Characterization of extreme Gibbs measures for a Categorical Approach to Statistical Mechanics

Grégoire Sergeant-Perthuis*[1]

[1]Laboratoire de Biologie Computationnelle et Quantitative, Sorbonne Université, Paris , France, , ORCiD: 0000-0002-2079-3410

February 22, 2024

Abstract

In recent work we proposed to follow a categorical approach to Statistical Mechanics. In this paper we continue in this direction and give a categorical formulation of the decomposition of Gibbs measures by extending the characterization of extreme Gibbs measures to the categorical setting.

Keywords: Statistical mechanics, category theory, Extreme Gibbs measures, Zero-one laws

MSC: 18D99,60F20,60A99

1 Introduction

Statistical Mechanics is the field interested in the statistical study of interacting particles ($X_i \in E_i, i \in \mathbb{N}$). A landmark rigorous formulation of Statistical Mechanics can be found in Georgii's *Gibbs Measures and Phase Transition* [1]. Such a framework is needed to define phases or, more generally, *Gibbs measures* of *statistical systems*. Phases only exist for infinitely many interacting particles and do not exist for finite-size systems. The main constructions of such a framework rely on the necessity to have a *universe*, denoted Ω, that encompasses all possible configurations of the system, i.e., all the possible joint configurations of the particles $\Omega := \prod_{i\in\mathbb{N}} E_i$. Not referring to a global space is problematic as exhibited by Giry's seminal work on *A Categorical Approach to Probability Theory* [2] and can be summarized by the fact that the Giry monad, $\mathbb{P}$, does not commute to limits. We are interested in describing statistical systems for which it is not possible to have

*gregoireserper@gmail.com

complete knowledge of the states of the particles at the same time; we want to forget about the 'Ω' and we want to study statistical systems 'locally'.

In [3, 4], we proposed to reformulate Statistical Mechanics and Gibbs measures introducing an appropriate category. Among others, in [5], we gave a characterization of independent variables in terms of projective objects in this category and an easy-to-verify condition that characterizes such objects (for the injective case see [6–9]). In [10], we proposed an Entropy functional for the categorical version of statistical systems. In this article, we show how the characterization of extreme Gibbs measures (Theorem 7.7 [1]), one of the steps for proving a zero-one law for the extreme Gibbs measures, transfers in the categorical setting. In the classical theory of rigorous statistical mechanics, the tail σ-algebra generates the observables for which a 'generalized' law of large numbers (zero-one law) holds. In this article, we give a candidate for such a σ-algebra in the categorical setting and show the associated extreme Gibbs measures decomposition. We will discuss in a follow-up paper the generalization of the zero-one law of specifications to $\mathscr{A}$-specifications.

2 $\mathscr{A}$-specifications and their Gibbs measures

In classical theory, a statistical system is defined by a specification (Definition 1.23 [1]). We proposed in [3,4] to reformulate such constructions as a couple of a presheaf and a functor over a poset, denoted as $\mathscr{A}$; we call such a couple an $\mathscr{A}$-specification. Let us recall their definition. Let us denote **Mes** the category that has as objects measurable spaces and as morphisms measurable maps, and **Kern** the category that has as objects measurable spaces and as morphisms Markov kernels.

Definition 2.1 ($\mathscr{A}$-specifications (Definition 7 [3])). Let $\mathscr{A}$ be a poset. An $\mathscr{A}$-specification is a couple (G, F) of a presheaf and a functor $G : \mathscr{A}^{op} \to$ **Mes** and $F : \mathscr{A} \to$ **Kern** such that for any $a, b \in \mathscr{A}$ with $b \leq a$,

$$G^a_b \circ F^b_a = \mathrm{id} \tag{1}$$

We denote $\alpha \circ \beta$ simply as $\alpha\beta$ in what follows.

Let us now recall the categorical generalization of the classical Gibbs measures (Definition 1.23 [1]) that we propose.

Definition 2.2 (Gibbs measure for $\mathscr{A}$-specifications Definition 8 [3]). Let $\gamma := (G, F)$ be an $\mathscr{A}$-specification,

$$\mathscr{G}(\gamma) := \{(p_a \in \mathbb{P}(G(a)), a \in \mathscr{A}) | \forall b \leq a, F^b_a p_b = p_a\} \tag{2}$$

with $\mathbb{P}(G(a))$ the (measurable) space of probability measures over $G(a)$.

3 Categorical Version of the Tail σ-Algebra

Measurable maps are particular Markov kernels. Let $F : E_b \to E_a$ be a Markov kernel; let us denote it also as $F(\omega_a|\omega_b)$, where $\omega_b \in E_b$ and $\omega_a \in E_a$. One associates to a Markov kernel $F : E_b \to E_a$ a linear map $\pi : L^\infty(E_a) \to L^\infty(E_b)$ defined as follows: for any $f \in L^\infty(E_a)$,

$$\forall \omega_b \in E_b, \quad \pi(f)(\omega_b) = \int f(\omega_a) F(d\omega_a|\omega_b) \tag{3}$$

This association is 'functorial', we may denote the underlying functor $L^\infty : \mathbf{Kern}^{op} \to \mathbf{Vect}$ which is presheaf from the category of Markov kernels to the category of vector spaces. It is the presheaf that associates spaces to their space of observables. Let us denote $L^\infty \circ G : \mathscr{A} \to \mathbf{Vect}$ as i and $L^\infty \circ F : \mathscr{A}^{op} \to \mathbf{Vect}$ as π. In these notation one has that for any $a, b \in \mathscr{A}$ such that $b \leq a$ then $\pi^a_b \circ i^b_a = \mathrm{id}$.

Let us recall the definition of the tail σ-algebra in the classical formulation. Let $\Omega = \prod_{i\in\mathbb{N}} E_i$, where E_i are measurable spaces; let us denote $\mathscr{F}_{\geq k}$ as the σ-algebra generated by the cylinders $\prod_{n\geq k} E_k$. The tail σ-algebra is defined as $\bigcap_{k\in\mathbb{N}} \mathscr{F}_{\geq k}$. A more general definition holds when I is any set and $\mathscr{F}_{\supseteq a}$ is indexed by a subset $a \subseteq I$ that is co-finite, i.e., which has a finite complement. For a functor from $\mathscr{A}$ to $\mathbf{Mes}$, let us denote $\sigma(G(a))$ as the sigma algebra of the measurable space $G(a)$, where $a \in \mathscr{A}$, and $\sigma(G)$ as the underlying functor defined as $\sigma(G)^b_a A_b := G^{a\,-1}_b A_b$, with $b \leq a$. We propose that one candidate that plays the role of the tail σ-algebra for a given specification $\gamma = (G, F)$ is $\lim \sigma(G)$ defined as,

$$\lim \sigma(G) := \{(A_a \in \sigma(G(a)), a \in \mathscr{A}) | \forall a, b \in \mathscr{A}, \quad A_a = G^{a\,-1}_b A_b\} \tag{4}$$

Let us denote $1_A : E \to \{0, 1\}$ the indicator function over the set A that sends $\omega \in A$ to 1 and $\omega \notin A$ to 0. Let us remark that $1_{A_b} \circ G^a_b = 1[G^a_b(\omega_a) \in A_b] = 1_{G^{a\,-1}_b A_b}$. Remark that $A \in \lim \sigma(G)$ is equivalent to $1_A \in \lim i$; in other words, $\lim \sigma(G)$ is the restriction of $\lim i$ to indicator functions of the form $1_{A_a}, a \in \mathscr{A}$.

Finally, we also need to recall that for any $f \in L^\infty(E)$ and $\mu \in \mathbb{P}(E)$, one can define a measure $f \cdot \mu$ as $f \cdot \mu(d\omega) = f(\omega) d\omega$.

The key proposition of this document is Proposition 3.1; the proof of that proposition is given for $G(a), a \in \mathscr{A}$ finite measurable sets. Therefore, we assume in what follows that the measurable sets are finite. However, there is no finiteness constraint on $\mathscr{A}$. We will say that $F > 0$ when for any $a, b \in \mathscr{A}$, such that $b \leq a$, $F(\omega_a|\omega_b) > 0$ for any ω_b such that $G^a_b(\omega_a) = \omega_b$; $G \circ F = \mathrm{id}$ requires that $F(\omega_a|\omega_b) = 0$ when $G^a_b(\omega_a) \neq \omega_b$.

The following lemma is an extension of the classical result that states that conditioning over a σ-subalgebra $\mathscr{F}_1 \subseteq \mathscr{F}$ defines a morphism of modules when finer ($\mathscr{F}$ measurable) observables are seen as modules over the coarser ($\mathscr{F}_1$ measurable observables).

Lemma 3.1. *Let E_1, E_2 be two measurable spaces, let $g : E_2 \to E_1$ be a measurable map and $f : E_1 \to E_2$ be a Markov kernel so that, $f \circ g = \mathrm{id}$. Let us denote respectively i and π the induced linear maps on $L^\infty(E_1), L^\infty(E_2)$. Let $h \in L^\infty(E_2)$ and $k \in L^\infty(E_1)$, then,*

$$\pi(h).k = \pi(h.i(k)) \tag{5}$$

Proof. Let us first prove the result in the particular case when $h = 1_B$ with $B \in \sigma(E_2)$ and $k = 1_A$ with $A \in \sigma(E_1)$. Let us denote $\overline{A}$ the complement of A, then $1_A + 1_{\overline{A}} = 1$ and $1_A.1_{\overline{A}} = 0$. Furthermore $i(1_A) = g^{-1}A$

$$\pi(1_B) = \pi(1_B.1_{g^{-1}A}) + \pi(1_B.1_{\overline{g^{-1}A}}) \tag{6}$$

and

$$\pi(1_B.1_{g^{-1}A}) \leq \pi(1_{g^{-1}A}) = \pi \circ i(1_A) = 1_A \tag{7}$$

$$\pi(1_B.1_{\overline{g^{-1}A}}) \leq \pi(1_{\overline{g^{-1}A}}) = 1_{\overline{A}} \tag{8}$$

Therefore,

$$\pi(1_B)1_A = \pi(1_B.1_{g^{-1}A})1_A + \pi(1_B.1_{g^{-1}A})1_A \tag{9}$$

But, $\pi(1_B.1_{g^{-1}A})1_A \leq 1_{\overline{A}}.1_A = 0$ so, $\pi(1_B)1_A = \pi(1_B.1_{g^{-1}A})1_A$. Furthermore, $\pi(1_B.1_{g^{-1}A}) = \pi(1_B.1_{g^{-1}A})1_A + \pi(1_B.1_{g^{-1}A})1_{\overline{A}}$ therefore,

$$\pi(1_B.1_{g^{-1}A})1_{\overline{A}} \leq \pi(i(1_A)).1_{\overline{A}} = 0 \tag{10}$$

We just showed that,

$$\pi(1_B.1_{g^{-1}A}) = \pi(1_B.1_{g^{-1}A})1_{\overline{A}} \tag{11}$$

So $\pi(1_B).1_A = \pi(1_B.i(1_A))$. The result then extends by linearity directly to $h = \sum_{k \leq n} 1_{B_k}$ and $k = \sum_{k \leq n_1} 1_{A_n}$, which ends the proof.

□

Let us remark that if $A \in \lim \sigma(G)$ then $\overline{A} := (\overline{A}_a, a \in \mathscr{A})$ is also in $\lim i$.

Proposition 3.1. *Let $\gamma = (G, F)$ be a specification, let $G(a)$ be finite sets for any $a \in \mathscr{A}$, let $F > 0$. Let $\mu \in \mathscr{G}(\gamma)$, for any $f \in \prod_{a \in A} L^\infty(G(a))$, such that $\forall a \in \mathscr{A}, \mu_a(f_a) = 1$,*

$$f.\mu \in \mathscr{G}(\gamma) \iff \exists \tilde{f} \in \lim i, \ s.t. \ f.\mu = \tilde{f}.\mu \tag{12}$$

Proof. Let us assume that $f.\mu \in \mathscr{G}(\gamma)$, then for any $a, b \in \mathscr{A}$ such that $b \leq a$, and for any $g_a \in L^\infty(G(a))$, by hypothesis, $(f.\mu)_b \pi^a_b(g_a) = (f.\mu)_a(g_a)$; it can be rewritten as,

$$\mu_b \left(\pi^a_b(g_a).f_b\right) = \mu_a \left(f_a.g_a\right) \tag{13}$$

By Lemma 3.1, $\pi^a_b(g_a).f_b = \pi^a_b(g_a.i^b_a f_b)$; therefore,

$$\mu_b \pi^a_b(g_a.i^b_a f_b) = \mu_a \left(f_a.g_a\right) \tag{14}$$

Therefore $f_a = i^b_a f_b$ μ_a-almost surely.

We will now show that there is $\tilde{f} \in \lim i$ such that $\tilde{f}.\mu = f.\mu$. It is in this part of the proof that we assume that $G(a)$ are finite sets and that $F > 0$. Let's call $S_a := \mathrm{supp}\mu_a$, the support of μ_a, i.e., the set $\mathrm{supp}\mu_a := \{\omega_a \in G(a) | \mu_a(\omega_a) > 0\}$. Let us denote $N_a := \overline{S_a}$ its complement and $M_a = 1_{\overline{N_a}}$. We will now show that $(N_a, a \in \mathscr{A}) \in \lim \sigma(G)$.

For any $b, a \in \mathscr{A}$ such that $b \leq a$, $\mu_a i^b_a = \mu_b$; therefore, as $\mu_b(M_b) = 0$, one has that $\mu_a i^b_a(M_b) = 0$. Recall that $i^b_a(M_b)$ is the indicator function of the set $G^{a\,-1}_b N_b$; the previous remark implies that $i^b_a(M_b) \leq M_a$. Furthermore, $\mu_b \pi^a_b(M_a) = \mu_a(M_a) = 0$; therefore, $\pi^a_b(M_a) \leq M_b$.

Hence, as $\pi^a_b(i^b_a M_b) = M_b$ and $i^b_a M_b \leq M_a$, then by applying π^a_b on both sides, $M_b \leq \pi^a_b(M_a)$. And so $\pi^a_b(M_a) = M_b$.

Recall that we showed that $\pi^a_b(M_a) = M_b$ and $i^b_a(M_b) \leq M_a$. In particular, $M_a - i^b_a(M_b) \geq 0$; furthermore, $\pi^a_b(M_a - i^b_a(M_b)) = 0$ so $M_a = i^b_a(M_b)$. To be more explicit: $\forall \omega_b \in G(b)$,

$$\pi^a_b(M_a - i^b_a(M_b))(\omega_b) = \sum_{\omega_a \in G(a)} F(\omega_a|\omega_b)[M_a - i^b_a(M_b)](\omega_a) \tag{15}$$

As $F(\omega_a|\omega_b) > 0$ by hypothesis, then $M_a = i^b_a(M_b)$. This implies that $M \in \lim i$ and $N \in \lim \sigma(G)$. This also implies that $S \in \lim \sigma(G)$.

Let $\tilde{f} = f 1_S$. Then $f 1_S \in \lim F$ and for any $a \in \mathscr{A}$, $f_a = \tilde{f}_a$ μ_a-a.s., which ends the proof.

□

One remarks that $\lim i$ is a subset of $\lim \pi$: for any $f \in \lim i$, by definition for any $a, b \in \mathscr{A}$ such that $b \leq a$, $i^b_a f_b = f_a$ and so $\pi^a_b i^b_a f_b = \pi^a_b f_a$ so $f_b = \pi^a_b f_a$.

Let us also remark that for any $b \leq a \in \mathscr{A}$, $\mu_a(A_a) = \mu_b(A_b)$ for $\mu \in \mathscr{G}(\gamma)$, $A \in \lim G$.

Theorem 3.1 (Extreme measure characterisation (Generalisation of Theorem 7.7 [1])). *Let $\gamma = (G, F)$ be a specification, let $G(a)$ be finite sets for any $a \in \mathscr{A}$, let $F > 0$. $G(\gamma)$ is a convex set. Each $\mu \in \mathscr{G}(\gamma)$ is uniquely determined by it's restriction to $\lim \sigma(G)$. Furthermore μ is extreme in $\mathscr{G}(\gamma)$ if and only if for any $A \in \lim \sigma(G)$, $\forall a \in \mathscr{A}$, $\mu_a(A_a) = 0$ or 1.*

Proof. Let us denote π^* the functor from $\mathscr{A}$ to **Vect** for which for any $b \leq a$, $\pi^{*b}_{\ a} : (L^\infty F(b))^* \to (L^\infty F(a))^*$ is the dual of π^a_b that send linear forms to linear forms. Then $\mathscr{G}(\gamma)$ is a subspace of the vector space $\lim F^*$ and furthermore for any $a \in \mathscr{A}$ and $p \in [0,1]$, $p\mu_a + (1-p)\nu_a \in \mathbb{P}(G(a))$ whenever $\mu_a, \nu_a \in \mathbb{P}(G(a))$. Therefore $\mathscr{G}(\gamma)$ is a convex set.

Proposition 3.1, allows us to apply a similar proof, when done with caution, to the one found of Theorem 7.7 in [1]. Let us recall the proof. Let $\mu, \nu \in \mathbb{G}(\gamma)$ such that $\mu_{|\lim i} = \nu_{|\lim i}$. Let $\overline{\mu} = \frac{\mu+\nu}{2}$, then $\overline{\mu} \in \mathscr{G}(\gamma)$. But μ and ν are absolutely continuous with respect to $\overline{\mu}$ therefore for any $a \in \mathscr{A}$ there is $f_a, g_a \in L^\infty(G(a))$ such that $\mu_a = f_a\overline{\mu}_a$ and $\nu_a = g_a\overline{\mu}_a$. By Proposition 3.1, $f, g \in \lim i$. By hypothesis, for any $h \in \lim i$ $\overline{\mu}_a(h_a) = \mu_a(h_a) = \nu_a(h_a)$. Importantly i^b_a is a ring morphism of $L^\infty(G(b))$, i.e. $i^b_a(k_b.h_b) = i^b_a(k_b).i^b_a(h_b)$. Therefore for any $h, k \in \lim i$, $k.h \in \lim i$; as $f - g \in \lim i$, then it is also the case that $(f-g)^2 \in \lim i$; but for any $a \in \mathscr{A}$,

$$\overline{\mu}_a[(f_a - g_a)^2] = 0 \tag{16}$$

so $f_a = g_a \quad \overline{\mu}_a -$ a.s. Therefore $f\overline{\mu} = g\overline{\mu}$ and $\mu = \nu$.

Showing that $\mu \in \mathscr{G}(\gamma)$ is extreme is equivalent to μ being trivial on $\lim i$ is a direct generalization of Corollary 7.4 [1] thanks to Proposition 3.1. Let $\mu \in \mathscr{G}(\gamma)$ be not trivial on $\lim \sigma(G)$ then there is $A = (A_a, a \in \mathscr{A}) \in \lim \sigma(G)$ such that,

$$\forall a \in \mathscr{A}, \quad 0 < \mu_a(A_a) < 1 \tag{17}$$

Therefore for any $a \in \mathscr{A}$,

$$\mu_a = \mu_a(A_a)\frac{1_{A_a}.\mu_a}{\mu_a(A_a)} + \mu_a(\overline{A}_a)\frac{1_{\overline{A}_a}.\mu_a}{\mu_a(\overline{A}_a)} \tag{18}$$

Furthermore, for any $b \leq a$,

$$i^b_a \frac{1_{A_b}}{\mu_a(A_b)} = \frac{1_{A_a}}{\mu_a(A_b)} = \frac{1_{A_a}}{\mu_a(A_a)} \tag{19}$$

Therefore $\frac{1_{A_a}}{\mu_a(A_a)}, a \in \mathscr{A}$ is in $\lim i$ and so by Lemma 3.1, $\left(\frac{1_{A_a}.\mu_a}{\mu_a(A_a)}, a \in \mathscr{A}\right) \in \mathscr{G}(\gamma)$. Similarly $\left(\frac{1_{\overline{A}_a}.\mu_a}{\mu_a(\overline{A}_a)}, a \in \mathscr{A}\right) \in \mathscr{G}(\gamma)$. In particular there is $0 < p < 1$ so that $\mu = p\nu + (1-p)\nu_1$ with $\nu, \nu_1 \in \mathscr{G}(\gamma)$. Therefore μ is not an extreme measure.

Assume now that $\mu \in \mathscr{G}(\gamma)$ is such that for any $A \in \lim \sigma(A)$, and any $a \in \mathscr{A}$, $\mu_a(A_a) = 0$ or 1. Suppose that there is $0 < p < 1$ such that $\mu = p\nu + (1-p)\nu_1$ with $\nu, \nu_1 \in \mathscr{G}(\gamma)$. Then for any $a \in \mathscr{A}$, ν_a, ν_{1a} is absolutely continuous with respect to μ_a. Therefore, there are $(f_a \geq 0, a \in \mathscr{A})$, $(g_a \geq 0, a \in \mathscr{A})$ both in $\prod_{a \in \mathscr{A}} L^\infty(G(a))$ such that $\nu = f\mu$ and $\nu_1 = g\mu$. As $\nu, \nu_1 \in \mathscr{G}(\gamma)$, then by Lemma 3.1, $f, g \in \lim i$. Therefore, for all $a \in \mathscr{A}$,

$\mu_a(f_a) = 0$ or for all $a \in \mathscr{A}$, $\mu_a(g_a) = 0$. So, $\mu = \nu$ or $\mu = \nu_1$ and μ is extreme in $\mathscr{G}(\gamma)$.

$\square$

Let us remark that if $\mathscr{A}$ has only one connected component, then for $A \in \lim \sigma(G)$, satisfying $\forall a \in \mathscr{A}, \mu_a(A_a) = 0$ or 1 is equivalent to $\exists a \in \mathscr{A}, \mu_a(A_a) = 0$ or 1. Indeed, if a, b are in the same connected component, i.e., $a \leq b$ or $b \leq a$, then $\mu_a(A_a) = \mu_b(A_b)$.

References

[1] H.-O. Georgii, *Gibbs Measures and Phase Transitions.* de Gruyter, 1988.

[2] M. Giry, "A categorical approach to probability theory," in *Categorical aspects of topology and analysis*, vol. 915 of *Lecture notes in Mathematics*, pp. 68–85, Springer, 1982.

[3] G. Sergeant-Perthuis, "A categorical approach to statistical mechanics," in *Geometric Science of Information* (F. Nielsen and F. Barbaresco, eds.), (Cham), pp. 258–267, Springer Nature Switzerland, 2023.

[4] G. Sergeant-Perthuis, *Intersection property, interaction decomposition, regionalized optimization and applications.* PhD thesis, Université de Paris, 2021.

[5] G. Sergeant-Perthuis, "Interaction decomposition for presheaves," 2021.

[6] G. Sergeant-Perthuis, "Bayesian/graphoid intersection property for factorisation spaces," 2021.

[7] G. Sergeant-Perthuis, "Intersection property and interaction decomposition," 2021.

[8] G. Sergeant-Perthuis, "Interaction decomposition for hilbert spaces," 2021.

[9] D. Bennequin, O. Peltre, G. Sergeant-Perthuis, and J. P. Vigneaux, "Extra-fine sheaves and interaction decompositions," 2020.

[10] G. Sergeant-Perthuis, "Regionalized optimization," 2022.