ON THE SPECIFICATION OF PROBABILITIES BY REGULAR g-FUNCTIONS

S. FRIEDLI

Abstract. We consider measures on $\{\pm 1\}^{\mathbb{Z}}$ for which the conditional probability that the spin at $n+1$ takes the value $+1$, given the values of all spins at sites $n, n-1, \dots$, is specified by some a priori given function g. We first show how the regularity of g (continuity and uniform non-nullness) allows to construct explicitely measures with such dependencies. We then consider the problem of extreme decomposition for the set of measures specified by g. Finally, we expose the uniqueness result of Johansson and Öberg [15]: when the variation of g is ℓ^2-summable, then there exists a unique invariant measure. These notes are in progress.

Contents

1. Introduction

Consider the product space associated to the finite alphabet $\{\pm 1\}$, $\Omega = \{\pm 1\}^{\mathbb{Z}}$, whose elements are the sequences $\omega = (\omega_n)_{n\in\mathbb{Z}}$. For each $\Lambda \subset \mathbb{Z}$, consider the set $\Omega_\Lambda = \{\pm 1\}^\Lambda$, together with the canonical projection $\pi_\Lambda : \Omega \to \Omega_\Lambda$ defined by $\pi_\Lambda(\omega)_k := \omega_k\ \forall k \in \Lambda$. Then define $\mathcal{C}(\Lambda) := \{\pi_\Lambda^{-1}(A) : A \subset \Omega_\Lambda\}$, called the algebra of cylinders with base Λ. For $S \subset \mathbb{Z}$ (possibly infinite), we consider the algebra of cylinders (with base in S), defined by $\mathcal{C}_S := \bigcup\{\mathcal{C}(\Lambda) : \Lambda \subset S, \text{finite}\}$, and the σ-algebra generated by the cylinders (with base in S): $\mathcal{F}_S := \sigma(\mathcal{C}_S)$. When $S = \mathbb{Z}$ we simply write $\mathcal{C} \equiv \mathcal{C}_{\mathbb{Z}}$ and $\mathcal{F} \equiv \mathcal{F}_{\mathbb{Z}}$. We denote by $\mathcal{M}$ the set of probability measures on $(\Omega, \mathcal{F})$.

The shift is the invertible map $T : \Omega \to \Omega$ defined by

$$(T\omega)_k := \omega_{k+1} \quad \forall k \in \mathbb{Z}\,.$$

A probability measure $\mu \in \mathcal{M}$ is invariant if $\mu \circ T^{-1} = \mu$, i.e. if $\mu(T^{-1}A) = \mu(A)$ for all $A \in \mathcal{F}$. The set of invariant probability measures is denoted $\mathcal{M}_T$.

The object of these notes is the study of measures μ for which the state of the present, when conditionned on the past, is specified by an a priori given function. Consider, for all $k \in \mathbb{Z}$, the σ-algebra $\mathcal{F}_{(-\infty,k]}$ (the past of $k+1$), and define the cylinder $[+]_k := \{\omega : \omega_k = +1\}$. Consider the random variable

$$g_k(\omega) := \mu([+]_{k+1}|\mathcal{F}_{(-\infty,k]})(\omega)\,.$$

Since g_k is $\mathcal{F}_{(-\infty,k]}$-measurable, it does not depend on the value of $\omega_{k'}$ when $k' > k$. Our purpose here is, like in the Theory of Gibbs States in Statistical Mechanics [6], to study measures for which the functions g_k are given a priori.

Definition 1.1. *Let, $g = (g_k)_{k\in\mathbb{Z}}$ where for each k, $g_k : \Omega \to [0,1]$ is $\mathcal{F}_{(-\infty,k]}$-measurable. A probability measure $\mu \in \mathcal{M}$ is specified by g if, for all $k \in \mathbb{Z}$,*

$$\mu([+]_{k+1}|\mathcal{F}_{(-\infty,k]})(\omega) = g_k(\omega) \tag{1.1}$$

for μ-almost all ω. The set of probability measures specified by g is denoted $\mathcal{M}(g)$, and $\mathcal{M}_T(g) := \mathcal{M}(g) \cap \mathcal{M}_T$.

We will only be interested in the case where the transition probabilities don't depend on k: we assume from now on that we are given a single $\mathcal{F}_{(-\infty,0]}$-measurable function g, such that $g_k = g \circ T^k$ for all k. We will use the following abuse of notation: $g_k(\omega) = g(T^k\omega) = g(\omega_k, \omega_{k-1}, \dots)$.

A function g playing the role defined above is called a g-function, and processes with dependencies as in (1.1), i.e. the elements of $\mathcal{M}(g)$ (if any), were originally called chains with complete connections [7] or g-measures [16]. Observe that if g depends, say, only on its first coordinate, i.e. $g = g(x_1)$, then the process is a simple Markov Chain. We will be concerned with the more interesting case where the state of the process at a given time depends on the whole past.

Let us give examples of g-functions.

Example 1.1. A natural way of defining $\mu([+]_{k+1}|\mathcal{F}_{(-\infty,k]})(\omega)$, i.e. determining the state of ω_{k+1} when conditionned on the past, is to look at a random distance $R \in \mathbb{N}$ back in the past and, according to a certain rule depending only on $\omega_k, \omega_{k-1}, \dots, \omega_{k-R+1}$, determine the probability that $\omega_{k+1} = +1$. This is done by defining a probability distribution for the random distance R, $Q(R = n) := p_n$, with $\sum_{n\geq 1} p_n = 1$, and, for each n, defining some function $P_n(\omega_k, \dots, \omega_{k-n+1}) \in [0,1]$ which gives the probability that $\omega_{k+1} = +1$ once $R = n$. By summing over all possibilities for the range R, we get:

$$g(\omega_k, \omega_{k-1}, \dots) := \sum_{n\geq 1} p_n P_n(\omega_k, \dots, \omega_{k-n+1})\,.$$

A natural choice for P_n is to follow the majority, in which case $\{\omega_{k+1} = +1\}$ is favorized if most of the spins $\omega_k, \dots, \omega_{k-n+1}$ are $+1$. This is naturally obtained by requiring that P_n is a function of the mean $\frac{1}{n}\sum_{k=0}^{n-1}\omega_{k-j}$:

$$P_n(\omega_k, \dots, \omega_{k-n}) := \varphi\Big(\frac{1}{n}\sum_{j=0}^{n-1}\omega_{k-j}\Big)\,,$$

for some non-decreasing measurable $\varphi : [-1,+1] \to [0,1]$, with, say, $\varphi(x) > \frac{1}{2}$ for $x > 0$ and $\varphi(x) < \frac{1}{2}$ for $x < 0$. For example, Bramson and Kalikow [5] considered the function

$$\varphi(x) = \begin{cases} 1-\epsilon & \text{if } x \geq 0\,, \\ \epsilon & \text{if } x < 0\,. \end{cases} \tag{1.2}$$

A natural class of sequences $(p_n)_{n\geq 1}$ is, for example, of those in which p_n behaves, for large n, as

$$p_n \sim \frac{1}{n^{1+\alpha}}\,, \tag{1.3}$$

where $\alpha > 0$. In this case we have

$$\mathsf{E}_Q(R) \begin{cases} < \infty & \text{if } 1 < \alpha < \infty\,, \\ = \infty & \text{if } 0 < \alpha \leq 1\,. \end{cases}$$

As will be seen in Section 4, $\mathsf{E}_Q(R) < \infty$ (with (1.2)) implies uniqueness of the invariant measure specified by g.

Example 1.2. Our second example is inspired by lattice systems of statistical mechanics. Let $(J_n)_{n\geq 1}$ be a summable sequence of non-negative real numbers. Then, define $g = g(\omega_k, \omega_{k-1}, \dots)$ by

$$g := \frac{e^{-\beta H}}{e^{+\beta H} + e^{-\beta H}}\,, \tag{1.4}$$

where $H = H(\omega_k, \omega_{k-1}, \dots) := -\sum_{n\geq 1} J_n \omega_{k-n+1}$, and $\beta > 0$ is the inverse temperature. Again, a natural type of sequence $(J_n)_{n\geq 1}$ is the one for which

$$J_n \sim \frac{1}{n^{1+\gamma}}\,, \tag{1.5}$$

where $\gamma > 0$. We will see in Section 4 that $\gamma > \frac{1}{2}$ implies uniqueness of the invariant measure.

The purpose of these notes is to expose some general "well-known" facts about processes described by g-functions of the type given above. In Section 2 we start by giving a standard existence result, based on a compactness argument. Rather than just giving existence of measures specified by g-functions, it also allows to *prepare* measures with prescribed boundary conditions, as in Statistical Mechanics. Then, in Section 3, we take a closer look at the convex structure of the sets $\mathcal{M}(g)$ and $\mathcal{M}_T(g)$, by showing that they are completely determined by their extreme elements. In Section 4 we expose a robust uniqueness criterium due to Johansson and Öberg. Section 5, still under construction, will be devoted to the problem of non-uniqueness.

Besides two results in Section 3.6 which we take from [10], the text is self contained. The proofs of some easy lemmas have been defered to the end of the text, in Appendix A.

2. Existence

A first problem is to find conditions on g which ensure that $\mathcal{M}(g) \neq \emptyset$. As in the Theory of Gibbs States, existence of specified probability measures is standard, and guaranteed under natural assumptions, namely continuity and non-nullness. We pursue this in the present section, which I originally wrote in order to understand the second paragraph on p. 156 of [5].

Definition 2.1. *A g-function g is* **continuous** *if $\mathrm{var}_k(g) \to 0$ when $k \to \infty$, where $\mathrm{var}_k(g)$ is the* **variation of g of order k**, *defined by*

$$\mathrm{var}_k(g) := \sup\left\{|g(\sigma) - g(\sigma')| : \sigma_l = \sigma_l\,,\ 1 \leq l \leq k\right\}.$$

Therefore, if a probability measure is specified by a continuous g-function, then the dependence of the present on the remote past is weak. We verify that the g-functions defined in the two examples of the previous section are continuous. In Example 1.1, we have:

$$\mathrm{var}_k(g) \leq \sum_{n>k} p_n \,, \tag{2.1}$$

which goes to zero when $k \to \infty$. In Example 1.2, a simple computation leads to:

$$\mathrm{var}_k(g) \leq \big(2\beta \sum_{n>k} J_n\big) \exp\big(2\beta \sum_{n>k} J_n\big) \,. \tag{2.2}$$

Since $(J_n)_{n\geq 1}$ is summable, this upper bound goes to zero when $k \to \infty$.

Definition 2.2. *A g-function g is* uniformly non-null *if it is uniformly bounded away from zero and one, i.e. if there exists $\epsilon > 0$ such that*

$$\epsilon \leq \inf_\sigma g(\sigma) \leq \sup_\sigma g(\sigma) \leq 1 - \epsilon \,. \tag{2.3}$$

In other words, non-nullness means that uniformly in its past, a spin always has a positive probability of changing sign. The g-functions of Examples 1.1 and 1.2 are uniformly non-null. From now on, we shall only consider g-functions which satisfy simultaneously the two properties defined above:

Definition 2.3. *A g-function g is* regular [5] *if it is both continuous and uniformly non-null.*

We start by showing that there always exists at least one probability measure speficied by a regular g-function. The argument is standard and follows from the compactness of $\mathcal{M}$ in the weak topology. The latter is defined as follows: a sequence $(\mu_n)_{n\geq 1}$, $\mu_n \in \mathcal{M}$ converges weakly to $\mu \in \mathcal{M}$ (denoted $\mu_n \Rightarrow \mu$) if $\mu_n(A) \to \mu(A)$ for all cylinder $A \in \mathcal{C}$. As well known, this convergence turns $\mathcal{M}$ into a sequentially compact topological space. In the sequel, all topological considerations about $\mathcal{M}$ will be with respect to this convergence. Observe that $\mathcal{M}_T \subset \mathcal{M}$ is closed, and hence compact.

Theorem 2.1. *Assume g is regular. Then*

(1) *$\mathcal{M}(g)$ and $\mathcal{M}_T(g)$ are closed, and hence compact.*
(2) *$\mathcal{M}(g) \neq \emptyset$, $\mathcal{M}_T(g) \neq \emptyset$.*

Before starting the proof, let us introduce a few notations and conventions. If $\omega \in \Omega$, $-\infty < a < b < +\infty$, then $\omega_a^b := (\omega_b, \omega_{b-1}, \dots, \omega_a)$. Similarly, $\omega_{-\infty}^b := (\omega_b, \omega_{b-1}, \dots)$ We will use and concatenate such words in many ways. For example, $\omega_a^b \sigma_{-\infty}^{a-1} := (\omega_b, \dots, \omega_a, \sigma_{a-1}, \dots)$. Also, if $\sigma \in \{\pm 1\}^{\mathbb{N}}$, then $\omega_a^b \sigma := (\omega_b, \dots, \omega_a, \sigma_1, \sigma_2, \dots)$.

Analogously, $[\omega]_a^b := \{\omega' : \omega'_k = \omega_k,\, a \leq k \leq b\}$ is called a thin cylinder. In the degenerate case $a = b$, we simply write $[\omega]_a = \{\omega' : \omega'_a = \omega_a\}$. We shall use the notations $[\pm]_a^b$ when ω is the constant configuration $\omega_k = \pm 1$

$\forall k$. Similarly, on defines $[\omega]_{-\infty}^{b} := \{\omega' : \omega'_k = \omega_k \,\forall k \leq b\}$. To avoid heavy notations, we will sometimes write $[\omega]_a^b[\sigma]_c^{a-1}$ instead of $[\omega]_a^b \cap [\sigma]_c^{a-1}$.

The study of $\mathcal{M}(g)$ will use repeatedly the following Martingale Convergence Theorem: since $\mathcal{F}_{(-\infty,k]}$ is the smallest σ-algebra containing $\bigcup_{l\geq 1}\mathcal{F}_{[-l,k]}$, then for any $A \in \mathcal{F}$,

$$\mu(A|\mathcal{F}_{(-\infty,k]}) = \lim_{l\to\infty}\mu(A|\mathcal{F}_{[-l,k]}) \quad \mu\text{-a.s.} \tag{2.4}$$

We will also use the definition of conditional probability: if $\mu \in \mathcal{M}(g)$, then for all $A \in \mathcal{F}_{(-\infty,k]}$,

$$\begin{aligned}\mu([+]_{k+1}\cap A) &= \int_A \mu([+]_{k+1}|\mathcal{F}_{(-\infty,k]})d\mu \\ &= \int_A g(\omega_{-\infty}^k)\mu(d\omega)\,. \end{aligned}\tag{2.5}$$

One can show, for example,

Lemma 2.1. *Let $\mu \in \mathcal{M}(g)$, where g is uniformly non-null with constant ϵ. Then $\mu([\sigma]_a^b) \geq \epsilon^{b-a+1} > 0$ for all thin cylinder $[\sigma]_a^b$.*

Proof of Theorem 2.1. The following is largely inspired by the proof of Dobrushin [6] for random fields on $\mathbb{Z}^d$. To show that $\mathcal{M}(g)$ is closed, consider a sequence $(\mu_n)_{n\geq 1}$, $\mu_n \in \mathcal{M}(g)$, and assume $\mu_n \Rightarrow \mu$ for some $\mu \in \mathcal{M}$. We show that $\mu \in \mathcal{M}(g)$. By (2.4), $\mu([+]_{k+1}|\mathcal{F}_{(-\infty,k]})$ equals, for μ-almost all ω,

$$\lim_{l\to\infty}\mu([+]_{k+1}|\mathcal{F}_{[-l,k]})(\omega) = \lim_{l\to\infty}\frac{\mu([+]_{k+1}[\omega]_{-l}^k)}{\mu([\omega]_{-l}^k)} = \lim_{l\to\infty}\lim_{n\to\infty}\frac{\mu_n([+]_{k+1}[\omega]_{-l}^k)}{\mu_n([\omega]_{-l}^k)}$$

We used Lemma 2.1. Using (2.5) for each μ_n,

$$\mu_n([+]_{k+1}[\omega]_{-l}^k) = \int_{[\omega]_{-l}^k} g(\omega_{-l}^k\sigma_{-\infty}^{-l-1})\mu_n(d\sigma)$$

We rewrite this last integral as

$$g(\omega_{-\infty}^k)\mu_n(\omega_{-l}^k) + \int_{\omega_{-l}^k}\big[g(\omega_{-l}^k\sigma_{-\infty}^{-l-1}) - g(\omega_{-l}^k\omega_{-\infty}^{-l-1})\big]\mu_n(d\sigma)$$

The difference in the integral can be bounded using the variation of g, which leads to

$$\Big|\frac{\mu_n([+]_{k+1}[\omega]_{-l}^k)}{\mu_n([\omega]_{-l}^k)} - g(\omega_{-\infty}^k)\Big| \leq \mathrm{var}_{k+l}(g)\,,$$

uniformly in n. Therefore we get

$$\mu([+]_{k+1}|\mathcal{F}_{(-\infty,k]})(\omega) = g(\omega_{-\infty}^k)\,,$$

which proves the first statement. To show that $\mathcal{M}_T(g)$ is closed, let $\mu_n \in \mathcal{M}_T(g)$, $\mu_n \Rightarrow \mu$. We know that $\mu \in \mathcal{M}(g)$. Moreover, for any cylinder $B \in \mathcal{C}$,

$$\mu(T^{-1}B) = \lim_{n\to\infty} \mu_n(T^{-1}B) = \lim_{n\to\infty} \mu_n(B) = \mu(B)\,.$$

The following lemma allows to conclude that $\mu \in \mathcal{M}_T$.

Lemma 2.2. *Let $\nu \in \mathcal{M}$ be such that $\nu(T^{-1}B) = \nu(B)$ for all cylinder $B \in \mathcal{C}$. Then $\nu \in \mathcal{M}_T$.*

For the second statement of the theorem, we start by constructing a measure ν^σ on $(\{\pm 1\}^{\mathbb{Z}_+}, \mathcal{F}_{\mathbb{Z}_+})$ ($\mathbb{Z}_+ := \mathbb{Z} \cap [1, +\infty)$), associated to a boundary condition $\sigma \in \{\pm 1\}^{\mathbb{Z}_-}$ ($\mathbb{Z}_- := \mathbb{Z} \cap (-\infty, 0]$). First, define ν^σ on thin cylinders:

$$\nu^\sigma([\omega]_1^n) := \prod_{k=1}^n \widehat{g}(\omega_1^k \sigma)\,, \tag{2.6}$$

where

$$\widehat{g}(\omega_1^k\sigma) := \begin{cases} g(\omega_1^{k-1}\sigma) & \text{if } \omega_k = +1\,, \\ 1 - g(\omega_1^{k-1}\sigma) & \text{if } \omega_k = -1\,. \end{cases} \tag{2.7}$$

By Kolmogorov's Extension Theorem, this defines ν^σ uniquely. To extend ν^σ to a measure on $(\Omega, \mathcal{F})$, write $\Omega = \{\pm 1\}^{\mathbb{Z}_-} \times \{\pm 1\}^{\mathbb{Z}_+}$, and define $\mu^\sigma \in \mathcal{M}$ by

$$\mu^\sigma := \delta_\sigma \otimes \nu^\sigma\,, \tag{2.8}$$

where δ_σ is the Dirac mass at σ. By the compactness of $\mathcal{M}$ we have existence, at least along a subsequence, of the weak limit

$$\mu^\sigma \circ T^{-n} \Rightarrow \mu_*^\sigma\,. \tag{2.9}$$

To show that μ_*^σ is specified by g, we fix $k \in \mathbb{Z}$ and use again (2.4):

$$\mu_*^\sigma([+]_{k+1} | \mathcal{F}_{(-\infty,k]}) = \lim_{l\to\infty} \mu_*^\sigma([+]_{k+1}|\mathcal{F}_{[-l,k]})$$

μ_*^σ-almost surely. By Lemma 2.1,

$$\begin{aligned} \mu_*^\sigma([+]_{k+1}|\mathcal{F}_{[-l,k]})(\omega) &= \frac{\mu_*^\sigma([+]_{k+1}[\omega]_{-l}^k)}{\mu_*^\sigma([\omega]_{-l}^k)} \\ &= \lim_{n\to\infty} \frac{\mu^\sigma([+]_{k+1+n}[T^{-n}\omega]_{-l+n}^{k+n})}{\mu^\sigma([T^{-n}\omega]_{-l+n}^{k+n})} \\ &= \lim_{n\to\infty} \frac{\nu^\sigma([+]_{k+1+n}[T^{-n}\omega]_{-l+n}^{k+n})}{\nu^\sigma([T^{-n}\omega]_{-l+n}^{k+n})} \end{aligned}$$

By summing over $[\eta]_1^{-l+n-1}$, the values of the configuration on the interval $[1,-l+n-1]$ (we are assuming $n \gg l$), we express the numerator as

$$\begin{aligned}&\sum \nu^\sigma([+]_{k+1+n}[T^{-n}\omega]_{-l+n}^{k+n}[\eta]_1^{-l+n-1})\\ &\qquad = \sum g((T^{-n}\omega)_{-l+n}^{k+n}\eta_1^{-l+n-1}\sigma)\nu^\sigma([T^{-n}\omega]_{-l+n}^{k+n}[\eta]_1^{-l+n-1})\,.\end{aligned}$$

The same resummation for $\nu^\sigma([T^{-n}\omega]_{-l+n}^{k+n})$ and

$$|g((T^{-n}\omega)_{-l+n}^{k+n}\eta_1^{-l+n-1}\sigma) - g((T^{-n}\omega)_1^{k+n}\sigma)| \leq \mathrm{var}_{k+l}(g)$$

lead to

$$\Big|\frac{\nu^\sigma([+]_{k+1+n}[T^{-n}\omega]_{-l+n}^{k+n})}{\nu^\sigma([T^{-n}\omega]_{-l+n}^{k+n})} - g((T^{-n}\omega)_1^{k+n}\sigma)\Big| \leq \mathrm{var}_{k+l}(g)\,.$$

Since, again by continuity,

$$\lim_{n\to\infty} g((T^{-n}\omega)_1^{k+n}\sigma) = g(\omega_{-\infty}^k)\,,$$

we have shown that $\mu_*^\sigma \in \mathcal{M}(g)$. Finally, in order to show that $\mathcal{M}_T(g) \neq \emptyset$, we use the measure μ_*^σ constructed above and define the averages

$$\mathcal{A}_n\mu_*^\sigma := \frac{1}{n}\sum_{k=0}^{n-1}\mu_*^\sigma \circ T^{-k}\,,$$

By the following lemma and since $\mathcal{M}(g)$ is convex, $\mathcal{A}_n\mu_*^\sigma \in \mathcal{M}(g)$ for all $n \geq 1$.

Lemma 2.3. *If $\nu \in \mathcal{M}(g)$, then $\nu \circ T^{-1} \in \mathcal{M}(g)$.*

We can thus consider some accumulation point $\mu_{\mathrm{inv}}^\sigma \in \mathcal{M}(g)$ such that

$$\mathcal{A}_n\mu_*^\sigma \Rightarrow \mu_{\mathrm{inv}}^\sigma \tag{2.10}$$

at least along a subsequence. Since

$$\mu_{\mathrm{inv}}^\sigma(T^{-1}B) = \mu_{\mathrm{inv}}^\sigma(B) + \lim_{n\to\infty}\frac{\mu_*^\sigma(T^{-n}B) - \mu_*^\sigma(B)}{n} = \mu_{\mathrm{inv}}^\sigma(B)$$

holds for all cylinder $B \in \mathcal{C}$, Lemma 2.2 implies that $\mu_{\mathrm{inv}}^\sigma \in \mathcal{M}_T$. Therefore, $\mathcal{M}(g) \cap \mathcal{M}_T \neq \emptyset$, which finishes the proof of Theorem 2.1. □

The probability measures μ_*^σ (and $\mu_{\mathrm{inv}}^\sigma$) constructed in the preceding proof are said to be prepared with the boundary condition σ, and an interesting question is to understand their dependence on σ. Observe that when there exists a unique measure specified by g (see Section 4), then all these measures coincide: $\mu_*^\sigma = \mu_*^{\sigma'}$ for all σ, σ'. This should be interpreted as a *loss of memory*: all the information contained in the boundary condition σ (at $-\infty$) is lost at $+\infty$. On the other side, when the dependence on the boundary condition is non-trivial, i.e. when there exist two boundary conditions σ, σ' for which $\mu_*^\sigma \neq \mu_*^{\sigma'}$, we say that there is a *phase transition*.

This phenomenon was exhibited for the first time, for a regular g-function, by Bramson and Kalikow [5]. It will be the subject of Section 5.

2.1. **Attractive g-functions.** We constructed the probability measure μ_*^σ, in (2.9), without knowing if it was invariant (due to the fact that the existence of the limit is guaranteed only along a subsequence), and the averaging procedure in (2.10) was necessary. In view of getting a control of the shift invariance after the first limit (2.9), we introduce a natural class of g-functions, whose additional property is similar to the one of specifications of random fields describing ferromagnets. We remind the usual partial order on Ω (or $\{\pm 1\}^{\mathbb{N}}$): $\sigma \leq \sigma'$ if and only if $\sigma_k \leq \sigma'_k$ for all k.

Definition 2.4. *A g-function g is* **attractive** [1] *if $g(\sigma) \leq g(\sigma')$ whenever $\sigma \leq \sigma'$.*

The g-functions defined in Examples 1.1 and 1.2 are both attractive. Systems specified by attractive g-functions have the property that the presence of pluses in the past favorizes the presence of pluses in the future. As will now be seen, this allows to say more about the two g-measures with conditions $\sigma \equiv +$ and $\sigma \equiv -$ respectively.

Proposition 2.1. *Assume g is regular and attractive. Let μ^+ and μ^- be defined as in* (2.8). *Then the sequences $(\mu^\pm \circ T^{-n})_{n\geq 1}$ have weakly limits,*

$$\mu^\pm \circ T^{-n} \Rightarrow \mu_*^\pm \,, \tag{2.11}$$

and these are invariant: $\mu_*^\pm \in \mathcal{M}_T(g)$.

Proof. Invariance will follow immediately from the existence of the weak limit (2.11). We treat the case μ^+. We start by showing existence, for each thin cylinder $[\sigma]_a^b$, of

$$\mu_*^+([+]_a^b) := \lim_{n\to\infty} \mu^+(T^{-n}[+]_a^b)\,. \tag{2.12}$$

This will allow to define $\mu_*^+(B)$ for any $B \in \mathcal{C}$, since the indicator of a cylinder can always be written as a linear combination of indicators of cylinders of the form $[+]_a^b$. The existence of (2.12) will follow by monotonicity: when k is large enough,

$$\mu^+(T^{-(k+1)}[+]_a^b) \leq \mu^+(T^{-k}[+]_a^b)\,. \tag{2.13}$$

Let us verify (2.13) in the simple case where $a = b = 0$ (the general case is proved similarly). We have $\mu^+(T^{-(k+1)}[+]_0) = \mu^+([+]_{k+1})$. Then, we condition with respect to the value taken by the spin at position $k = 1$:

$$\mu^+([+]_{k+1}) = \mu^+([+]_{k+1}[+]_1) + \mu^+([+]_{k+1}[-]_1)\,.$$

[1] This terminology apparently originated in the monograph of Preston [17], and was used again by Hulse in [13].

By the definition of μ^+ and the attractivity of g,

$$\mu^+([+]_{k+1}[-]_1) = \int_{[-]_1} g(\omega_2^k -_1 +_{-\infty}^0)\mu^+(d\omega)$$
$$\leq \int_{[-]_1} g(\omega_2^k +_1 +_{-\infty}^0)\mu^+(d\omega) \equiv \int_{[-]_1} g(\omega_2^k +_{-\infty}^1)\mu^+(d\omega)\,.$$

Since $[-]_1 = [+]_1^c$, the last integral equals

$$\int g(\omega_2^k +_{-\infty}^1)\mu^+(d\omega) - \int_{[+]_1} g(\omega_2^k +_{-\infty}^1)\mu^+(d\omega)$$

By a change of variable, the first integral equals

$$\int g(\omega_2^k +_{-\infty}^1)\mu^+(d\omega) = \int g(\omega_1^{k-1} +_{-\infty}^0)\mu^+(d\omega) = \mu^+([+]_k)\,.$$

For the second, we proceed as above to recompose

$$\int_{[+]_1} g(\omega_2^k +_{-\infty}^1)\mu^+(d\omega) = \mu^+([+]_{k+1}[+]_1)\,.$$

We have thus shown (2.13). To verify that μ_*^+ is σ-additive on $\mathcal{C}$, consider a decreasing family of cylinders such that $\bigcap_{n\geq 1} B_n = \emptyset$.

Lemma 2.4. *Let $\{B_n\}_{n\geq 1}$ be a decreasing ($B_{n+1} \subset B_n$) family of cylinders such that $\bigcap_{n\geq 1} B_n = \emptyset$. Then $B_n = \emptyset$ for all sufficiently large n*

By Lemma 2.4, $B_n = \emptyset$ for large enough n, which implies $\lim_{n\to\infty} \mu_*^+(B_n) = 0$. The Extension Theorem of Carathéodory finishes the proof. □

The standard extension argument used above will be used again in Sections 3.4 and 3.6. More properties of the measures $\mu_*^\pm$ can be found in [13].

3. Decompositions

In this section we take a closer look at the convex structure of the sets $\mathcal{M}(g)$ and $\mathcal{M}_T(g)$, where g is a regular g-function. More precisely, we consider the decomposition of any element of these sets into a convex (infinite dimensional) combination of extreme elements. Although such results are usually obtained via non-constructive functional-analytic arguments (Krein-Millman Theorem, see e.g. [18]), we will follow the measure-theoretic approach developped by Dynkin [8], simplified by Georgii [12].

3.1. **Heuristics.** Let $\mathcal{P} \subset \mathcal{M}$ be any nonempty set of probability measures. Since we will later be interested in the case where $\mathcal{P}$ is either $\mathcal{M}_T$, $\mathcal{M}(g)$ or $\mathcal{M}_T(g)$, we can consider the elements of $\mathcal{P}$ as being the measures of $\mathcal{M}$ which share a common property. In $\mathcal{M}_T$, the common property is "to be invariant under T"; in $\mathcal{M}(g)$ it is "to be specified by g".

Since $\mathcal{P}$ is convex in the three cases under consideration, it is natural to ask if it contains extreme elements, $\mathrm{ex}\,\mathcal{P} \subset \mathcal{P}$, and if each $\mu \in \mathcal{P}$ can be decomposed into a convex combination of these extreme elements. Since spaces of measures are infinite-dimensional, such convex combinations must involve some sort of probability measure on $\mathrm{ex}\,\mathcal{P}$, denoted α_μ. The decomposition we expect is of the following form:

$$\forall A \in \mathcal{F}\,, \quad \mu(A) = \int_{\mathrm{ex}\,\mathcal{P}} \nu(A)\alpha_\mu(d\nu)\,. \tag{3.1}$$

In a finite-dimensional setting, α_μ would correspond to the coefficients of the decomposition of μ, and $\alpha_\mu(\mathrm{ex}\,\mathcal{P}) = 1$ reflects the fact that these add up to 1.

In concrete cases, the common property satisfied by the measures of $\mathcal{P}$ has a natural sub-σ-algebra $\mathcal{A} \subset \mathcal{F}$ associated to it, and the extreme elements of $\mathcal{P}$ happen to have a simple characterization in terms of $\mathcal{A}$.

Definition 3.1. *Let $\mathcal{A} \subset \mathcal{F}$ be a sub-σ-algebra. A probability $\nu \in \mathcal{M}$ is trivial on $\mathcal{A}$ if $\nu(A) \in \{0,1\}$ for all $A \in \mathcal{A}$. If $\mathcal{P} \subset \mathcal{M}$, define*

$$\mathcal{P}_{\mathcal{A}} := \{\mu \in \mathcal{P} : \mu \text{ is trivial on } \mathcal{A}\}\,. \tag{3.2}$$

For example, when $\mathcal{P} = \mathcal{M}_T$, $\mathcal{A}$ is the σ-algebra of invariant events, which appears naturally in the Ergodic Theorem, and the extreme elements of $\mathcal{P}$ are the ergodic measures, which are trivial on $\mathcal{A}$ (see Section (3.4)). Therefore, we assume from now on that the extreme elements of $\mathcal{P}$ are characterized by their triviality on a sub-σ-algebra $\mathcal{A}$, and always denote $\mathrm{ex}\,\mathcal{P}$ by $\mathcal{P}_{\mathcal{A}}$. Observe that a priori, it seems non-trivial that $\mathcal{P}_{\mathcal{A}} \neq \emptyset$.

It happens that triviality on $\mathcal{A}$ is equivalent to conditionning with respect to $\mathcal{A}$, as shown in the next lemma.

Lemma 3.1. *Let $\mathcal{A} \subset \mathcal{F}$ be a sub-σ-algebra, $\mu \in \mathcal{M}$. Then μ is trivial on $\mathcal{A}$ if and only if for all $B \in \mathcal{F}$, $\mu(B|\mathcal{A}) = \mu(B)$ μ-a.s.*

Therefore, the description of $\mathcal{P}_{\mathcal{A}}$ is done by studying conditional probabilities of elements of $\mathcal{P}$ with respect to $\mathcal{A}$. So, to start with, let $\mu \in \mathcal{P}$ and assume the existence of a regular conditional distribution of μ with respect to $\mathcal{A}$. That is, assume

$$B \mapsto Q^\omega(B) := \mu(B|\mathcal{A})(\omega) \tag{3.3}$$

defines a probability measure for μ-almost all ω. At points ω at which Q^ω is not a probability measure, define Q^ω in an arbitrary way (For example, define $Q^\omega := \mu_0$, where μ_0 is a fixed probability measure in $\mathcal{M}$.) The identity $\mathsf{E}_\mu(1_B) = \mathsf{E}_\mu(\mathsf{E}_\mu(1_B|\mathcal{A}))$ can then be written as follows:

$$\mu(B) = \int Q^\omega(B)\mu(d\omega)\,. \tag{3.4}$$

This integral can already be interpreted as a decomposition of μ into a convex combination of the measures Q^ω, $\omega \in \Omega$. Nevertheless, we need to show that the relevant measures involved in the decomposition are, in the sense of μ, concentrated on $\mathcal{P}_\mathcal{A}$. A simple argument is in favor of this: observe that when $A \in \mathcal{A}$, then of course

$$Q^{\cdot}(A) = 1_A(\cdot) \quad \mu\text{-a.s.}\,. \tag{3.5}$$

Ideally, we would like a reversed statement: that for μ-almost all ω, $Q^\omega(A) \in \{0,1\}$ for each $A \in \mathcal{A}$:

$$\mu(\{\omega : Q^\omega \in \mathcal{P}_\mathcal{A}\}) = 1\,. \tag{3.6}$$

We will see later that this central concentration property does indeed hold in the cases in which we are interested. Assuming (3.6), the integral (3.4) can be restricted to $\{Q^{\cdot} \in \mathcal{P}_\mathcal{A}\}$:

$$\mu(B) = \int_{Q^{\cdot} \in \mathcal{P}_\mathcal{A}} Q^\omega(B)\mu(d\omega)\,. \tag{3.7}$$

Then, consider the measure-valued random variable $\mathsf{X} : \Omega \to \mathcal{M}$, $\mathsf{X}(\omega) := Q^\omega$. We can express μ as the expectation of X:

$$\mu = \int_{\mathsf{X} \in \mathcal{P}_\mathcal{A}} \mathsf{X}(\omega)\mu(d\omega) \tag{3.8}$$

Clearly, this requires the definition of a measurable structure on $\mathcal{M}$. In elementary probability, the expectation of a real-valued random variable $X : \Omega \to \mathbb{R}$ can be transformed into

$$\int_\Omega X(\omega)\mu(d\omega) = \int_\mathbb{R} x\mu_X(dx)\,, \tag{3.9}$$

where μ_X is the distribution of X, defined for any Borel set $I \subset \mathbb{R}$ by

$$\mu_X(I) := \mu(\{\omega : X(\omega) \in I\})\,.$$

The same can done for (3.8): for each measurable set $M \subset \mathcal{P}_\mathcal{A}$, define the distribution

$$\alpha_\mu(M) := \mu(\{\omega : \mathsf{X}(\omega) \in M\})\,.$$

Since $\alpha_\mu(\mathcal{P}_\mathcal{A}) = 1$ by (3.6), one can proceed as in (3.9) and push the integration of X onto $\mathcal{P}_\mathcal{A}$:

$$\int_{\mathsf{X} \in \mathcal{P}_\mathcal{A}} \mathsf{X}(\omega)\mu(d\omega) = \int_{\mathcal{P}_\mathcal{A}} \nu\alpha_\mu(d\nu)\,. \tag{3.10}$$

This gives the wanted decomposition of μ.

The above decomposition thus relies on the existence of the regular conditional distribution Q, and on the justification of the steps that led to the definition of the probability measure α_μ. In particular, (3.6) is crucial. Observe that this program cannot be guaranteed to succeed in general, and requires that $(\Omega, \mathcal{F})$ have some particular topological structure. In our case, our construction of Q will rely heavily on the fact that our σ-algebra $\mathcal{F}$ is generated by the cylinders, which are countable.

In Section 3.2 we define the appropriate measurable structure on sets of probability measures. In Section 3.3 we make the previous informal argument rigorous. In subsequent sections we apply this result to the three situations of interest in this paper, namely where $\mathcal{P}$ is respectively $\mathcal{M}_T$, $\mathcal{M}(g)$ and $\mathcal{M}_T(g)$.

3.2. **Measurable sets of Probability Measures.** Consider a non-empty subset $\mathcal{M}_0 \subset \mathcal{M}$. A standard way of defining a measurable structure on $\mathcal{M}_0$ is to consider a family of real functions on $\mathcal{M}_0$ and to require each of these to be measurable.

Definition 3.2. *For each $A \in \mathcal{F}$, the* evaluation map *$e_A : \mathcal{M}_0 \to [0,1]$ is defined by $e_A(\mu) := \mu(A)$. The* evaluation σ-algebra on *$\mathcal{M}_0$ is the smallest σ-algebra of subsets of $\mathcal{M}_0$ for which each evaluation map e_A, $A \in \mathcal{F}$, is measurable.*

Remark 3.1. In other words, $e(\mathcal{M}_0)$ is the σ-algebra generated by all the sets of the form $\{\nu \in \mathcal{M}_0 : e_A(\nu) \leq c\}$, for $A \in \mathcal{F}$, $c \in [0,1]$.

It will also be necessary to consider, for each bounded $\mathcal{F}$-measurable $f : \Omega \to \mathbb{R}$, the map $e_f : \mathcal{M}_0 \to \mathbb{R}$ defined by $e_f(\mu) := \mathsf{E}_\mu(f)$. By expressing f as a uniform limit of simple functions and using Dominated Convergence, one obtains

Lemma 3.2. *If $f : \Omega \to \mathbb{R}$ is bounded and measurable, then $e_f : \mathcal{M}_0 \to \mathbb{R}$ is $e(\mathcal{M}_0)$-measurable.*

We denote by $\mathcal{M}_1^+(\mathcal{M}_0, e(\mathcal{M}_0))$ the set of probability measures on the measurable space $(\mathcal{M}_0, e(\mathcal{M}_0))$. If $\alpha \in \mathcal{M}_1^+(\mathcal{M}_0, e(\mathcal{M}_0))$, the integral of an $e(\mathcal{M}_0)$-measurable function $F : \mathcal{M}_0 \to \mathbb{R}$ with respect to μ is denoted

$$\int_{\mathcal{M}_0} F(\nu)\alpha(d\nu)\,.$$

In particular, by choosing for each $A \in \mathcal{F}$ the evaluation map $F = e_A$, one can define a probability measure $\mu \in \mathcal{M}$, called the barycenter of α:

$$\forall A \in \mathcal{F}\,, \quad \mu(A) := \int_{\mathcal{M}_0} \nu(A)\alpha(d\nu)\,. \tag{3.11}$$

3.3. **The Abstract Decomposition.** This section is without reference to any particular structure on the underlying measurable space, besides $\mathcal{F}$ being countably generated. We will nevertheless continue using the notations used so far. Our exposition follows that of Georgii [12], with slight differences.

Regular conditional distributions are well described by probability kernels, of which re remind the definition.

Definition 3.3. *A probability kernel from* $(\Omega, \mathcal{A})$ *to* $(\Omega, \mathcal{F})$ *is a map* $Q : \Omega \times \mathcal{F} \to [0,1]$ *satisfying the following conditions:*

(1) *for all* $\omega \in \Omega$, $Q(\omega, \cdot) \in \mathcal{M}$,
(2) *for all* $B \in \mathcal{F}$, $Q(\cdot, B)$ *is* $\mathcal{A}$*-measurable.*

In this section, the symbol Q will always denote a probability kernel from $(\Omega, \mathcal{A})$ to $(\Omega, \mathcal{F})$. We will also write $Q^\omega(B)$ in place of $Q(\omega, B)$, and since the kernel defines a map that associates to each $\omega \in \Omega$ a probability measure Q^ω on $(\Omega, \mathcal{F})$, we denote it by $\mathsf{X} : \Omega \to \mathcal{M}$, $\mathsf{X}(\omega) := Q^\omega$.

Lemma 3.3. *Let* $\mathcal{M}$ *be endowed with the evaluation* σ*-algebra* $e(\mathcal{M})$. *Then* $\mathsf{X} : \Omega \to \mathcal{M}$ *is* $\mathcal{A}$*-measurable:* $\mathsf{X}^{-1}(M) \in \mathcal{A}$ *for all* $M \in e(\mathcal{M})$.

Proof. Define $\mathcal{D} := \{M \in e(\mathcal{M}) : \mathsf{X}^{-1}(M) \in \mathcal{A}\}$. It suffices (see Remark 3.1) to show that $\mathcal{D}$ contains all sets of the form $\{\nu \in \mathcal{M} : \nu(B) \le c\}$, with $B \in \mathcal{F}$. But $\mathsf{X}^{-1}(\{\nu \in \mathcal{M} : \nu(B) \le c\}) = \{\omega : Q^\omega(B) \le c\} \in \mathcal{A}$ by the $\mathcal{A}$-measurability of $\omega \mapsto Q^\omega(B)$. □

Our discussion of Section 3.1 should make the following definition natural.

Definition 3.4. *Let* $\mathcal{P} \subset \mathcal{M}$ *be non-empty and* $\mathcal{A} \subset \mathcal{F}$ *be a sub-*σ*-algebra. A probability kernel from* $(\Omega, \mathcal{A})$ *to* $(\Omega, \mathcal{F})$ *is called a superkernel for the pair* $(\mathcal{P}, \mathcal{A})$ *if* $\{Q^\cdot \in \mathcal{P}\} \in \mathcal{A}$, *and if for all* $\mu \in \mathcal{P}$ *the following conditions hold:*

(1) $\forall B \in \mathcal{F}$, $\mu(B|\mathcal{A})(\cdot) = Q^\cdot(B)$ μ*-a.s.*
(2) $\mu(Q^\cdot \in \mathcal{P}) = 1$.

Condition (1) means that the kernel Q is suited for the description of the conditional distribution of *each* $\mu \in \mathcal{P}$ with respect to $\mathcal{A}$, and Condition (2) will be necessary to justify (3.6). Now goes the main theorem.

Theorem 3.1. *Assume* $\mathcal{F}$ *is countably generated. Let* $\mathcal{P} \subset \mathcal{M}$ *be non-empty,* $\mathcal{A} \subset \mathcal{F}$ *a sub-*σ*-algebra, and let* $\mathcal{P}_\mathcal{A}$ *denote the set of elements of* $\mathcal{P}$ *which are trivial on* $\mathcal{A}$, *defined in* (3.2). *Suppose there exists a superkernel* Q *for the pair* $(\mathcal{P}, \mathcal{A})$. *Then*

(1) $\mathcal{P}_\mathcal{A} \neq \emptyset$, *and*
(2) *for all* $\mu \in \mathcal{P}$, $M \in e(\mathcal{P}_\mathcal{A})$, $\alpha_\mu(M) := \mu(Q^\cdot \in M)$ *is well defined,* $\alpha_\mu \in \mathcal{M}_1^+(\mathcal{P}_\mathcal{A}, e(\mathcal{P}_\mathcal{A}))$, *and the following decomposition holds:*

$$\forall B \in \mathcal{F}\,, \quad \mu(B) = \int_{\mathcal{P}_\mathcal{A}} \nu(B) \alpha_\mu(d\nu)\,. \tag{3.12}$$

The measure α_μ *is the unique measure of* $\mathcal{M}_1^+(\mathcal{P}_\mathcal{A}, e(\mathcal{P}_\mathcal{A}))$ *whose barycenter is* μ, *i.e. for which* (3.12) *holds.*

The following maps will be useful in the proof. For all $B \in \mathcal{F}$, define the variance of $Q^{\cdot}(B)$, $v_B : \mathcal{M} \to \mathbb{R}$, by

$$\begin{aligned} v_B(\mu) :=& \mathsf{E}_\mu\left[(Q^{\cdot}(B) - \mathsf{E}_\mu(Q^{\cdot}(B)))^2\right] \\ =& \mathsf{E}_\mu[Q^{\cdot}(B)^2] - \mathsf{E}_\mu[Q^{\cdot}(B)]^2 \\ \equiv& e_{Q^{\cdot}(B)^2}(\mu) - e_{Q^{\cdot}(B)}(\mu)^2 \end{aligned}$$

Since $\omega \mapsto Q^\omega(B)$ is measurable and bounded, Lemma 3.2 says that v_B is $e(\mathcal{M})$-measurable. Further,

$$v_B(\mu) = \mathsf{E}_\mu[(Q^{\cdot}(B) - \mu(B))^2] \quad \forall \mu \in \mathcal{P}\,. \tag{3.13}$$

Namely, if $\mu \in \mathcal{P}$ then $\mathsf{E}_\mu[Q^{\cdot}(B)] = \mathsf{E}_\mu[\mathsf{E}_\mu(1_B|\mathcal{A})] = \mathsf{E}_\mu(1_B) = \mu(B)$. The core of the proof given below is to prove the following simple characterization of $\mathcal{P}_\mathcal{A}$ in terms of these variances: $\mu \in \mathcal{P}_\mathcal{A}$ if and only if $\mu \in \mathcal{P}$ and $v_B(\mu) = 0$ for all cylinder $B \in \mathcal{C}$.

Proof of Theorem 3.1. Denote by $\mathcal{C}$ the countable generator of $\mathcal{F}$. The proof is divided in three parts.
Step 1: Study of the set $\mathcal{P}_\mathcal{A}$. We characterize $\mathcal{P}_\mathcal{A}$ using the maps v_B:

$$\begin{aligned} \mathcal{P}_\mathcal{A} &= \{\mu \in \mathcal{P} : \forall B \in \mathcal{F}\,, \mu(B|\mathcal{A}) = \mu(B) \ \mu\text{-a.s.}\} \\ &= \{\mu \in \mathcal{P} : \forall B \in \mathcal{F}\,, Q^{\cdot}(B) = \mu(B) \ \mu\text{-a.s.}\} \\ &= \{\mu \in \mathcal{P} : \forall B \in \mathcal{C}\,, Q^{\cdot}(B) = \mu(B) \ \mu\text{-a.s.}\} && (3.14) \\ &= \{\mu \in \mathcal{P} : \forall B \in \mathcal{C}, v_B(\mu) = 0\} && (3.15) \\ &\equiv \bigcap_{B \in \mathcal{C}} \{\mu \in \mathcal{P} : v_B(\mu) = 0\}\,. && (3.16) \end{aligned}$$

The first equality follows from Lemma 3.1, the second from the definition of the superkernel, and the third from the fact that $\{B \in \mathcal{F} : Q^{\cdot}(B) = \mu(B)\,, \mu - a.s.\}$ is a Dynkin system containing $\mathcal{C}$ and from Theorem B.1. In (3.15) we used (3.13). For the same reason as before, the restriction $v_B : \mathcal{P} \to \mathbb{R}$ is $e(\mathcal{P})$-measurable. Therefore, (3.16) shows that $\mathcal{P}_\mathcal{A} \in e(\mathcal{P})$. Using again (3.16), we have

$$\begin{aligned} \{Q^{\cdot} \in \mathcal{P}_\mathcal{A}\} &= \bigcap_{B \in \mathcal{C}} \{Q^{\cdot} \in \mathcal{P}, v_B(Q^{\cdot}) = 0\} \\ &= \{Q^{\cdot} \in \mathcal{P}\} \cap \bigcap_{B \in \mathcal{C}} \mathsf{X}^{-1}\{\mu \in \mathcal{P} : v_B(\mu) = 0\} \\ &= \{Q^{\cdot} \in \mathcal{P}\} \cap \bigcap_{B \in \mathcal{C}} \mathsf{X}^{-1}\{\mu \in \mathcal{M} : v_B(\mu) = 0\}\,. && (3.17) \end{aligned}$$

Since we assumed $\{Q^{\cdot} \in \mathcal{P}\} \in \mathcal{A}$ this implies, using measurability of v_B and of X (Lemma 3.3) in each of the terms of the intersection over $B \in \mathcal{C}$, that $\{Q^{\cdot} \in \mathcal{P}_\mathcal{A}\} \in \mathcal{A}$.

Step 2: The probability measure α_μ. Let $\mu \in \mathcal{P}$. Let, for $M \in e(\mathcal{P}_\mathcal{A})$,

$$\alpha_\mu(M) := \mu(\{\omega : \mathsf{X}(\omega) \in M\}) = \mu(\mathsf{X}^{-1}(M)) . \tag{3.18}$$

Let us verify that this set function is well defined, i.e. that $\mathsf{X}^{-1}(M) \in \mathcal{A}$ for all $M \in e(\mathcal{P}_\mathcal{A})$. We proceed as in the proof of Lemma 3.1. Let $\mathcal{D} := \{M \in e(\mathcal{P}_\mathcal{A}) : \mathsf{X}^{-1}(M) \in \mathcal{A}\}$. Since $\mathcal{D} \subset e(\mathcal{P}_\mathcal{A})$ and $\mathcal{D}$ is a σ-algebra on $\mathcal{P}_\mathcal{A}$ (as can be easily verified), it suffices to see that $\mathcal{D}$ contains all sets of the form $\{\nu \in \mathcal{P}_\mathcal{A} : e_B(\nu) \leq c\}$, with $B \in \mathcal{F}$, $c \in [0,1]$ (see Remark 3.1). But

$$\mathsf{X}^{-1}(\{\nu \in \mathcal{P}_\mathcal{A} : e_B(\nu) \leq c\}) = \{Q^\cdot \in \mathcal{P}_\mathcal{A}\} \cap \{Q^\cdot(B) \leq c\} \in \mathcal{A} ,$$

by Step 1 and since $\omega \mapsto Q^\omega(B)$ is $\mathcal{A}$-measurable. This shows that α_μ is well defined. It is immediate from (3.18) that α_μ is σ-additive. The point is to verify that α_μ is a probability, i.e. that X concentrates on $\mathcal{P}_\mathcal{A}$ (remember (3.6)):

$$\alpha_\mu(\mathcal{P}_\mathcal{A}) = \mu(\mathsf{X} \in \mathcal{P}_\mathcal{A}) = \mu(Q^\cdot \in \mathcal{P}_\mathcal{A}) = 1 .$$

Consider the event $\{Q^\cdot \in \mathcal{P}_\mathcal{A}\}$ expressed as in (3.17). Since $\mu(Q^\cdot \in \mathcal{P}) = 1$ by hypothesis, it remains to show that for all $B \in \mathcal{C}$,

$$\mu(\mathsf{X}^{-1}\{\mu \in \mathcal{M} : v_B(\mu) = 0\}) \equiv \mu(v_B(Q^\cdot) = 0) = 1 .$$

Since $v_B(Q^\cdot) \geq 0$, it is enough to show that $\mathsf{E}_\mu(v_B(Q^\cdot)) = 0$. Remembering that $\mu(Q^\cdot \in \mathcal{P}) = 1$ and using (3.13), we get

$$\begin{aligned} \mathsf{E}_\mu(v_B(Q^\cdot)) &= \int_{Q^\cdot \in \mathcal{P}} v_B(Q^\omega)\mu(d\omega) \\ &= \int_{Q^\cdot \in \mathcal{P}} [\mathsf{E}_{Q^\omega}(Q^\cdot(B)^2) - Q^\omega(B)^2]\mu(d\omega) \\ &= \int [\mathsf{E}_{Q^\omega}(Q^\cdot(B)^2) - Q^\omega(B)^2]\mu(d\omega) . \end{aligned} \tag{3.19}$$

Now, observe that for all bounded measurable $f : \Omega \to \mathbb{R}$,

$$\int \mathsf{E}_{Q^\omega}(f)\mu(d\omega) = \int \mathsf{E}_\mu(f|\mathcal{A})(\omega)\mu(d\omega) .$$

This follows from the definition of $Q^\cdot$. Namely, for $f = 1_E$, $E \in \mathcal{F}$,

$$\int \mathsf{E}_{Q^\omega}(1_E)\mu(d\omega) = \int Q^\omega(E)\mu(d\omega) \equiv \int \mathsf{E}_\mu(1_E|\mathcal{A})(\omega)\mu(d\omega) .$$

The extension to bounded functions is standard. Therefore, when $f = Q^\cdot(B)^2$,

$$\int \mathsf{E}_{Q^\omega}(Q^\cdot(B)^2)\mu(d\omega) = \int \mathsf{E}_\mu(Q^\cdot(B)^2|\mathcal{A})(\omega)\mu(d\omega) = \int Q^\omega(B)^2\mu(d\omega) .$$

This shows that (3.19), i.e. $\mathsf{E}_\mu(v_B(Q^\cdot))$, is zero.

Step 3: The decomposition and its uniqueness. Let $\mu \in \mathcal{P}$. To represent μ as the barycenter of α_μ, proceed as follows (the idea was already explained in the heuristic discussion of Section 3.1). First, express (3.18) using indicator functions:

$$\int_{\mathcal{P}_{\mathcal{A}}} 1_M(\nu)\alpha_\mu(d\nu) = \int 1_M(Q^\omega)\mu(d\omega)\,.$$

By a standard extension argument, this implies that for any bounded $e(\mathcal{P}_{\mathcal{A}})$-measurable function $F : \mathcal{P}_{\mathcal{A}} \to \mathbb{R}$,

$$\int_{\mathcal{P}_{\mathcal{A}}} F(\nu)\alpha_\mu(d\nu) = \int F(Q^\omega)\mu(d\omega)\,.$$

In particular, for the evaluation maps $F = e_B$, $B \in \mathcal{F}$, one gets

$$\int_{\mathcal{P}_{\mathcal{A}}} \nu(B)\alpha_\mu(d\nu) = \int Q^\omega(B)\mu(d\omega)\,.$$

We have thus shown (3.12). To verify that α_μ is the unique measure of $\mathcal{M}_1^+(\mathcal{P}_{\mathcal{A}}, e(\mathcal{P}_{\mathcal{A}}))$ representing μ, observe, first, that

$$\{Q^{\cdot} = \mu\} = \{Q^{\cdot}(B) = \mu(B)\,\forall B \in \mathcal{C}\}$$

This follows from the Extension Theorem of Carathéodory. So (3.14) can also be written as:

$$\mathcal{P}_{\mathcal{A}} = \{\mu \in \mathcal{P} : \mu(Q^{\cdot} = \mu) = 1\}\,. \tag{3.20}$$

Assume $\alpha'_\mu \in \mathcal{M}_1^+(\mathcal{P}_{\mathcal{A}}, e(\mathcal{P}_{\mathcal{A}}))$ also represents μ. Consider any $M \in e(\mathcal{P}_{\mathcal{A}})$ and write

$$\alpha'_\mu(M) = \int_{\mathcal{P}_{\mathcal{A}}} 1_M(\nu)\alpha'_\mu(d\nu)\,.$$

Since $\nu \in \mathcal{P}_{\mathcal{A}}$, (3.20) clearly gives $\nu(Q^{\cdot} \in M) = 1_M(\nu)$. Since we are also assuming that α'_μ represents μ,

$$\int_{\mathcal{P}_{\mathcal{A}}} 1_M(\nu)\alpha'_\mu(d\nu) = \int_{\mathcal{P}_{\mathcal{A}}} \nu(Q^{\cdot} \in M)\alpha'_\mu(d\nu) = \mu(Q^{\cdot} \in M) \equiv \alpha_\mu(M)\,,$$

which implies $\alpha_\mu(M) = \alpha'_\mu(M)$. This proves Theorem 3.1. □

3.4. **Extreme Decomposition for $\mathcal{M}_T$.** We apply the general result of last section to the case where $\mathcal{P} = \mathcal{M}_T$, the set of measures which are invariant under T:

$$\mathcal{M}_T = \{\mu \in \mathcal{M} : \mu \circ T^{-1} = \mu\}\,.$$

The central ingredient, in the study of invariant measures, is the Ergodic Theorem of Birkhoff, which we state for the sake of completeness:

Theorem 3.2. *Let $\mu \in \mathcal{M}_T$. For any $f \in L^1(\Omega, \mathcal{F}, \mu)$, as $n \to \infty$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k \to \mathsf{E}_\mu(f|\mathcal{I}) \quad \mu\text{-a.s.} \tag{3.21}$$

Here, $\mathcal{I}$ is the σ-algebra of invariant sets, defined by

$$\mathcal{I} := \{A \in \mathcal{F} : TA = A\} .$$

It is well known (see e.g. [3]) that the extreme elements of $\mathcal{M}_T$ are the invariant measures which are trivial on $\mathcal{I}$. We therefore define

$$\mathcal{M}_T^{\mathrm{erg}} := \{\mu \in \mathcal{M}_T : \mu \text{ is trivial on } \mathcal{I}\} . \tag{3.22}$$

In the notations of the previous section, $\mathcal{M}_T^{\mathrm{erg}} = (\mathcal{M}_T)_{\mathcal{I}}$. The elements of $\mathcal{M}_T^{\mathrm{erg}}$ are called ergodic measures. We can now state the well-known result saying that each invariant measure can be decomposed into a convex combination of ergodic measures.

Theorem 3.3. *Let $\mu \in \mathcal{M}_T$. Then there exists a unique probability measure $\alpha_\mu \in \mathcal{M}_1^+(\mathcal{M}_T^{\mathrm{erg}}, e(\mathcal{M}_T^{\mathrm{erg}}))$, such that*

$$\forall B \in \mathcal{F}, \quad \mu(B) = \int_{\mathcal{M}_T^{\mathrm{erg}}} \nu(B)\alpha_\mu(d\nu) . \tag{3.23}$$

Each integral in (3.23) *is called the ergodic decomposition of $\mu(B)$.*

This theorem is a direct application of Theorem 3.1, once we have a superkernel for the pair $(\mathcal{M}_T, \mathcal{I})$. As in the construction of measures specified by a given g-function, we will rely on the compactness of Ω.

Proposition 3.1. *There exists a superkernel for the pair $(\mathcal{M}_T, \mathcal{I})$, denoted Q.*

Proof. For each ω, $n \geq 1$, define the empirical measure

$$Q_n^\omega := \frac{1}{n}\sum_{k=0}^{n-1} \delta_{T^k\omega} ,$$

where δ_ω is the Dirac mass at ω. Taking $B \in \mathcal{F}$, one can also write

$$Q_n^\omega(B) = \frac{1}{n}\sum_{k=0}^{n-1} 1_B \circ T^k(\omega) . \tag{3.24}$$

Consider the set

$$\Omega_0 := \bigcap_{B \in \mathcal{C}} \bigl\{\omega \in \Omega : \lim_{n\to\infty} Q_n^\omega(B) \text{ exists}\bigr\} . \tag{3.25}$$

The following can be verified easily: for each $\omega \in \Omega_0$, $\lim_{n\to\infty} Q_n^{T^{-1}\omega}(B)$ exists for all $B \in \mathcal{C}$ and equals

$$\lim_{n\to\infty} Q_n^{T^{-1}\omega}(B) = \lim_{n\to\infty} Q_n^\omega(B) . \tag{3.26}$$

In particular, $\Omega_0 \in \mathcal{I}$. Let μ_0 be an arbitrary probability measure in $\mathcal{M}_T$ (for example, a product measure). Define, for each $\omega \in \Omega$ and each cylinder $B \in \mathcal{C}$,

$$Q^\omega(B) := \begin{cases} \lim_{n\to\infty} Q^\omega_n(B) & \text{if } \omega \in \Omega_0\,, \\ \mu_0(B) & \text{otherwise}\,. \end{cases}$$

Lemma 3.4. *For all ω, Q^ω can be extended uniquely to a probability measure on $(\Omega, \mathcal{F})$.*

Proof. When $\omega \in \Omega_0^c$, there is nothing to show: just set $Q^\omega(B) := \mu_0(B)$ for all $B \in \mathcal{F}$. When $\omega \in \Omega_0$, we see that Q^ω is finitely additive by considering, for two disjoint cylinders B_1, B_2,

$$\begin{aligned} Q^\omega(B_1 \cup B_2) &= \lim_{n\to\infty} Q^\omega_n(B_1 \cup B_2) \\ &= \lim_{n\to\infty} Q^\omega_n(B_1) + \lim_{n\to\infty} Q^\omega_n(B_2) = Q^\omega(B_1) + Q^\omega(B_2)\,. \end{aligned}$$

To verify σ-additivity, consider a decreasing family of cylinders such that $\bigcap_{n\geq 1} B_n = \emptyset$. By Lemma 2.4, $B_n = \emptyset$ for large enough n, which implies $\lim_{n\to\infty} Q^\omega(B_n) = 0$. By the Extension Theorem of Carathéodory, there exists a unique extension of Q^ω. □

Let $B \in \mathcal{C}$, $c \in [0,1]$. By (3.26), $\{\omega : Q^\omega(B) \leq c\}$ is invariant. This shows that $\omega \to Q^\omega(B)$ is $\mathcal{I}$-measurable. Since, as can be easily verified, $\{B \in \mathcal{F} : \omega \mapsto Q^\omega(B) \text{ is } \mathcal{I}\text{-measurable }\}$ is a Dynkin system containing the cylinders, and since the cylinders are stable under intersection, this shows that $\omega \mapsto Q^\omega(B)$ is $\mathcal{I}$-measurable for all $B \in \mathcal{F}$ (Theorem B.1). Therefore, Q is a probability kernel from $(\Omega, \mathcal{I})$ to $(\Omega, \mathcal{F})$.

We must then verify that the kernel Q provides, for each $\mu \in \mathcal{M}_T$, a regular conditional distribution with respect to $\mathcal{I}$. So take $\mu \in \mathcal{M}_T$, let $B \in \mathcal{C}$ and write $Q^\cdot(B)$ as in (3.24). By the Ergodic Theorem, as $n \to \infty$,

$$Q^\cdot_n(B) \to \mu(B|\mathcal{I})\,, \quad \mu\text{-a.s.}$$

Therefore, $\mu(B|\mathcal{I}) = Q^\cdot(B)$ μ-a.s. This also implies $\mu(\Omega_0) = 1$, and again, since $\{B \in \mathcal{F} : \mu(B|\mathcal{I}) = Q^\cdot \ \mu\text{-a.s.}\}$ is a Dynkin system containing the cylinders, we have $\mu(B|\mathcal{I}) = Q^\cdot(B)$ μ-a.s. for all $B \in \mathcal{F}$, which is Condition 1 of Definition 3.4. Now

$$\{Q^\cdot \in \mathcal{M}_T\} = \bigcap_{B\in\mathcal{C}} \{Q^\cdot(T^{-1}B) = Q^\cdot(B)\}\,. \tag{3.27}$$

This identity follows from Lemma 2.2. Since $\{Q^\cdot(T^{-1}B) = Q^\cdot(B)\} \in \mathcal{I}$ and $\mu(Q^\cdot(T^{-1}B) = Q^\cdot(B)) = 1$ for each $B \in \mathcal{C}$, (3.27) implies $\{Q^\cdot \in \mathcal{M}_T\} \in \mathcal{I}$ and $\mu(Q^\cdot \in \mathcal{M}_T) = 1$, which is Condition (2) of Definition 3.4. We have thus shown that Q is a superkernel for the pair $(\mathcal{M}_T, \mathcal{I})$. □

The kernel Q constructed above will be used again in Section 3.6.

3.5. **Extreme Decomposition for $\mathcal{M}(g)$.** In this section, we consider the case where $\mathcal{P}$ is $\mathcal{M}(g)$,

$$\mathcal{M}(g) := \{\mu \in \mathcal{M} : \mu \text{ is specified by } g\} ,$$

where g is a regular g-function. We will adapt the method exposed by Georgii in [12] for Gibbs measures and specifications on $\mathbb{Z}^d$.

The property defining elements $\mu \in \mathcal{M}(g)$ involves a conditionning on each of the pasts $\mathcal{F}_{(-\infty,k]}$. Therefore, a natural candidate for the σ-algebra which we need to study extreme elements of $\mathcal{M}(g)$ is the left-tail-σ-field:

$$\mathcal{T}_{-\infty} := \bigcap_{k\geq 1} \mathcal{F}_{(-\infty,-k]} ,$$

It was shown in [10] that indeed, the extreme elements of $\mathcal{M}(g)$ are characterized by triviality on $\mathcal{T}_{-\infty}$:

$$\operatorname{ex}\mathcal{M}(g) = \{\mu \in \mathcal{M}(g) : \mu \text{ is trivial on } \mathcal{T}_{-\infty}\} .$$

Theorem 3.4. *Let $\mu \in \mathcal{M}(g)$, where g is regular. Then there exists a unique probability measure $\pi_\mu \in \mathcal{M}_1^+(\operatorname{ex}\mathcal{M}(g), e(\operatorname{ex}\mathcal{M}(g)))$, such that*

$$\forall B \in \mathcal{F}, \quad \mu(B) = \int_{\operatorname{ex}\mathcal{M}(g)} \nu(B)\pi_\mu(d\nu) . \tag{3.28}$$

In particular, $\operatorname{ex}\mathcal{M}(g) \neq \emptyset$.

Again, this theorem is a direct application of Theorem 3.1, once we have a superkernel for the pair $(\mathcal{M}(g), \mathcal{T}_{-\infty})$.

Proposition 3.2. *Let g be regular. Then there exists a superkernel for the pair $(\mathcal{M}(g), \mathcal{T}_{-\infty})$, denoted* Π.

In the last section, the central ingredient in the construction of the superkernel for the pair $(\mathcal{M}_T, \mathcal{I})$ was the Ergodic Theorem. Here, the key convergence result will be the Backward Martingale Convergence Theorem: for all $A \in \mathcal{F}$,

$$\mu(A|\mathcal{T}_{-\infty}) = \lim_{n\to\infty} \mu(A|\mathcal{F}_{(-\infty,-n]}) \quad \mu\text{-a.s.} \tag{3.29}$$

Proof of Proposition 3.2: Let $\omega \in \Omega$, $n \geq 1$, $b \geq -n$. Define first, for a cylinder $[\sigma]_{-n+1}^b$,

$$\gamma_{-n}^\omega([\sigma]_{-n+1}^b) := \prod_{j=-n+1}^{b} \widehat{g}(\sigma_{-n+1}^j \omega_{-\infty}^n) , \tag{3.30}$$

where $\widehat{g}$ was defined in (2.7). Then, for any thin cylinder $[\sigma]_a^b$, $-n < a \leq b$,

$$\Pi_{-n}^\omega([\sigma]_a^b) := \sum_{[\eta]_{-n+1}^{a-1}} \gamma_{-n}^\omega([\sigma]_a^b[\eta]_{-n+1}^{a-1}) . \tag{3.31}$$

The definition of Π^ω_{-n} extends immediately to any cylinder $B \in \mathcal{C}$ whose base lies in $[-n+1, +\infty)$. Then, define

$$\Omega_1 := \bigcap_{B\in\mathcal{C}} \left\{\omega : \lim_{n\to\infty} \Pi^\omega_{-n}(B) \text{ exists}\right\}.$$

Clearly, $\omega \mapsto \Pi^\omega_{-n}(B)$ is $\mathcal{F}_{(-\infty,-n]}$-measurable for each thin cylinder, and so $\{\omega : \lim_{n\to\infty} \Pi^\omega_{-n}(B) \text{ exists}\} \in \mathcal{T}_{-\infty}$ for each $B \in \mathcal{C}$, which implies that $\Omega_1 \in \mathcal{T}_{-\infty}$. Let μ_0 be an arbitrary probability measure in $\mathcal{M}_T$. To be precise, take a product measure for which $\mu_0([-]_k) = 1 - \mu_0([+]_k) = \epsilon$ for all k, where ϵ is the constant that appears in the uniform non-nullness of g. Define, for each $\omega \in \Omega$ and each cylinder $B \in \mathcal{C}$,

$$\Pi^\omega(B) := \begin{cases} \lim_{n\to\infty} \Pi^\omega_n(B) & \text{if } \omega \in \Omega_1\,, \\ \mu_0(B) & \text{otherwise}\,. \end{cases}$$

As was done in the proof of Proposition (3.1), Π^ω can be uniquely extended to a probability measure on $(\Omega, \mathcal{F})$. To verify that $\omega \mapsto \Pi^\omega(B)$ is $\mathcal{T}_{-\infty}$-measurable for each $B \in \mathcal{F}$, start by writing, for $B \in \mathcal{C}$ and $c \in [0,1]$,

$$\begin{aligned} \{\omega : &\Pi^\omega(B) \le c\} = \big(\{\omega : \Pi^\omega(B) \le c\} \cap \Omega_1\big) \cup \big(\{\omega : \Pi^\omega(B) \le c\} \cap \Omega_1^c\big) \\ &= \big(\{\omega : \mu_0(B) \le c\} \cap \Omega_1\big) \cup \big(\{\omega : \lim_{n\to\infty} \Pi^\omega_{-n}(B) \le c\} \cap \Omega_1^c\big)\,, \end{aligned}$$

from which $\omega \mapsto \Pi^\omega(B)$ is obviously $\mathcal{T}_{-\infty}$-measurable. As in the proof of Proposition (3.1), it can be shown that this measurability extends to all $B \in \mathcal{F}$. Π is therefore a probability kernel from $(\Omega, \mathcal{T}_{-\infty})$ to $(\Omega, \mathcal{F})$. We show in a few steps that Π is a superkernel for the pair $(\mathcal{M}(g), \mathcal{T}_{-\infty})$.

Claim 3.1. *Let $\mu \in \mathcal{M}(g)$. Then $\forall A \in \mathcal{F}$, $\mu(A|\mathcal{T}_{-\infty}) = \Pi^{\cdot}(A)$ μ-a.s.*

Proof. Let $\mu \in \mathcal{M}(g)$. Consider a thin cylinder $[\sigma]_a^b$. By the Backward Martingale Convergence Theorem (3.29),

$$\mu([\sigma]_a^b|\mathcal{T}_{-\infty}) = \lim_{l\to\infty} \mu([\sigma]_a^b|\mathcal{F}_{(-\infty,-l]}) \quad \mu\text{-a.s.}$$

By summing over the possible values of the configuration on $[-l+1a-1]$,

$$\mu([\sigma]_a^b|\mathcal{F}_{(-\infty,-l]}) = \sum_{[\eta]_{-l+1}^{a-1}} \mu([\sigma]_a^b[\eta]_{-l+1}^{a-1}|\mathcal{F}_{(-\infty,-l]}) \quad \mu\text{-a.s.}$$

Lemma 3.5. *Let $\mu \in \mathcal{M}(g)$. Then for all thin cylinder $[\sigma]_{-n+1}^b$,*

$$\mu([\sigma]_{-n+1}^b|\mathcal{F}_{(-\infty,-n]}) = \gamma^{\cdot}_{-n}([\sigma]_{-n+1}^b)\,, \quad \mu\text{-a.s.} \tag{3.32}$$

Therefore, μ-a.s.,

$$\sum_{[\eta]_{-l+1}^{a-1}} \mu([\sigma]_a^b[\eta]_{-l+1}^{a-1}|\mathcal{F}_{(-\infty,-l]}) = \sum_{[\eta]_{-l+1}^{a-1}} \gamma^{\cdot}_{-l}([\sigma]_a^b[\eta]_{-l+1}^{a-1}) \equiv \Pi^{\cdot}_{-l}([\sigma]_a^b)\,.$$

This gives

$$\mu([\sigma]_a^b|\mathcal{T}_{-\infty}) = \lim_{l\to\infty} \Pi^{\cdot}_{-l}([\sigma]_a^b) = \Pi^{\cdot}([\sigma]_a^b) \quad \mu\text{-a.s.} \tag{3.33}$$

This extends to all cylinder $B \in \mathcal{C}$. It can easily be verified that $\mathcal{D} = \{A \in \mathcal{F} : \mu(A|\mathcal{T}_{-\infty}) = \Pi^{\cdot}(A)\, \mu\text{-a.s.}\}$ is a Dynkin system. Since it contains the cylinders, the claim is proved. □

An important consequence of (3.33) is that $\mu(\Omega_1) = 1$.

Claim 3.2. *For all $\omega \in \Omega_1$, $\Pi^\omega \in \mathcal{M}(g)$.*

Proof. Let $\omega \in \Omega_1$. By (2.4) we have, for Π^ω-almost all σ,

$$\begin{aligned}\Pi^\omega([+]_{k+1}|\mathcal{F}_{(-\infty,k]})(\sigma) &= \lim_{l\to\infty} \Pi^\omega([+]_{k+1}|\mathcal{F}_{[-l,k]})(\sigma)\\ &= \lim_{l\to\infty}\lim_{n\to\infty} \frac{\Pi^\omega_{-n}([+]_{k+1}[\sigma]_{-l}^k)}{\Pi^\omega_{-n}([\sigma]_{-l}^k)}\end{aligned}$$

The quotient is well defined by

Lemma 3.6. *For all ω, $\Pi^\omega([\sigma]_a^b) \geq \epsilon^{b-a+1} > 0$.*

As can be verified by direct computation using the definition of Π^ω_{-n},

$$\Big|\frac{\Pi^\omega_{-n}([+]_{k+1}[\sigma]_{-l}^k)}{\Pi^\omega_{-n}([\sigma]_{-l}^k)} - g(\sigma_{-n+1}^k\omega_{-\infty}^{-n})\Big| \leq \mathrm{var}_{k+l}(g)\,, \tag{3.34}$$

uniformly in n. This shows that

$$\Pi^\omega([+]_{k+1}|\mathcal{F}_{(-\infty,k]}) = g(\sigma_{-\infty}^k) \quad ,$$

and therefore $\Pi^\omega \in \mathcal{M}(g)$. □

Claim 3.3. $\{\Pi^{\cdot} \in \mathcal{M}(g)\} \in \mathcal{T}_{-\infty}$

Proof. Express $\{\omega : \Pi^\omega \in \mathcal{M}(g)\}$ as follows:

$$\begin{aligned}&\bigcap_{k\in\mathbb{Z}}\bigcap_{A\in\mathcal{F}_{(-\infty,k]}} \Big\{\omega : \Pi^\omega([+]_{k+1}\cap A) = \int_A g(\sigma_{-\infty}^k)\Pi^\omega(d\sigma)\Big\}\\ &\qquad = \bigcap_{k\in\mathbb{Z}}\bigcap_{B\in\mathcal{C}_{(-\infty,k]}} \Big\{\omega : \Pi^\omega([+]_{k+1}\cap B) = \int_B g(\sigma_{-\infty}^k)\Pi^\omega(d\sigma)\Big\}\,.\end{aligned}$$

This identity follows from the fact that for all k and all ω, the collection

$$\mathcal{D}_k^\omega := \Big\{A \in \mathcal{F}_{(-\infty,k]} : \Pi^\omega([+]_{k+1}\cap A) = \int_A g(\sigma_{-\infty}^k)\Pi^\omega(d\sigma)\Big\}$$

is a Dynkin system. But it is then easy to verify that

$$\Big\{\omega : \Pi^\omega([+]_{k+1}\cap B) < \int_B g(\sigma_{-\infty}^k)\Pi^\omega(d\sigma)\Big\} \in \mathcal{T}_{-\infty}\,,$$

which follows from the $\mathcal{T}_{-\infty}$-measurability of $\Pi^{\cdot}$. The other set with $>$ in place of $<$ belongs to $\mathcal{T}_{-\infty}$ for the same reason. □

Claim 3.1 shows that Π satisfies (1) of Definition 3.4. Moreover, since $\mu(\Omega_1) = 1$ and since $\Omega_1 \subset \{\Pi^{\cdot} \in \mathcal{M}(g)\}$ by Claim 3.2, we have $\mu(\Pi^{\cdot} \in \mathcal{M}(g)) = 1$, which is (2) of Definition 3.4. This finishes the proof. $\square$

3.6. **Extreme Decomposition for $\mathcal{M}_T(g)$.** Finally, we consider the case where $\mathcal{P} = \mathcal{M}_T(g)$, whose extreme elements we denote by $\operatorname{ex}\mathcal{M}_T(g)$. Surprisingly, we won't need to go through the construction of a new superkernel.

Theorem 3.5. *Let $\mu \in \mathcal{M}_T(g)$, where g is regular. Consider the probability measure $\alpha_\mu \in \mathcal{M}_1^+(\mathcal{M}_T^{\mathrm{erg}}, e(\mathcal{M}_T^{\mathrm{erg}}))$ of Theorem 3.3. Then*

$$\forall B \in \mathcal{F}, \quad \mu(B) = \int_{\operatorname{ex}\mathcal{M}_T(g)} \nu(B)\alpha_\mu(d\nu)\,. \tag{3.35}$$

In particular, $\operatorname{ex}\mathcal{M}_T(g) \neq \emptyset$.

The fact that the measure α_μ can be used follows from the fact that the σ-algebras $\mathcal{I}$ and $\mathcal{T}_{-\infty}$, which appeared naturally in the constructions of Q and Π respectively, are intimately related. This will be the content of Lemma 3.8 below.

Proof of Theorem 3.5: Let $\mu \in \mathcal{M}_T(g)$. Since μ is invariant, we can use Theorem 3.3 to decompose μ into its ergodic components:

$$\forall B \in \mathcal{F}, \quad \mu(B) = \int_{\mathcal{M}_T^{\mathrm{erg}}} \nu(B)\alpha_\mu(d\nu)\,. \tag{3.36}$$

We will see in Lemma 3.7 that α_μ concentrates on $\mathcal{M}(g)$: $\alpha_\mu(\mathcal{M}(g)) = 1$. Therefore,

$$\forall B \in \mathcal{F}, \quad \mu(B) = \int_{\mathcal{M}_T^{\mathrm{erg}}\cap\mathcal{M}(g)} \nu(B)\alpha_\mu(d\nu)\,, \tag{3.37}$$

But $\mathcal{M}_T^{\mathrm{erg}}\cap\mathcal{M}(g) = \operatorname{ex}\mathcal{M}_T(g)$, as was shown in [10]. This gives (3.35). $\square$

Lemma 3.7. *Let $\mu \in \mathcal{M}_T(g)$, with g regular. Then $\alpha_\mu(\mathcal{M}(g)) = 1$.*

Proof. Remember that $\alpha_\mu(M) = \mu(Q^{\cdot} \in M)$ for all $M \in e(\mathcal{M}_T^{\mathrm{erg}})$, where $Q^{\cdot}$ is the superkernel of Section 3.4. We compute, for Q^ω-almost all σ,

$$Q^\omega([+]_{k+1}|\mathcal{F}_{(-\infty,k]})(\sigma) = \lim_{l\to\infty} \frac{Q^\omega([+]_{k+1}[\sigma]_{-l}^k)}{Q^\omega([\sigma]_{-l}^k)}$$

Since $\mu \in \mathcal{M}_T$,

$$\frac{Q^\omega([+]_{k+1}[\sigma]_{-l}^k)}{Q^\omega([\sigma]_{-l}^k)} = \frac{\mu([+]_{k+1}[\sigma]_{-l}^k|\mathcal{I})(\omega)}{\mu([\sigma]_{-l}^k|\mathcal{I})(\omega)} \quad \mu\text{-a.s.}$$

The following can be found in [10], [12].

Lemma 3.8. *Let $\mu \in \mathcal{M}_T$. Then $\mathcal{I} \subset \mathcal{T}_{-\infty}$ μ-a.s. That is there exists, for all $A \in \mathcal{I}$, a set $B \in \mathcal{T}_{-\infty}$ such that $\mu(A\triangle B) = 0$.*

Therefore, we can write

$$\mu([+]_{k+1}[\sigma]_{-l}^{k}|\mathcal{I}) = \mu\big(\mu([+]_{k+1}[\sigma]_{-l}^{k}|\mathcal{T}_{-\infty})\big|\mathcal{I}\big) \quad \mu\text{-a.s.}$$

Since $\mu \in \mathcal{M}(g)$, we can describe $\mu([+]_{k+1}[\sigma]_{-l}^{k}|\mathcal{T}_{-\infty})$ with the superkernel Π:

$$\begin{aligned}\mu([+]_{k+1}[\sigma]_{-l}^{k}|\mathcal{T}_{-\infty}) &= \Pi^{\cdot}([+]_{k+1}[\sigma]_{-l}^{k}) \quad \mu\text{-a.s.}\\ &= \lim_{n\to\infty} \Pi^{\cdot}_{-n}([+]_{k+1}[\sigma]_{-l}^{k}) \quad \mu\text{-a.s.}\end{aligned} \tag{3.38}$$

But, as we already saw in (3.34),

$$\Pi^{\cdot}_{-n}([+]_{k+1}[\sigma]_{-l}^{k}) = \big[g(\sigma_{-\infty}^{k}) + O(\mathrm{var}_{k+l}(g))\big]\Pi^{\cdot}_{-n}([\sigma]_{-l}^{k})\,.$$

Using the same identity as above in the other way, $\lim_{n\to\infty}\Pi^{\cdot}_{-n}([\sigma]_{-l}^{k}) = \mu([\sigma]_{-l}^{k}|\mathcal{T}_{-\infty})$ μ-a.s., we get

$$Q^{\omega}([+]_{k+1}|\mathcal{F}_{(-\infty,k]})(\sigma) = g([\sigma]_{-\infty}^{k})\,,$$

which proves that $Q^{\omega} \in \mathcal{M}(g)$. □

4. Uniqueness: the Johansson-Öberg Criterium

In this section, we consider a kind of condition that g must satisfy in order to guarantee $|\mathcal{M}_T(g)| = 1$. We have seen, in Section 2, how the continuity of g, i.e. $\mathrm{var}_k(g) \to 0$, implies existence of measures specified by g. Since the dependence on the past is weaker when $\mathrm{var}_k(g)$ converges to 0 faster, a natural problem is to understand if uniqueness of the invariant measure can be obtained by imposing some condition on the speed at which $\mathrm{var}_k(g)$ goes to zero.

The uniqueness problem was considered by Doeblin and Fortet in their pioneering paper [7] (see also [14]), via a coupling argument, under the condition that $(\mathrm{var}_k(g))_{k\geq 1} \in \ell^1$, i.e.

$$\sum_{k\geq 1} \mathrm{var}_k(g) < \infty\,. \tag{4.1}$$

In Example 1.1, the bound (2.1) shows that

$$\sum_{k\geq 1} \mathrm{var}_k(g) \leq \sum_{k\geq 1} Q(R > k) = \mathsf{E}_Q(R)\,.$$

Therefore, [7] gives uniqueness when $\mathsf{E}_Q(R) < \infty$, and $\mathsf{E}_Q(R) = +\infty$ is a necessary ingredient for non-uniqueness.

Other significant results on uniqueness were obtained by Bowen [4], Walters [19], Berbee [1] and Hulse [13]. More recently, Johansson and Öberg [15] gave the following criterium:

Theorem 4.1 ([15]). *Assume g is regular and $(\mathrm{var}_k(g))_{k\geq 1} \in \ell^2$, i.e.*

$$\sum_{k\geq 1} \mathrm{var}_k(g)^2 < \infty \,. \tag{4.2}$$

Then there exists a unique invariant measure specified by g: $|\mathcal{M}_T(g)| = 1$.
Since $\ell^1 \subset \ell^2$, this result is a significant improvement of [7]. It was shown recently by Berger, Hoffman and Sidoravicius [2] that when regarding summability of variations, Theorem 4.1 is optimal in the following sense: for any $\epsilon > 0$, there exists a g-function with $(\mathrm{var}_k(g))_{k\geq 1} \in \ell^{2+\epsilon}$ which has at least two invariant measures specified by g [2].

Consider the g-function of Example 1.2. Since

$$\sum_{n>k} J_n \sim \frac{1}{k^\gamma}\,,$$

the upper bound (2.1) and Theorem 4.1 show that uniqueness is guaranteed when $\gamma > \frac{1}{2}$. This might seem surprising since the one-dimensional Ising model with long range ferromagnetic interactions given by a sequence $(J_n)_{n\geq 1}$ with the asymptotic behaviour as in (1.5), exhibits a phase transition at low temperature for all values $0 < \gamma < 1$ [9]. As a matter of fact, it was shown recently by Fernández and Maillard [11], using [9], that if invariance is dropped, then the Johansson-Öberg Criterium does not hold.

Theorem 4.1 is a corollary of

Proposition 4.1. *Assume g is uniformly non-null and $(\mathrm{var}_k(g))_{k\geq 1} \in \ell^2$. Then any two measures $\mu, \nu \in \mathcal{M}_T(g)$ are absolutely continuous with respect to the other: $\mu \ll \nu$ and $\nu \ll \mu$.*

Proof of Theorem 4.1: By the decomposition Theorem 3.5 for $\mathcal{M}_T(g)$, we know that the extreme elements of $\mathcal{M}_T(g)$, which are ergodic, determine $\mathcal{M}_T(g)$ completely. But if μ, ν are two distinct ergodic measures, then they are singular [3]. By Proposition 4.1, this is impossible. Therefore there can exist at most one ergodic measure specified by g, proving the theorem. □

Proof of Proposition 4.1: Since (4.2) implies continuity of g, there exists at least one invariant measure specified by g (Theorem 2.1). Assume $\mu, \nu \in \mathcal{M}_T(g)$. To show that $\mu \ll \nu$ ($\nu \ll \mu$ is obtained in the same way), the idea is to compare these measures on cylinders with large bases, by considering the following random variables:

$$M_n(\omega) := \frac{\mu([\omega]_1^n)}{\nu([\omega]_1^n)}\,.$$

[2] Observe, nevertheless, that the alphabet considered in [2] contains more than two letters.

The variables M_n are often called likelihood ratios; they are well defined by Lemma 2.1. The starting point is the content of the following lemma.

Lemma 4.1. *Let $A \in \mathcal{F}_{[1,m]}$. Then, for all $n \geq m$*

$$\mu(A) = \int_A M_n \, d\nu \,. \tag{4.3}$$

In particular, the sequence $(M_n)_{n\geq 1}$ is a martingale with respect to the filtration $(\mathcal{F}_{[1,n]})_{n\geq 1}$ and to the measure ν.

Proof. Clearly, M_n is $\mathcal{F}_{[1,n]}$-measurable. If (4.3) holds, then in particular

$$\int_A M_n \, d\nu = \int_A M_m \, d\nu \,, \tag{4.4}$$

which implies that $(M_n)_{n\geq 1}$ is a martingale with respect to $(\mathcal{F}_{[1,n]})_{n\geq 1}$ and to ν. To show (4.3), we decompose A into thin cylinders $[\omega_0]_1^m$ with $\omega_0 \in \{\pm 1\}^{[1,m]}$, and for each ω_0, we resum over all configurations ω_1 in $\{\pm 1\}^{[m+1,n]}$:

$$\begin{aligned}
\int_{[\omega_0]_1^m} M_n \, d\nu = \sum_{\omega_1} \int_{[\omega_0\omega_1]_1^n} M_n \, d\nu &= \sum_{\omega_1} \frac{\mu([\omega_0\omega_1]_1^n)}{\nu([\omega_0\omega_1]_1^n)} \int_{[\omega_0\omega_1]_1^n} d\nu \\
&= \sum_{\omega_1} \mu([\omega_0\omega_1]_1^n) \\
&= \mu([\omega_0]_1^m) \,,
\end{aligned}$$

which yields

$$\int_A M_n \, d\nu = \sum_{\omega_0} \int_{[\omega_0]_1^m} M_n \, d\nu = \sum_{\omega_0} \mu([\omega_0]_1^m) = \mu(A) \,,$$

proving (4.3). □

The identity (4.3) is interesting for the following reason: M_n is a candidate for the construction of a density of μ with respect to ν. The strategy of the proof is thus the following. We shall first show how the Johansson-Öberg Condition (4.2) implies that the martingale $(M_n)_{n\geq 1}$ is uniformly integrable, leading to the ν-almost sure existence of the limit $M_n \to M_\infty$. This will imply, by taking $n \to \infty$ in (4.3), that

$$\mu(A) = \int_A M_\infty \, d\nu \tag{4.5}$$

for all cylinder A whose base lies in the half space $[1, +\infty)$. In order to extend (4.5) to any cylinder (i.e. with base in $\mathbb{Z}$), we will average translates of M_∞ by constructing

$$\overline{M}_\infty := \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} M_\infty \circ T^k \,, \tag{4.6}$$

whose existence is guaranteed by the Ergodic Theorem. $\overline{M}_\infty$ will then be shown to be the Radon-Nikodým derivative of μ with respect to ν, concluding the proof.

Let us show that $(M_n)_{n\geq 1}$ is uniformly integrable with respect to ν, which means (remember that $M_n > 0$)

$$\lim_{K\to\infty} \sup_{n\geq 1} \int_{M_n \geq K} M_n \, d\nu = 0\,.$$

By taking $A = \{M_n \geq K\}$ in (4.3), we have

$$\int_{M_n\geq K} M_n \, d\nu = \mu(M_n \geq K)\,.$$

Uniform integrability with respect to ν is thus equivalent to tightness with respect to μ:

$$\lim_{K\to\infty} \sup_{n\geq 1} \mu(M_n \geq K) = 0\,. \tag{4.7}$$

Lemma 4.2. *If g satisfies the Johansson-Öberg Condition* (4.2)*, then $(M_n)_{n\geq 1}$ is tight with respect to μ.*

Proof. Since $M_n > 0$, it suffices to show that

$$\lim_{K\to\infty} \sup_{n\geq 1} \mu(\log M_n \geq K) = 0\,. \tag{4.8}$$

Consider the decomposition:

$$\mu([\omega]_1^n) = \prod_{k=1}^{n} \pi_k(\mu)\,,$$

where the $\pi_k(\mu) = \pi_k(\mu)(\omega)$ are defined by (use again Lemma 2.1)

$$\pi_k(\mu) := \mu([\omega]_k | [\omega]_1^{k-1})\,, \quad k = 2, \ldots, n\,,$$

and $\pi_1(\mu) := \mu([\omega]_1)$. We define $\pi_k(\nu)$ similarly.

Lemma 4.3. *We have* $\inf_\omega \pi_k(\cdot) \geq \epsilon > 0$ *and*

$$\sup_\omega |\pi_k(\mu) - \pi_k(\nu)| \leq 2\,\mathrm{var}_k(g)\,.$$

Proof. We use the fact that μ and ν are specified by g, which is assumed to be regular. For example, in the case where $\omega_k = +1$, (2.5) gives

$$\begin{aligned}\pi_k(\mu) &= \frac{1}{\mu([\omega]_1^{k-1})} \int_{[\omega]_1^{k-1}} g([\omega]_1^{k-1}[\sigma]_{-\infty}^0)\, \mu(d\sigma) \\ &= g(\omega_{-\infty}^k) + O(\mathrm{var}_k(g))\,. \end{aligned} \tag{4.9}$$

Doing the same with $\pi_k(\nu)$, we obtain the result. □

Now, write

$$\log M_n = \sum_{k=1}^{n} \log \frac{\pi_k(\mu)}{\pi_k(\nu)} = \sum_{k=1}^{n} \log \Big(1 + \frac{\pi_k(\mu) - \pi_k(\nu)}{\pi_k(\nu)}\Big)$$

$$\leq \sum_{k=1}^{n} \frac{\pi_k(\mu) - \pi_k(\nu)}{\pi_k(\nu)}\,, \tag{4.10}$$

since $\log(1+x) \leq x$ for all $x > -1$.

Remark 4.1. Using Lemma 4.3 in (4.10), we get

$$\log M_n \leq \frac{2}{\epsilon} \sum_{k=1}^{n} \mathrm{var}_k(g) \leq \frac{2}{\epsilon} \sum_{k=1}^{\infty} \mathrm{var}_k(g)$$

Therefore, if one assumes that the variation of g is in ℓ^1 (the condition of Doeblin and Fortet), one obtains immediately tightness of $(\log M_n)_{n\geq 1}$.

The key to Johansson and Öberg's generalization is to consider the following identity:

$$\sum_{k=1}^{n} \frac{\pi_k(\mu) - \pi_k(\nu)}{\pi_k(\nu)} = \sum_{k=1}^{n} \frac{\pi_k(\mu) - \pi_k(\nu)}{\pi_k(\mu)} + \sum_{k=1}^{n} \frac{(\pi_k(\mu) - \pi_k(\nu))^2}{\pi_k(\mu)\pi_k(\nu)}\,.$$

The second term is bounded using (4.2):

$$\sum_{k=1}^{n} \frac{(\pi_k(\mu) - \pi_k(\nu))^2}{\pi_k(\mu)\pi_k(\nu)} \leq \frac{4}{\epsilon^2} \sum_{k=1}^{n} \mathrm{var}_k(g)^2 \leq \frac{4}{\epsilon^2} \sum_{k\geq 1} \mathrm{var}_k(g)^2 < \infty\,.$$

The point is that in the first term, which we denote by

$$Z_n := \sum_{k=1}^{n} \frac{\pi_k(\mu) - \pi_k(\nu)}{\pi_k(\mu)}\,,$$

the measure appearing in the denominator is μ in place of ν. This is convenient, since we need to show tightness *with respect to* μ.

Lemma 4.4. *$(Z_n)_{n\geq 1}$ is a martingale with respect to the filtration $(\mathcal{F}_{[1,n]})_{n\geq 1}$ and to μ.*

Proof. We show that

$$E_\mu\big((Z_k - Z_{k-1})|\mathcal{F}_{[1,k-1]}\big) = 0 \tag{4.11}$$

for all $k \geq 1$. Denote by $[\omega^j]_1^{k-1}$, $j = 1, 2, \dots, 2^{k-1}$ the atoms generating $\mathcal{F}_{[1,k-1]}$. On $[\omega^j]_1^{k-1}$, we have

$$E_\mu\big((Z_k - Z_{k-1})|\mathcal{F}_{[1,k-1]}\big) = \frac{1}{\mu([\omega^j]_1^{k-1})} \int_{[\omega^j]_1^{k-1}} \frac{\pi_k(\mu) - \pi_k(\nu)}{\pi_k(\mu)}\, d\mu$$

This last integral can be computed explicitely:

$$\begin{aligned}
&\int_{[\omega^j]_1^{k-1}} \frac{\pi_k(\mu)-\pi_k(\nu)}{\pi_k(\mu)}\, d\mu \\
&\qquad = \sum_{\alpha=\pm} \frac{\mu([\alpha]_k|[\omega^j]_1^{k-1})-\nu([\alpha]_k|[\omega^j]_1^{k-1})}{\mu([\alpha]_k|[\omega^j]_1^{k-1})}\mu([\alpha]_k[\omega^j]_1^{k-1}) \\
&\qquad = \mu([\omega^j]_1^{k-1})\sum_{\alpha=\pm}\big(\mu([\alpha]_k|[\omega^j]_1^{k-1})-\nu([\alpha]_k|[\omega^j]_1^{k-1})\big) = 0\,.
\end{aligned}$$

Notice the cancellations of the terms involving μ, which occured precisely because Z_n was defined with μ in the denominator. $\square$

Now, since $\log M_n \le Z_n + C$, $C<\infty$, tightness of $(Z_n)_{n\ge1}$ (with respect to μ) implies tightness of $(\log M_n)_{n\ge1}$. To show that $(Z_n)_{n\ge1}$ is tight, it is sufficient to show that it is bounded in $L^2(\Omega,\mathcal{F},\mu)$, i.e. $\sup_{n\ge1}\|Z_n\|_2<\infty$. Indeed, for any $K>0$,

$$\mu(Z_n\ge K) = \frac{1}{K^2}\int_{Z_n\ge K} K^2\, d\mu \le \frac{1}{K^2}\int_{Z_n\ge K} Z_n^2\, d\mu \le \frac{1}{K^2}\|Z_n\|_2^2\,.$$

Now, setting $Z_0 := 0$ and writing $Z_n = \sum_{k=1}^n (Z_k - Z_{k-1})$, we get

$$\|Z_n\|_2^2 = \sum_{k=1}^n E_\mu[(Z_k-Z_{k-1})^2] + 2\sum_{1\le j<k\le n} E_\mu[(Z_j-Z_{j-1})(Z_k-Z_{k-1})]\,.$$

By the definition of Z_k and Lemma 4.3,

$$\begin{aligned}
\sum_{k=1}^n E_\mu[(Z_k-Z_{k-1})^2] &= \sum_{k=1}^n E_\mu\Big[\Big(\frac{\pi_k(\mu)-\pi_k(\nu)}{\pi_k(\mu)}\Big)^2\Big] \\
&\le \frac{4}{\epsilon^2}\sum_{k=1}^n \mathrm{var}_k(g)^2 \le \frac{4}{\epsilon^2}\sum_{k\ge1}\mathrm{var}_k(g)^2 < \infty\,.
\end{aligned}$$

For each pair $j<k$, $Z_j - Z_{j-1}$ is $\mathcal{F}_{[1,k-1]}$-measurable since $\mathcal{F}_{[1,j-1]} \subset \mathcal{F}_{[1,j]} \subset \mathcal{F}_{[1,k-1]}$, and so

$$\begin{aligned}
&E_\mu[(Z_j-Z_{j-1})(Z_k-Z_{k-1})] \\
&\qquad = E_\mu\big[E_\mu\big((Z_j-Z_{j-1})(Z_k-Z_{k-1})|\mathcal{F}_{[1,k-1]}\big)\,\big] \\
&\qquad = E_\mu\big[(Z_j-Z_{j-1})E_\mu\big((Z_k-Z_{k-1})|\mathcal{F}_{[1,k-1]}\big)\big] = 0
\end{aligned}$$

by (4.11). We have thus shown that $\sup_{n\ge1}\|Z_n\|_2<\infty$, which in turn implies that $(M_n)_{n\ge1}$ is tight, and proves Lemma 4.2. $\square$

Since $(M_n)_{n\ge1}$ is tight with respect to μ, it is uniformly integrable with respect to ν, as we have seen. Therefore (see [20]), there exists an $\mathcal{F}_{[1,+\infty)}$-measurable random variable $M_\infty\ge0$ such that $M_n\to M_\infty$ in $L^1(\Omega,\mathcal{F},\nu)$.

Taking the limit $n \to \infty$ in (4.3), we get

$$\mu(A) = \int_A M_\infty \, d\nu \tag{4.12}$$

for all cylinder whose base lies in $[1, +\infty)$, i.e. for all $A \in \bigcup_{n\geq 1} \mathcal{F}_{[1,n]}$. To obtain a similar representation for a generic cylinder $A \in \mathcal{C}$, we use the Ergodic Theorem: since ν is invariant,

$$\frac{1}{n} \sum_{k=0}^{n-1} M_\infty \circ T^k \longrightarrow \overline{M}_\infty \qquad \text{in } L^1(\Omega, \mathcal{F}, \nu)\,,$$

$\overline{M}_\infty$ is a version of the conditional expectation of M_∞ with respect to $\mathcal{I}$. Now for any cylinder $A \in \mathcal{C}$, we compute

$$\int_A \overline{M}_\infty \, d\nu = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_A M_\infty \circ T^k \, d\nu\,. \tag{4.13}$$

Assume k is large enough so that the base of $T^{-k}A$ lies in $[1, +\infty)$. By a change of variable, (4.12) and the invariance of μ,

$$\int_A M_\infty \circ T^k \, d\nu = \int_{T^{-k}A} M_\infty \, d\nu = \mu(T^{-k}A) = \mu(A)$$

Since this can be used for essentially all the terms in the sums appearing in (4.13), we get

$$\mu(A) = \int_A \overline{M}_\infty \, d\nu\,.$$

But the cylinders generate $\mathcal{F}$, and so this shows that μ is absolutely continuous with respect to ν, with density $\overline{M}_\infty$, which finishes the proof of Proposition 4.1. $\square$

5. Non-Uniqueness

The Johannson-Öberg Criterium described in Section 4 shows that when the variation of g decreases fast enough, i.e. at least when $(\mathrm{var}_k(g))_{k\geq 1} \in \ell^2$, then there exists a unique invariant measure specified by g. It had actually been an open problem for some time to decide whether continuity of g was enough to guarantee this uniqueness, without requiring anything on the speed of convergence of the variation to zero. For the first time in [5], Bramson and Kalikow showed the existence of regular g-functions for which there exist at least two invariant measures, i.e. $|\mathcal{M}_T(g)| > 1$. We will describe their result in Section 5.1 below. More recently, Berger, Hoffman and Sidoravicius [2] constructed another example of non-uniqueness, showing the optimality of the Johannson-Öberg Criterium. In [11], Fernández and Maillard exhibited g-functions which satisfy the Johannson-Öberg Criterium, but for which there exist at least two measures specified by g; such measures being necessarily non-invariant.

5.1. **The Bramson-Kalikow Mechanism.** Consider the type of g-function described in Example 1.1 of the Introduction:

$$g(\omega_{-\infty}^{k}) := \sum_{n\geq 1} p_n \varphi\Big(\frac{1}{n}\sum_{j=0}^{n-1}\omega_{k-j}\Big)\,, \tag{5.1}$$

where φ is a non-decreasing and bounded away from zero and one, in order to guarantee regularity of g. Bramson and Kalikow considered

$$\varphi(x) = \begin{cases} 1-\epsilon & \text{if } x \geq 0\,,\\ \epsilon & \text{if } x < 0\,,\end{cases} \tag{5.2}$$

where $\epsilon > 0$. Functions of the type (5.1) are clearly attractive, and so we can construct two measures $\mu_*^{\pm}$, prepared respectively with the boundary condition $\sigma \equiv +$ and $\sigma \equiv -$, as in (2.11).

Theorem 5.1. [5] *Let φ be as in* (5.2) *with $0 < \epsilon < \frac{1}{4}$. There exists a sequence $(p_n)_{n\geq 1}$ (depending on ϵ) such that $\mu_*^+ \neq \mu_*^-$.*

The mechanism leading to non-uniqueness invented by Bramson and Kalikow relies on the use of a sequence $(p_n)_{n\geq 1}$ with a highly lacunary structure, i.e. such that

$$g(\omega_{-\infty}^{k}) := \sum_{k\geq 1} p_{m_k} \varphi\Big(\frac{1}{m_k}\sum_{j=0}^{m_k-1}\omega_{k-j}\Big)\,, \tag{5.3}$$

where $(m_k)_{k\geq 1}$ is a rapidly increasing sequence of integers. Let us sketch the argument showing why non-uniqueness can be obtained under this assumption.

Appendix A. Proofs of auxiliary results

Proof of Lemma 2.1: We proceed by induction on the size of the base. For a thin cylinder whose base is of size one, say $[\sigma]_k$ with $\sigma_k = +1$, one has by (2.5):

$$\mu([\sigma]_k) = \int g(\omega_{-\infty}^{k-1})\mu(d\omega) \geq \epsilon\,.$$

If one assumes that $\mu([\sigma]_a^b) \geq \epsilon^{b-a+1}$, then for the cylinder $[\sigma]_a^{b+1}$ one gets easily, in case $\sigma_{b+1} = +1$,

$$\begin{aligned}\mu([\sigma]_a^{b+1}) = \mu([\sigma]_{b+1}\cap[\sigma]_a^b) &= \int_{[\sigma]_a^b}\mu([+]_{b+1}|\mathcal{F}_{(-\infty,b]})d\mu\\ &\geq \epsilon\mu([\sigma]_a^b) \geq \epsilon\epsilon^{b-a+1} = \epsilon^{b+1-a+1}\,,\end{aligned}$$

which proves the lemma. □

Proof of Lemma 2.2: Define $\mathcal{E} = \{B \in \mathcal{F} : \nu(T^{-1}B) = \nu(B)\}$. By hypothesis, $\mathcal{C} \subset \mathcal{E} \subset \mathcal{F}$. Moreover, if $B_n \in \mathcal{E}$, $B_n \nearrow B$, then

$$\nu(T^{-1}B) = \nu\big(\bigcup_{n\geq 1} T^{-1}B_n\big) = \lim_{n\to\infty} \nu(T^{-1}B_n) = \lim_{n\to\infty} \nu(B_n) = \nu(B)\,,$$

i.e. $B \in \mathcal{E}$. The same can be done for decreasing sequences of cylinders. By the Monotone Class Theorem, $\mathcal{E} = \mathcal{F}$. $\square$

Proof of Lemma 2.3: For $\nu \circ T^{-1}$-almost all ω,

$$\nu \circ T^{-1}([+]_{k+1}|\mathcal{F}_{(-\infty,k]}) = \lim_{l\to\infty} \frac{\nu \circ T^{-1}([+]_{k+1}[\omega]^k_{-l})}{\nu \circ T^{-1}([\omega]^k_{-l})}\,.$$

Now $\nu \circ T^{-1}([+]_{k+1}[\omega]^k_{-l}) = \nu([+]_{k+2}[T^{-1}\omega]^{k+1}_{-l+1})$, which equals

$$\begin{aligned}\nu([+]_{k+2}[T^{-1}\omega]^{k+1}_{-l+1}) &= \int_{[T^{-1}\omega]^{k+1}_{-l+1}} g((T^{-1}\omega)^{k+1}_{-l+1}\sigma^{-l}_{-\infty})\nu(d\sigma)\\ &= \big[g((T^{-1}\omega)^{k+1}_{-\infty}) + O(\mathrm{var}_{k+l}(g))\big]\nu([T^{-1}\omega]^{k+1}_{-l+1})\\ &= \big[g(\omega^k_{-\infty}) + O(\mathrm{var}_{k+l}(g))\big]\nu \circ T^{-1}([\omega]^k_{-l})\end{aligned}$$

$\square$

Proof of Lemma 2.4: Let $i_1, i_2, \dots$ be an enumeration of $\mathbb{Z}$ and define, for all $\omega, \omega' \in \Omega$,

$$d(\omega,\omega') := \sum_{n\geq 1} 2^{-n} d_0(\omega_{i_n}, \omega'_{i_n})\,,$$

where $d_0(a,a') = 0$ if $a = a'$, 1 otherwise. This metric turns Ω into a compact metric space. Observe that in this topology, cylinders are at the same time open and close.
Assume $B_n \neq \emptyset$ for all n. Then $\bigcap_{n\in\mathcal{N}} B_n \neq \emptyset$ for all finite non-empty set $\mathcal{N} \subset \mathbb{N}$. This implies that $\{B_n\}_{n\geq 1}$ has the finite intersection property. Since cylinders are closed and Ω is compact, this implies that $\bigcap_{n\geq 1} B_n \neq \emptyset$, a contradiction. $\square$

Proof of Lemma 3.1: Assume μ is trivial on $\mathcal{A}$. Fix some $B \in \mathcal{F}$. Then for all $A \in \mathcal{A}$,

$$\int_A \mu(B)d\mu = \mu(B)\mu(A) = \mu(B\cap A) = \int_A \mu(B|\mathcal{A})d\mu = \mu(B\cap A)\,,$$

where the second inequality follows from the triviality of μ. Conversely, assuming that for all $B \in \mathcal{F}$, $\mu(B|\mathcal{A}) = \mu(B)$ μ-a.s., we get, for $B = A \in \mathcal{A}$,

$$\mu(A) = \mu(A\cap A) = \int_A \mu(A|\mathcal{A})d\mu = \int_A \mu(A)d\mu = \mu(A)^2\,,$$

and therefore $\mu(A) = 0$ or 1. $\square$

Proof of Lemma 3.5: W verify for example that

$$\mu([+]_1^2|\mathcal{F}_{(-\infty,0]})(\omega) = g(+_1\omega_{-\infty}^0)g(\omega_{-\infty}^0) \tag{A.1}$$

for μ-almost all ω. As usual, μ-a.s.

$$\mu([+]_1^2|\mathcal{F}_{(-\infty,0]})(\omega) = \lim_{l\to\infty} \frac{\mu([+]_2[+]_1[\omega]_{-l}^0)}{\mu([\omega]_{-l}^0)} .$$

But, since $\mu \in \mathcal{M}(g)$,

$$\begin{aligned}\mu([+]_2[+]_1[\omega]_{-l}^0) &= \int_{[+]_1[\omega]_{-l}^0} g(+_1\omega_{-l}^0\sigma_{-\infty}^{-l-1})\mu(d\sigma)\\ &= \big[g(+_1[\omega]_{-\infty}^0) + O(\mathrm{var}_{l+1}(g))\big]\mu([+]_1[\omega]_{-l}^0) .\end{aligned}$$

In the same way, one gets

$$\mu([+]_1[\omega]_{-l}^0) = \big[g(\omega_{-\infty}^0) + O(\mathrm{var}_l(g))\big]\mu([\omega]_{-l}^0) .$$

This shows (A.1). The general case is done in the same way. □

Proof of Lemma 3.6: For simplicity, consider a cylinder $[+]_a^b$. Then since $g \ge \epsilon$,

$$\begin{aligned}\Pi_{-n}^{\omega}([+]_a^b) &= \sum_{[\eta]_{-n+1}^{a-1}} \gamma_{-n}^{\omega}([+]_a^b[\eta]_{-n+1}^{a-1})\\ &\ge \epsilon^{b-a+1} \sum_{[\eta]_{-n+1}^{a-1}} \gamma_{-n}^{\omega}([\eta]_{-n+1}^{a-1}) \equiv \epsilon^{b-a+1} ,\end{aligned}$$

uniformly in n. □

Appendix B. Dynkin Systems

Let Ω be any non-empty set. We denote by 2^Ω the family of all subsets of Ω, including the emptyset.

Definition B.1. *A collection $\mathcal{D} \subset 2^\Omega$ is called a* Dynkin System *(or simply* D-system*) if the following conditions hold:*

(1) $\Omega \in \mathcal{D}$.
(2) *If $A, B \in \mathcal{D}$, $A \subset B$, then $B\backslash A \in \mathcal{D}$.*
(3) *If $A_n \in \mathcal{D}$ for all $n \ge 1$, $A_n \nearrow A$, then $A \in \mathcal{D}$*

Observe that D-systems are stable by complementation since $A \in \mathcal{D}$ implies $A^c = \Omega\backslash A \in \mathcal{D}$. Since in general $B\backslash A = B \cap A^c$, σ-algebras are D-systems. The only property which D-systems might not have in comparison to σ-algebras is stability under intersections.

Lemma B.1. *A collection $\mathcal{F} \subset 2^\Omega$ is a σ-algebra if and only if it is a D-system stable under intersection.*

Proof. The "only if" part is trivial. Then, assume $\mathcal{F}$ is a D-system stable under intersection. Let $A, B \in \mathcal{F}$. We have $A \cup B = (A^c \cap B^c)^c = \Omega \backslash (A^c \cap B^c) \in \mathcal{F}$. Let $A_n \in \mathcal{F}$, $B_n := \bigcup_{k=1}^n A_k$. Since $B_n \in \mathcal{F}$ and $B_n \nearrow \bigcup_{n\geq 1} B_n$, we have that $\bigcup_{n\geq 1} B_n \in \mathcal{F}$. This shows that $\mathcal{F}$ is a σ-algebra. □

As can be easily verified, the intersection of an arbitrary family of D-systems is a D-system. Therefore, given any collection $\mathcal{C} \subset 2^\Omega$, one can define the smallest D-system containing $\mathcal{C}$, called the D-system generated by $\mathcal{C}$, denoted $\mathcal{D}(\mathcal{C})$. In practice, it is interesting to compare the D-system $\mathcal{D}(\mathcal{C})$ with the σ-algebra $\sigma(\mathcal{C})$. One clearly has $\mathcal{D}(\mathcal{C}) \subset \sigma(\mathcal{C})$. The reverse inclusion is obtained by imposing that $\mathcal{C}$ be stable under intersections, as seen in the following lemma.

Theorem B.1. *If $\mathcal{C} \subset 2^\Omega$ is stable under intersection, then $\mathcal{D}(\mathcal{C}) = \sigma(\mathcal{C})$.*

Proof. To simplify the notations, denote $\mathcal{D}(\mathcal{C})$ by $\mathcal{D}$ and $\sigma(\mathcal{C})$ by $\mathcal{F}$. We already saw that $\mathcal{D} \subset \mathcal{F}$. To show that $\mathcal{D} \supset \mathcal{F}$, it suffices to verify that $\mathcal{D}$ is a σ-algebra. By Lemma B.1, it suffices to verify that $\mathcal{D}$ is stable under intersection.
Define $\mathcal{D}_1 := \{B \in \mathcal{D} : B \cap C \in \mathcal{D} \forall C \in \mathcal{C}\}$. We verify that $\mathcal{D}_1 = \mathcal{D}$. By definition, $\mathcal{D}_1 \subset \mathcal{D}$. To verify that $\mathcal{D}_1 \supset \mathcal{D}$, it suffices to see that $\mathcal{D}_1$ is a D-system containing $\mathcal{C}$. Now $\mathcal{D}_1 \supset \mathcal{C}$ follows from the fact that $\mathcal{C}$ is closed under intersection. This also implies that $\Omega \in \mathcal{D}_1$. Let $B_1, B_2 \in \mathcal{D}_1$, $B_1 \subset B_2$, $C \in \mathcal{C}$. Then

$$(B_2 \backslash B_1) \cap C = B_2 \cap C \cap (B_1^c \cup C^c) = (B_2 \cap C) \backslash (B_1 \cap C) \in \mathcal{D}$$

Then, if $B_n \in \mathcal{D}_1$, $B_n \nearrow B$, then $B \cap C = \bigcup_n (B_n \cap C) \in \mathcal{D}$, logo $B \in \mathcal{D}_1$. This proves that $\mathcal{D}$ is a D-system.
Define $\mathcal{D}_2 := \{A \in \mathcal{D} : A \cap B \in \mathcal{D} \forall B \in \mathcal{D}\}$. We verify that $\mathcal{D}_2 = \mathcal{D}$, which will show that $\mathcal{D}$ is stable under intersection. By the first step, $\mathcal{D}_2$ contains $\mathcal{C}$. As before, one can show that $\mathcal{D}_2 = \mathcal{D}$. This shows that $\mathcal{D}$ is stable under intersection, and finishes the proof of the theorem. □

References

[1] H. Berbee. Chains with infinite connections: Uniqueness and markov representation. *Prob. Theory and Rel. Fields*, 76:243–253, 1987.
[2] N. Berger, C. Hoffman, and V. Sidoravicius. Nonuniqueness for specifications in $\ell^{2+\epsilon}$. *preprint*, 2005.
[3] P. Billingsley. *Ergodic Theory and Information*. Wiley, New York, 1965.
[4] R. Bowen. Equilibrium states and the ergodic theory of Anosov diffeomorphisms. *Lecture Notes in Mathematics*, 470, 1975.
[5] M. Bramson and S. Kalikow. Non-uniqueness in g-functions. *Israel J. Math.*, 84:153–160, 1993.
[6] R.L. Dobrushin. The description of a random field by means of conditional probabilities and conditions of its regularity. *Theor. Prob. Appl.*, 13:197–224, 1968.

[7] W. Doeblin and R. Fortet. Sur les chaînes à liaisons complètes. *Bull. Soc. Math. France*, 65:132–148, 1937.
[8] E.B. Dynkin. Sufficient statistics and extreme points. *Ann. Prob.*, 6:705–730, 1978.
[9] Freeman J. Dyson. Existence of a phase-transition in a one-dimensional Ising ferromagnet. *Commun. Math. Phys.*, 12:91–107, 1969.
[10] R. Fernández and G. Maillard. Chains with complete connections : General theory, uniqueness, loss of memory and mixing properties. *Journ. Stat. Phys.*, 118:555–588, 2005.
[11] R. Fernández and G. Maillard. Uniqueness and non-uniqueness of chains on half-lines. *preprint*, 2007.
[12] H.-O. Georgii. *Gibbs Measures and Phase Transitions.* de Gruyter Studies in Mathematics, 1988.
[13] P. Hulse. Uniqueness and ergodic properties of attractive g-measures. *Ergod. Th. and Dynam. Sys*, 11:65–77, 1991.
[14] M. Iosifescu. A coupling method in the theory of dependence with complete connections according to Doeblin. *Rev. Roumaine Math. Pures Appl.*, 37:59–66, 1992.
[15] A. Johansson and A. Öberg. Square summability of variations of g-functions and uniqueness of g-measures. *Mathematical Research Letters*, 10:587–601, 2003.
[16] M. Keane. Strongly mixing g-measures. *Invent. Math.*, 16:309–324, 1972.
[17] C.J. Preston. Random fields. *Lecture Notes in Mathematics*, 465, 1975.
[18] H.L. Royden. *Real Analysis.* MacMillan Publishing Co, 1968.
[19] P. Walters. Ruelle's operator theorem and g-measures. *Trans. Amer. Math. Soc.*, 214:375–387, 1975.
[20] D. Williams. *Probability with Martingales.* Cambridge Mathematical Textbooks, 1991.

Departamento de Matemática, ICEx, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, C.P. 702, CEP 30123-970 Belo Horizonte, Brasil
E-mail address: sacha@mat.ufmg.br