the derivative of $\xi_t^{(k)}$ is bounded uniformly and independent of k, and that the length of J_k is smaller than $1/N_k|(P_t^{(k)}(M))^c \Delta (Q_t^{(k)}(M))^c|$,

$$l(\xi_t^{(k)}(J_k)) \to 0.$$

From the integrability of f on $(0, +\infty)$,

$$\int_{[0, l(\xi_t^{(k)}(J_k))]} f(x) \to 0.$$

As a consequence, there exists exists some $k_1 \geq k_0$ such that for all $k \geq k_1$,

$$\frac{1}{N_k}\left| \sum_{j \notin P_t^{(k)}(M)} f\left((\xi_t^{(k)})^{-1}\left(\frac{j}{N_k}\right)\right) - \sum_{j \notin Q_t^{(k)}(M)} f\left((\xi_t^{(k)})^{-1}\left(\frac{j}{N_k}\right)\right)\right| \leq \frac{\epsilon}{4}. \tag{38}$$

Approximating $\xi_t^{(k)}$ by $\xi_{t,d}$ *in the sum:*

(1) *Bounding the contribution in a neighborhood of* 0: With an argument similar to that used in the last point (bounding with integrals), there exists $\sigma > 0$ smaller than M such that for all k,

$$\frac{1}{N_k} \sum_{j \in (Q_t^{(k)}(\sigma))^c \cap (Q_t^{(k)}(M))^c} \left| f\left((\xi_t^{(k)})^{-1}\left(\frac{j}{N_k}\right)\right)\right| \leq \frac{\epsilon}{8}. \tag{39}$$

(2) *Using the convergence of* $\xi_t^{(k)}$ *on a compact away from* 0: There exists some $k_2 \geq k_1$ such that for all $k \geq k_2$,

$$\frac{1}{N_k} \sum_{j \in Q_t^{(k)}(\sigma) \cap (Q_t^{(k)}(M))^c} \left| f\left((\xi_t^{(k)})^{-1}\left(\frac{j}{N_k}\right)\right) - f\left(\boldsymbol{\xi}_{t,d}^{-1}\left(\frac{j}{N_k}\right)\right)\right| \leq \frac{\epsilon}{16}. \tag{40}$$

Indeed, for all the integers j in the sum, $\boldsymbol{\xi}_{t,d}^{-1}(j/N_k) \in [\sigma, M]$, and by uniform convergence of $(\xi_t^{(k)})^{-1}$ on the compact $\boldsymbol{\xi}_{t,d}([\sigma, M])$, for k great enough, the real numbers $(\xi_t^{(k)})^{-1}(j/N_k)$ and $\boldsymbol{\xi}_{t,d}^{-1}(j/N_k)$ for these integers k and these indexes j all lie in the same compact interval. Since f is continuous, there exists some $\eta > 0$ such that whenever x, y lie in this compact interval and $|x - y| \leq \eta$, then $|f(x) - f(y)| \leq \epsilon/8$. Since $(\xi_t^{(k)})^{-1}$ converges uniformly towards $\boldsymbol{\xi}_{t,d}^{-1}$ on the compact $\boldsymbol{\xi}_{t,d}([\sigma, M])$, there exists some $k_3 \geq k_2$ such that for all $k \geq k_3$, and for all j such that $j \in Q_t^{(k)}(\sigma\sigma)$ and $j \notin Q_t^{(k)}(M)$,

$$\left|(\xi_t^{(k)})^{-1}\left(\frac{j}{N_k}\right) - \boldsymbol{\xi}_{t,d}^{-1}\left(\frac{j}{N_k}\right)\right| \leq \eta.$$

As a consequence, we obtain the announced inequality.

Convergence of the remaining sum: The following sum is a Riemmann sum:

$$\frac{1}{N_k} \sum_{j \notin Q_t^{(k)}(M)} f\left(\boldsymbol{\xi}_{t,d}^{-1}\left(\frac{j}{N_k}\right)\right).$$

Indeed,

$$(Q_t^{(k)}(M))^c = \left\{ j \in [\![1, n_k]\!] : \boldsymbol{\xi}_{t,d}^{-1}\left(\frac{j}{N_k}\right) \in [-M, M]\right\}.$$

Let us also remember that we imposed also that the indexes in the sums of the proof are all in $[\![\lceil n_k/2\rceil + 1, n_k]\!]$. As a consequence, since $\boldsymbol{\xi}_{t,d}$ is increasing, the indexes are consecutive integers, from one that is at distance less than some constant from $N_k\boldsymbol{\xi}_{t,d}(0)$ and the last one is at distance less than this constant from $N_k\boldsymbol{\xi}_{t,d}(M)$.

As a consequence, if $d = 1/2$, this sum converges towards

$$\int_{\boldsymbol{\xi}_{t,d}(0)}^{\boldsymbol{\xi}_{t,d}(M)} f(\boldsymbol{\xi}_{t,d}^{-1}(\alpha))d\alpha = \int_0^M f(\alpha)\boldsymbol{\xi}'_{t,d}(\alpha)\, d\alpha$$

by a change of variable. If $d < 1/2$, it converges towards

$$\int_{\boldsymbol{\xi}_{t,d}(0)}^{d} f(\boldsymbol{\xi}_{t,d}^{-1}(\alpha))d\alpha = \int_0^d f(\alpha)\boldsymbol{\xi}'_{t,d}(\alpha)\, d\alpha,$$

since in this case, M was chosen to satisfy the inequality in equation (35).

As a consequence, there exists some $k_4 \geq k_3$ such that for all $k \geq k_4$, if $d = 1/2$,

$$\left| \sum_{j \notin Q_t^{(k)}(M)} f\left(\boldsymbol{\xi}_{t,d}^{-1}\left(\frac{j}{N_k}\right)\right) - \int_0^M f(\alpha)\boldsymbol{\xi}'_{t,d}(\alpha)\, d\alpha\right| \leq \frac{\epsilon}{16}. \tag{41}$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

If $d < 1/2$,

$$\left| \sum_{j \notin Q_t^{(k}(M)} f\left(\boldsymbol{\xi}_{t,d}^{-1}\left(\frac{j}{N_k}\right)\right) - \int_0^d f(\alpha)\boldsymbol{\xi}'_{t,d}(\alpha)\, d\alpha \right| \leq \frac{\epsilon}{16}. \tag{41)'$$

Assembling the inequalities: Putting together equations 36–41, if $d = 1/2$ (equation (41)′ if $d < 1/2$), we have for all $k \geq k_4$,

$$\left| \frac{1}{N_k} \sum_{j=\lceil n_k/2 \rceil + 1}^{n_k} f(\boldsymbol{\alpha}_j^{(k)}(t)) - \int_0^{\boldsymbol{\xi}_t^{-1}(d)} f(\alpha)\boldsymbol{\xi}'_{t,d}(\alpha)\, d\alpha \right| \leq \frac{\epsilon}{2} + \frac{\epsilon}{4} + \frac{\epsilon}{8} + \frac{\epsilon}{16} + \frac{\epsilon}{16} = \epsilon.$$

Since for all $\epsilon > 0$ there exists such an integer k_4, this proves the statement. □

7. *Computation of square ice entropy*

▷ Let us remember that the purpose of the paper is to compute the entropy of square ice, which is the entropy of the subshift X_s. In §5, we have expressed the entropy of the stripes subshifts $\overline{X}_N^s$ as a sum involving Bethe roots. To compute $h(X^s)$, we can use the following formula:

$$h(X^s) = \lim_N \frac{h(\overline{X}_N^s)}{N}.$$

In §6, we have dealt with the asymptotics of sums involving a positive decreasing integrable function and Bethe roots. We will thus combine the results of these two parts here.

Notation 17. For all $d \in [0, 1/2]$, we set

$$F(d) = -2 \int_0^{\boldsymbol{\xi}_{t,d}^{-1}(d)} \log_2(2|\sin(\kappa_t(\alpha)/2)|)\rho_t(\alpha)\, d\alpha.$$

LEMMA 13. *Let us consider (N_k) some sequence of integers, and (n_k) another sequence such that for all k, $2n_k + 1 \leq N_k/2$ and $(2n_k + 1)/N_k \to d \in [0, 1/2]$. Then*

$$\log_2(\lambda_{2n_k+1,N_k}(1)) \to F(d).$$

Proof. In this proof, for all k we set

$$(\mathbf{p}_j^{(k)})_j = (\kappa_t(\boldsymbol{\alpha}_j^{(k)}))_j,$$

the solution of the system of Bethe equations $(E_k)[1, 2n_k + 1, N_k]$. Then for all k (we use first Theorem 6 and then Theorem 3),

$$\lambda_{2n_k+1,N_k}(1) = \Lambda_{2n_k+1,N_k}[\mathbf{p}^{(k)}] = \left(2 + (N_k - 1) + \sum_{j \neq (n_k+1)} \frac{\partial \Theta_1}{\partial x}(0, \mathbf{p}_j^{(k)})\right) \prod_{j=1}^{n_k} M_1(e^{i\mathbf{p}_j^{(k)}}),$$

since by anti-symmetry of $\mathbf{p}^{(k)}$ (Theorem 4), and that the length $2n_k + 1$ of this tuple is odd, $\mathbf{p}_{n_k+1}^{(k)} = 0$ (the second case in Theorem 3 applies).

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

For all z such that $|z| = 1$,

$$M_1(z) = \frac{z}{z-1}.$$

By anti-symmetry of the sequences $\mathbf{p}^{(k)}$, for all k,

$$\prod_{j=1}^{n_k} e^{i\mathbf{p}_j^{(k)}} = \prod_{j=1}^{n_k} e^{i\mathbf{p}_j^{(k)}/2} = 1.$$

As a consequence,

$$\begin{aligned}\Lambda_{2n_k+1,N_k}[\mathbf{p}^{(k)}] &= \bigg(2 + (N_k - 1) + \sum_{j\neq(n_k+1)} \frac{\partial\Theta_1}{\partial x}(0, \mathbf{p}_j^{(k)})\bigg) \prod_{j=1}^{n_k} \frac{1}{e^{i\mathbf{p}_j^{(k)}} - 1} \\ &= \bigg(2 + (N_k - 1) + \sum_{j\neq(n_k+1)} \frac{\partial\Theta_1}{\partial x}(0, \mathbf{p}_j^{(k)})\bigg) \prod_{j=1}^{n_k} \frac{e^{-i\mathbf{p}_j^{(k)}/2}}{e^{i\mathbf{p}_j^{(k)}/2} - e^{-i\mathbf{p}_j^{(k)}/2}}.\end{aligned}$$

Since this number is positive (by Perron–Frobenius theorem),

$$\begin{aligned}\Lambda_{2n_k+1,N_k}[\mathbf{p}^{(k)}] &= |\Lambda_{2n_k+1,N_k}[\mathbf{p}^{(k)}]| \\ &= \bigg| 2 + (N_k - 1) + \sum_{j\neq(n_k+1)} \frac{\partial\Theta_1}{\partial x}(0, \mathbf{p}_j^{(k)})\bigg| \prod_{j=1}^{n_k} \frac{1}{2|\sin(\mathbf{p}_j^{(k)}/2)|}.\end{aligned}$$

As a consequence, since $\partial\Theta_1/\partial x$ is a bounded function,

$$\begin{aligned}\lim_k \log_2(\lambda_{2n_k+1,N_k}(1)) &= -\lim_k \bigg(\frac{1}{N_k}\sum_{j=1}^{n_k} \log_2(2|\sin(\mathbf{p}_j^{(k)}/2)|) + O\bigg(\frac{\log_2(N_k)}{N_k}\bigg)\bigg) \\ &= -2\lim_k \frac{1}{N_k} \sum_{j=\lceil n_k/2\rceil+1}^{n_k} \log_2(2|\sin(\kappa_t(\boldsymbol{\alpha}_k^{(k)})/2)|) \\ &= -2\int_0^{\boldsymbol{\xi}_{t,\mathbf{d}}^{-1}(\mathbf{d})} \log_2(2|\sin(\kappa_t(\alpha)/2)|)\rho_t(\alpha)\, d\alpha \\ &= F(\mathbf{d}),\end{aligned}$$

where $\rho_t = \boldsymbol{\xi}'_{t,d}$, and we used the anti-symmetry of the Bethe roots vectors in the second equality. For the other equalities, they are a consequence of Theorem 13, since the function defined as $\alpha \mapsto -\log_2(2|\sin(\kappa_t(\alpha)/2)|)$ on $(0, +\infty)$ is continuous, integrable, decreasing and positive:

(1) *Positive:* For all $\alpha > 0$, $\kappa_t(\alpha)$ is in

$$(0, \pi - \mu_t) = \bigg(0, \frac{\pi}{3}\bigg).$$

As a consequence, $2\sin(\kappa_t(\alpha)/2)$ is in $(0, 1)$, and this implies that for all $\alpha > 0$,

$$-\log_2(2|\sin(\kappa_t(\alpha)/2)|) > 0.$$

(2) *Decreasing:* This comes from the fact that $-\log_2$ is decreasing, and κ_t is increasing, and the sine is increasing on $(0, \pi/6)$.

(3) *Integrable:* Since $\kappa_t(0) = 0$ and $\kappa_t'(0) > 0$, for α positive sufficiently close to 0, $2\sin(\kappa_t(\alpha)/2) \leq 2\kappa_t'(0)\alpha$. As a consequence,

$$-\log_2(2|\sin(\kappa_t(\alpha)/2)|) \leq -\log_2(2\kappa_t'(0)\alpha).$$

Since the logarithm is integrable on any bounded neighborhood of 0, the function $\alpha \mapsto -\log_2(2|\sin(\kappa_t(\alpha)/2)|)$ is integrable.

The other limit is obtained by anti-symmetry of κ_t. □

THEOREM 1. *The entropy of square ice is*

$$h(X^s) = \frac{3}{2}\log_2\left(\frac{4}{3}\right).$$

Remark 8. This value corresponds to $\log_2(W)$ in [**L67**].

Proof. Here we fix $t = 1 \in (0, \sqrt{2})$. As a consequence, $\mu_t = 2\pi/3$.

Entropy of X^s and asymptotics of the maximal eigenvalue: Let us recall that the entropy of X^s is given by

$$h(X^s) = \lim_N \frac{1}{N} \max_{2n+1\leq N/2} \log_2(\lambda_{2n+1,N}(1)).$$

For all N, we denote by $\nu(N)$ the smallest j such that $2j + 1 \leq N/2$ and that for all n with $2n + 1 \leq N/2$,

$$\lambda_{2j+1,N}(1) \geq \lambda_{2n+1,N}(1).$$

By compactness, there exists an increasing sequence (N_k) such that $(2\nu(N_k) + 1)/N_k$ converges towards some non-negative real number $\mathbf{d}$. Since for all k, $2\nu(N_k) + 1 \leq N_k/2$, then $d \leq 1/2$. By virtue of Lemma 13, $h(X^s) = F(\mathbf{d})$.

Comparison with the asymptotics of other eigenvalues: Moreover, if d is another number in $[0, 1/2]$, there exists $\nu' : \mathbb{N} \to \mathbb{N}$ such that

$$(2\nu'(N) + 1)/N \to d.$$

For all k,

$$\lambda_{2\nu'(N_k)+1,N_k}(1) \leq \lambda_{2\nu(N_k)+1,N_k}(1).$$

Also by virtue of Lemma 13, $h(X^s) \geq F(d)$, and thus

$$F(\mathbf{d}) = \max_{d\in[0,1/2]} F(d).$$

This maximum is realized only for $d = 1/2$. As a consequence $d = 1/2$.

Rewritings: As a consequence,

$$h(X^s) = -2\int_0^{+\infty} \log_2(2|\sin(\kappa_t(\alpha)/2|)\rho_t(\alpha)\, d\alpha.$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

Let us rewrite this expression of $h(X^s)$ using

$$|\sin(x/2)| = \sqrt{\frac{1-\cos(x)}{2}}.$$

This leads to

$$h(X^s) = -\frac{\log_2(2)}{2}\int_{-\infty}^{+\infty} \rho_t(\alpha)\, d\alpha - \frac{1}{2}\int_{-\infty}^{+\infty} \log_2(1-\cos(\kappa_t(\alpha))).\rho_t(\alpha)\, d\alpha.$$

Thus,

$$h(X^s) = -\frac{1}{2}\int_{-\infty}^{+\infty} \log_2(2-2\cos(\kappa_t(\alpha))).\rho_t(\alpha)\, d\alpha.$$

Let us recall that for all α,

$$\rho(\alpha) = \frac{1}{4\mu_t \cosh(\pi\alpha/2\mu_t)} = \frac{3}{8\pi \cosh(3\alpha/4)}$$

$$\cos(\kappa_t(\alpha)) = \frac{\sin^2(\mu_t)}{\cosh(\alpha)-\cos(\mu_t)} - \cos(\mu_t) = \frac{3}{4(\cosh(\alpha)+1/2)} + \frac{1}{2}.$$

We thus have that

$$h(X^s) = -\frac{3}{16\pi}\int_{-\infty}^{+\infty} \log_2\left(1-\frac{3}{2\cosh(\alpha)+1}\right)\frac{1}{\cosh(3\alpha/4)}\, d\alpha.$$

Using the variable change $e^{\alpha} = x^4$, $d\alpha.x = 4\, dx$,

$$h(X^s) = -\frac{3}{16\pi}\int_{0}^{+\infty} \log_2\left(1-\frac{3}{x^4+1/x^4+1}\right)\frac{2}{(x^3+1/x^3)}\frac{4}{x}\, dx.$$

By symmetry of the integrand,

$$h(X^s) = -\frac{3}{4\pi}\int_{-\infty}^{+\infty} \frac{x^2\, dx}{x^6+1}\log_2\left(\frac{(2x^4-1-x^8)}{1+x^4+x^8}\right) dx$$

$$h(X^s) = -\frac{3}{4\pi}\int_{-\infty}^{+\infty} \frac{x^2\, dx}{x^6+1}\log_2\left(\frac{(x^2-1)^2(x^2+1)^2}{1+x^4+x^8}\right) dx.$$

Application of the residues theorem: In the following, we use the standard determination of the logarithm on $\mathbb{C}\backslash\mathbb{R}_-$.

We apply the residue theorem to obtain (the poles of the integrand are $e^{i\pi/6}$, $e^{i\pi/2}$, $e^{i5\pi/6}$)

$$\int_{-\infty}^{+\infty} \frac{x^2\log_2(x+i)}{x^6+1}\, dx = 2\pi i\left(\sum_{k=1,3,5} \frac{e^{ik\pi/3}\log_2(e^{ik\pi/6}+i)}{6e^{i5k\pi/6}}\right),$$

$$\int_{-\infty}^{+\infty} \frac{x^2\log_2(x-i)}{x^6+1}\, dx = -2\pi i\left(\sum_{k=7,9,11} \frac{e^{ik\pi/3}\log_2(e^{ik\pi/6}-i)}{6e^{i5k\pi/6}}\right).$$

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

By summing these two equations, we obtain that $\int_{-\infty}^{+\infty}(x^2\log_2(x^2+1)/(x^6+1))\,dx$ is equal to

$$\frac{\pi}{3}[\log_2(e^{i\pi/6}+i)-\log_2(e^{i\pi/2}+i)+\log_2(e^{i5\pi/6}+i)$$
$$+\log_2(e^{i7\pi/6}-i)-\log_2(e^{i9\pi/6}-i)+\log_2(e^{i11\pi/6}-i)].$$

This is equal to

$$\frac{\pi}{3}(\log_2(|e^{i\pi/6}+i|^2))-\log_2(|e^{i\pi/2}+i|)+\log_2(|e^{i5\pi/6}+i|^2)=\frac{2\pi}{3}\log_2\left(\frac{3}{2}\right).$$

Other computations: We do not include the following computation, since it is very similar to the previous one:

$$\int_{-\infty}^{+\infty}\frac{x^2}{x^6+1}\log_2(1+x^4x^8)\,dx=\frac{2\pi}{3}\log_2\left(\frac{8}{3}\right).$$

For the last integral, we write $\log_2((x^2-1)^2)=2\text{Re}(\log_2(x-1)+\log_2(x+1))$ and obtain

$$\int_{-\infty}^{+\infty}\frac{x^2}{x^6+1}\log_2((x^2-1)^2)=\text{Re}\bigg(\int_{-\infty}^{+\infty}\frac{x^2}{x^6+1}\log_2(x-1)$$
$$+\int_{-\infty}^{+\infty}\frac{x^2}{x^6+1}\log_2(x+1)\bigg)$$
$$=\frac{2\pi}{3}\log_2\left(\frac{1}{2}\right).$$

Summing these integrals: As a consequence,

$$h(X^s)=-\frac{3}{4\pi}\frac{2\pi}{3}\left(\log_2\left(\frac{1}{2}\right)+2\log_2\left(\frac{3}{2}\right)-\log_2\left(\frac{8}{3}\right)\right)$$
$$=\frac{1}{2}\log_2\left(\frac{4^3}{3^3}\right)=\frac{3}{2}\log_2\left(\frac{4}{3}\right).$$

□

8. *Comments*

This text is meant as a basis for further research that would aim at extending the computation method that we exposed to a broader set of multidimensional SFT, including, for instance, Kari–Culik tilings [**C96**], the monomer–dimer model [see, for instance, [**FP05**]], subshifts of square ice [**GS17**], the hard square shift [**P12**] or a three-dimensional version of the six-vertex model. Adaptations for these models may be possible, but would not be immediate at all. We explain here at which points the method has limitations, each of them coinciding with a specific property of square ice.

Let us recall that we called the Lieb path an analytic function of transfer matrices $t \mapsto V_N(t)$ such that for all t, $V_N(t)$ is an irreducible non-negative and symmetric matrix on Ω_N. Although the definition of transfer matrices admits straightforward generalization to multidimensional SFT and their non-negativity does not seem difficult to achieve, the

https://doi.org/10.1017/etds.2022.6 Published online by Cambridge University Press

property of *symmetry* of the matrices $V_N(t)$ relies on symmetries of the alphabet and local rules of the SFT. Friedland [**F97**] proved that under these symmetry constraints (which are verified, for instance, by the monomer–dimer and hard square models, but *a priori* not by Kari–Culik tilings), entropy is algorithmically computable, through a generalization of the gluing argument exposed in Lemma 1. Outside of the class of SFT defined by these symmetry restrictions, as far as we know, only strong mixing or measure theoretic conditions ensure algorithmic computability of entropy, leading, for instance, to relatively efficient algorithms approximating the hard square shift entropy [**P12**]. However, the *irreducibility* of the matrices $V_N(t)$ derives from the irreducibility property of the stripes subshifts X_N^s [Definition 2], that can be derived from the *linear block gluing property* of X^s [**GS17**]. This property consists in the possibility for any pair of patterns on $\mathbb{U}_N^{(2)}$ to be glued in any relative positions, provided that the distance between the two patterns is greater than a minimal distance, which is $O(N)$.

Furthermore, Lemma 1, which relies on a horizontal symmetry of the model, is a simplification in the proof of Theorem 2, whose implication is that the entropy of X^s can be computed through entropies of subshifts $\overline{X}_N^s$, and thus simplifies the algebraic Bethe ansatz, that we will expose in another text. One can see in [**VL19**] that it is possible to use the ansatz without Lemma 1. However, this application of the ansatz would lead to different Bethe equations, and it is not clear if these equations admit solutions, and if we can evaluate their asymptotic behavior. The symmetry is also involved in the equality of the entropy of $\overline{X}_{n,N}^s$ and the entropy of $\overline{X}_{N-n,N}^s$. Without this equality, we do not know how to identify the greatest eigenvalue of $V_N(t)$ with the candidate eigenvalue obtained via the ansatz.

Acknowledgements. The author was funded by the ANR project CoCoGro (ANR-16-CE40-0005) and is grateful to K.K. Kozlowski for helpful discussions.

REFERENCES

[B82] R. J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Academic Press, Cambridge, MA, 1982.

[C96] K. Culik. An aperiodic set of 13 Wang tiles. *Discrete Math.* **160** (1996), 245–251.

[DGHMT16] H. Duminil-Copin, M. Gagnebin, M. Harel, I. Manolescu and V. Tassion. Discontinuity of the phase transition for the planar random-cluster and Potts models with q > 4. *Preprint*, 2017, arXiv:1611.09877.

[DGHMT18] H. Duminil-Copin, M. Gagnebin, M. Harel, I. Manolescu and V. Tassion. The Bethe ansatz for the six-vertex and XXZ models: an exposition. *Probab. Surv.* **15** (2018), 102–130.

[F97] S. Friedland. On the entropy of $\mathbb{Z}^d$ subshifts of finite type. *Linear Algebra Appl.* **252** (1997), 199–220.

[FP05] S. Friedland and U. N. Peled. Theory of computation of multidimensional entropy with an application to the monomer-dimer problem. *Adv. Appl. Math.* **34** (2005), 486–522.

[GS17] S. Gangloff and M. Sablik. Quantified block gluing, aperiodicity and entropy of multidimensional SFT. *J. Anal. Math.* **144** (2021), 21–118.

[HKC92] L. P. Hurd, J. Kari and K. Culik. The topological entropy of cellular automata is uncomputable. *Ergod. Th. & Dynam. Sys.* **12** (1992), 2551–2065.

[HM10] M. Hochman and T. Meyerovitch. A characterization of the entropies of multidimensional shifts of finite type. *Ann. of Math. (2)* **171** (2010), 2011–2038.

[K18] K. Kozlowski. On condensation properties of Bethe roots associated with the XXZ spin chain. *Comm. Math. Phys.* **357** (2018), 1009–1069.

[K61] P. W. Kasteleyn. The statistics of dimers on a lattice: I. The number of dimer arrangements on a quadratic lattice. *Physica* **17** (1961), 1209–1225.

[K96] G. Kuperberg. Another proof of the alternating sign matrix conjecture. *Int. Math. Res. Not. IMRN* **1996** (1996), 139–150.

[L67] E. H. Lieb. Residual entropy of square ice. *Phys. Rev.* **162** (1967), 162.

[LM95] D. A. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, Cambridge, 1995.

[LSM61] E. H. Lieb, T. Shultz and D. Mattis. Two soluble models of an antiferromagnetic chain. *Ann. Physics* **16** (1961), 407–466.

[P12] R. Pavlov. Approximating the hard square entropy constant with probabilistic methods. *Ann. Probab.* **40** (2012), 2362–2399.

[PS15] R. Pavlov and M. Schraudner. Entropies realizable by block gluing shifts of finite type. *J. Anal. Math.* **126** (2015), 113–174.

[PV17] P. A. Pierce and A. Vittori-Orgeas. Yang–Baxter solution of Dimers as a free-fermion six-vertex model. *J. Phys. A* **50** (2017), 434001.

[VL19] R. S. Vieira and A. Lima-Santos. The algebraic Bethe ansatz and combinatorial trees. *J. Integrable Syst.* **4** (2019), xyz002.

[YY66a] C. N. Yang and C. P. Yang. One-dimensional chain of anisotropic spin-spin interactions. I. Proof of Bethe's hypothesis for ground state in a finite system. *Phys. Rev.* **150** (1966), 321.

[YY66b] C. N. Yang and C. P. Yang. One-dimensional chain of anisotropic spin-spin interactions. II. Properties of the ground-state energy per lattice site for an infinite system. *Phys. Rev.* **150** (1966), 327.