Random matrices and Lyapunov coefficients regularity

G.Gallavotti
INFN-Roma1 and Rutgers University
January 30, 2014

Abstract: *Analyticity and other properties of the largest or smallest Lyapunov exponent of a product of real matrices with a "cone property" are studied as functions of the matrices entries, as long as they vary without destroying the cone property. The result is applied to stability directions, Lyapunov coefficients and Lyapunov exponents of a class of products of random matrices and to dynamical systems. The method is based on the classical theory of the Mayer series in Statistical Mechanics of rarefied gases.*

1 Introduction and paradigm

Regularity of the Lyapunov exponents of products of random matrices has been studied thoroughly in [11]. The study dealt with various aspects and consequences of the following *cone property* (as it will be called here):

Definition 1: *A sequence* $\{T_j\}_{-\infty}^{\infty} = \dots, T_0, T_1, T_2, \dots$ *of* $d \times d$ *real invertible matrices has the* (Γ, Γ')*-cone property if there are proper, closed convex cones* $\Gamma, \Gamma' \subset R^d$*, with apex at the origin* O *and* $\Gamma' \subset \Gamma$*, such that* $T_j\Gamma \subset \Gamma'$ *and* Γ'/O *is contained in the interior* Γ^0 *of* Γ*.*

It will be convenient to imagine the matrix T_j attached to the point j of a lattice to which is also attached the linear space $E_n = R^d$ on which T_j acts transforming a vector $v \in E_j$ into $T_j v$ which is regarded as an element of E_{j-1}: $T_j E_j = E_{j-1}$.

The proofs in [11] are based on the implicit function theorem and on the results in [2]. Here the aim is to obtain most of the results in [11], relative to the finite dimensional case and product of matrices, with a different and self contained technique. The "cone property" importance for studies beyond the Lyapunov exponents, like decay of correlations in smooth and non smooth dynamical systems, has been developed in [7]: the new techniques of the latter work (and in the many stemming out of it) are also quite different from the ones presented here.

To present the main idea of this work imagine the T_n diagonalizable; but

the results will cover the general case. Then T_n will be written:

$$T_n = \sum_{\sigma=0}^{d-1} \lambda_{n,\sigma} \, |\sigma, n\rangle \langle \sigma, n| \tag{1.1}$$

where $\lambda_{n,\sigma}$ are the eigenvalues of T_n and the vectors $|n,\sigma\rangle, \langle n,\sigma|$ are the corresponding right and left eigenvectors which will be supposed normalized to $\langle n\sigma|n,\sigma'\rangle \equiv \delta_{\sigma\sigma'}$. The eigenvalues will be labeled by decreasing modulus $|\lambda_{n,0}| \geq |\lambda_{n,1}| \geq \ldots \geq |\lambda_{n,d-1}|$ and $\lambda_{n,\sigma} \neq 0$ (invertibility condition).

Notice that largest Lyapunov exponent vectors $h(n)$, assumed existing, have to be inside the cone Γ' and should satisfy $\Lambda(n,p)h(n) = T_n \cdots T_{p-1}h(p)$ for $p \geq n$ and $\Lambda(n,p)$ suitable. Therefore, suitably fixing a convenient normalization for $h(n)$, it should be obtainable as a limit as $N \to \infty$ of $H(n,N) \stackrel{def}{=} T_n \cdots T_N \; v, \; v \in \Gamma'$ which can be written as

$$H(n,N) = \sum_{\sigma_n,\ldots,\sigma_N} |n,\sigma_n\rangle \cdot \Big(\prod_{j=n}^{N} \lambda_{j,\sigma_j} \Big) \prod_{j=n}^{N} \langle j,\sigma_j|j+1,\sigma_{j+1}\rangle \tag{1.2}$$

with $|N+1,\sigma_{N+1}\rangle \equiv v$ (here σ_{N+1} is just a label and not an index of summation, being only used for uniformity of notation).

The representation in Eq.(1.2) suggests an alternative approach to the analysis of such products of matrices, directly inspired by the methods of 1-*dimensional* statistical mechanics of spin systems.

To bring it to a more familiar form: consider intervals of integers $J = [h,h'] = (h, h+1, \ldots, h')$ with $1 \leq h \leq h' \leq N$ and, on them, *spin configurations* $\boldsymbol{\sigma}_J = (\sigma_h, \ldots, \sigma_{h'})$ with $\sigma_j = 1, \ldots, d-1$ (*i.e.* $\sigma_j \neq 0$, see Eq.(1.1)).

Call $Y = (J, \boldsymbol{\sigma}_J)$ a "*polymer*" with base J and structure $\boldsymbol{\sigma}_J$ (if $|J| = \ell$; there are ℓ^{d-1} polymers with base J): then Eq.(1.2) can be interpreted as an expectation value evaluated in an ensemble of polymers as follows.

A "*configuration*" of polymers in $[1,N]$ will be $\mathbf{Y} = (J_1, \boldsymbol{\sigma}_{J_1}, \ldots, J_s, \boldsymbol{\sigma}_{J_s})$ with $J_1 < J_2 < \ldots < J_s$, $s > 0$: *i.e.* $\mathbf{Y}$ is a configuration of non overlapping polymers ("hard core polymers" of size $|J_i| \geq 1$). With each polymer $(J, \boldsymbol{\sigma}_J)$ associate an "activity" $I(J, \boldsymbol{\sigma}_J)$ (I might even be complex).

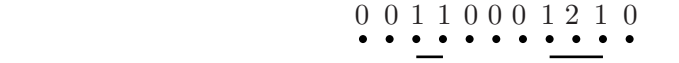

Fig.1: Spin configuration for 3×3 matrices containing two polymers, of sizes $2, 3$.

A formal probability distribution ("ensemble") on the $\mathbf{Y}$'s is obtained by attributing a *weight* $\zeta(\mathbf{Y}) \stackrel{def}{=} \prod_{i=1}^{s} I(J_i, \boldsymbol{\sigma}_i)$ equal to the product of the

activities; the empty configuration is given weight 1. The ensemble thus defined is formal as $I(J, \boldsymbol{\sigma}_J) \geq 0$ is not required .

After some meditation, it is recognized that Eq.(1.2), can be rewritten imagining the sites with $\sigma_i = 0$ as "empty sites" in polymer configurations, and it can be cast in the form

$$H(n,N) = \overline{\Lambda}(n,N) \sum_{\sigma_n,\dots,\sigma_N} \sum_{s\geq 0} \sum_{J_1<\dots<J_s} |n,\sigma_n\rangle \, \frac{\prod_{i=1}^{s} I(J_i, \boldsymbol{\sigma}_{J_i})}{\Omega(n,N)} \tag{1.3}$$

where $\overline{\Lambda}(n,N), \Omega(n,N)$ are normalization factors and

$$\begin{aligned} I(J,\boldsymbol{\sigma}_J) &\stackrel{def}{=} \Big(\prod_{j=h}^{h'} \frac{\lambda_{j,\sigma_j}}{\lambda_{j,0}} \Big) \Big(\prod_{j=h}^{h'-1} \frac{\langle j,\sigma_j | j+1,\sigma_{j+1}\rangle}{\langle j,0|j+1,0\rangle} \Big) \\ &\quad \cdot \Big(\frac{\langle h-1,0|h,\sigma_h\rangle}{\langle h-1,0|h,0\rangle} \Big)^{\delta_{h>n}} \Big(\frac{\langle h',\sigma_{h'}|h'+1,0\rangle}{\langle h',0|h'+1,0\rangle} \Big) \\ \Omega(n,N) &\stackrel{def}{=} \sum_{s\geq 0} \sum_{\substack{J_1<\dots<J_s \\ \boldsymbol{\sigma}_{J_1},\dots,\boldsymbol{\sigma}_{J_s}}} \prod_{i=1}^{s} I(J_i,\boldsymbol{\sigma}_{J_i}) \\ \overline{\Lambda}(n,N) &\stackrel{def}{=} \Omega(n,N) \prod_{j=n}^{N} \Big(\lambda_{j,0} \ \langle j,0|j+1,0\rangle \Big) \end{aligned} \tag{1.4}$$

where $|N+1,\sigma\rangle$ has to be interpreted, *instead*, as v, see Eq.(1.2), and, for $s > 0$, $J_1 < \dots < J_s$ are consecutive intervals in $[n,N]$ *not empty and disjoint*; $s = 0$ contributes 1 to $\Omega(n,N)$.

Eq.(1.3) maps the problem of studying $H(n,N)$ into the study of $\overline{\Lambda}(n,N)$ times a *formal average* in what is known in statistical mechanics as a *Fisher model*, [1]. The latter is well known, since [1], as a machine for examples and counterexamples in statistical mechanics and in dynamical systems, for some applications see [3], [5, Def. D7.3.1]).

Here Eq.(1.3) will be the starting point to obtain, with an alternative method, the following theorems, special cases of results already in [11]:

Theorem 1: *(1) Let T_n be a sequence as in definition 1 (hence with the (Γ,Γ')-cone property) with $||T_n|| < B_0$ for some $B_0 > 0$. Then given any sequence $\{v_j\}$ of unit vectors in Γ', the limits*

$$\begin{aligned} h(n) &= \lim_{N\to\infty} \frac{T_n T_{n+1} \cdots T_{N-1}\, v_N}{\overline{\Lambda}_{v_N}(n,N)} = \lim_{N\to\infty} \frac{H_{n,N} v_N}{\overline{\Lambda}_{v_N}(n,N)} \\ \Lambda(n,p) &= \lim_{N\to\infty} \frac{\overline{\Lambda}_{v_N}(p,N)}{\overline{\Lambda}_{v_N}(n,N)} > 0, \end{aligned} \tag{1.5}$$

exist $\forall 1 \leq n \leq p$ *and are independent of the sequence* $\{v_j\}_{j\geq 1}$.
(2) The limits are (real) analytic functions of each of the matrices entries, as long as their variations are small enough.
(3) The vectors $h(n)$ *are "eigenvectors" for the product of the inverse matrices* T_j^{-1}*, i.e. there are "Lyapunov coefficients",* $\Lambda(n,p)$ *such that:*

$$T_{p-1}^{-1} \dots T_n^{-1}\, h(n) \;=\; \Lambda(n,p)\, h(p), \qquad p > n \tag{1.6}$$

(4) There is B *such that* $B^{-1} \leq ||h(n)|| \leq B$*,* $\Lambda(n,p) > 0$*, and the upper and lower limits of* $\frac{1}{p}\log\Lambda(n,p)$*, as* $p \to \infty$*, are* n *independent.*

Remarks: (a) The uniqueness property implies that $h(n)$ is the *unique* eigenvector (*i.e.* invariant vector) with the largest Lyapunov exponent.
(b) $h_n = \frac{h(n)}{||h(n)||}$ is called *unstable unit (*or *direction) vector* at site n.
(c) Imagine the matrices $(T_j)_{j=1}^{\infty}$ be a sequence of random variables and that their entries are distributed with a distribution ρ which is invariant with respect to the (left) translations, and with samples restricted to keep the cone property, *i.e.* with the cones Γ, Γ' which do not depend on the choice of the matrices. Then the upper and lower limits of $\frac{1}{p}\log|\Lambda(n,p)|$ as $p \to \infty$ will be constant under translation, *i.e.* n-independent as a consequence of the n-independence in item (4). They will be shown below to exist almost everywhere, as a consequence of the ergodic theorem, hence if ρ is ergodic they will be equal and constant with ρ probability 1.
(d) A more general analyticity property holds:

Theorem 2: *Let* T_h *depend on a real parameter* $z \in [-\overline{\nu}, \overline{\nu}] = \Delta$*, admit a power series around each point in* Δ *with radius* ν *and, furthermore, for all* $z \in \Delta$ *they have the cone property with respect to* z*–independent cones* Γ, Γ'*. Then the vectors* b_h *and the Lyapunov coefficients* $\Lambda(n,p)$ *are holomorphic in* z *for* $|z - \Delta| < \nu'$ *for some* ν'*.*

Theorems 1,2 are paradigms: they are the basis for similar theorems for dynamical systems, *e.g.* [11], and Appendix C.

2 Theorems 1,2

The essence of the proof lies in understanding the case in which the matrices T_j are diagonalizable with real eigenvectors and eigenvalues uniformly (in j) pairwise separated, and possess a (Γ, Γ')-cone property; furthermore $\max_{n,\sigma=1,\dots,d-1} \frac{|\lambda_{\sigma,n}|}{|\lambda_{0,n}|} < \varepsilon_0$ with ε_0 small enough.

The more general case contemplated in theorem 1 (*i.e.* just the cone property and a uniform bound on $||T_n||$ is supposed) will be eventually

reduced to the latter one via the following algebraic lemmata, see proof in Appendix A:

Lemma 1: *Let T be a (single) $d \times d$ matrix with the (Γ, Γ')-cone property. Then:*
(a) $\frac{T^n v}{||T^n v||} \xrightarrow[n\to+\infty]{} b$ exponentially fast and $\frac{Tb}{||Tb||} = b$ for all $v \in \Gamma$. A corresponding property holds for the transposed T^* with b replaced by a b^*.
(b) The eigenvalue λ_0 of T with maximum modulus is simple and positive, hence there is $\gamma < 1$ such that $\max_{\sigma>0} \frac{|\lambda_\sigma|}{|\lambda_0|} \leq \gamma < 1$.

Definition 4: *Let λ the modulus of the largest modulus eigenvalue of a matrix and λ' the maximum modulus of the other eigenvalues; call, here, $\frac{\lambda'}{\lambda}$ the matrix "spectral gap".*

The proof of lemma 1 leads to, see Appendix A,

Lemma 2: *Suppose that the sequence $T_1, T_2, \ldots$ satisfies the cone property with respect to the pair of cones $\Gamma \supset \Gamma'$ and let $T' \stackrel{def}{=} T_1 \cdot T_2 \cdots T_p$. Then there are constants $c, \alpha > 0$ with $\alpha < 1$ such that the spectral gap of T' is $\leq c\alpha^p$. Furthermore the matrix elements of T' on the basis formed by b (see lemma 1) and by $d-1$ unit vectors in the plane orthogonal to b^* are all bounded by $c\alpha^p$ with respect to the entry $|T'^{[p]}_{0,0}| \geq \frac{1}{c}$. The c, α depend on the inclination ε and on the openings ϑ, ϑ' of the cones.*

This implies that, if p is large enough, the sequence $T'_n = T_{np+1} T_{np+2} \cdots T_{(n+1)p}, n = 0, 1, \ldots,$ satisfies the cone property with respect to the *same* pair of cones $\Gamma \supset \Gamma'$ and the spectral gap of the matrices T'_n can be made *as small as wished* by taking p large enough.

Hence it is sufficient to prove theorem 1 for matrices T'_j with the above defined spectral gap $\frac{\lambda'}{\lambda} = \gamma$ as small as needed. Once the sequence b'_n for T'_n is obtained the sequence b_n that has to be found will be obtaines setting $b_{np} \stackrel{def}{=} b'_n$ and:

$$b_{np-k} = T_{np+k} \cdots T_{np} b'_{np}, \quad k = 1, \ldots, p \tag{2.1}$$

and all the statements in theorem 1 will, as well, follow for the sequence T_n.

proof of theorem 1: Suppose *at first* that the matrices T_n have real eigenvalues *with reciprocal distance*, as n varies, greater than a positive lower bound. The $\Omega(n, N)$ in Eq.(1.4) can be interpreted as the partition function of a gas of polymers represented by (lattice) intervals $J = [h, h'] \subset [1, N]$ with "base" J, "structure" $\boldsymbol{\sigma}_J = (\sigma_h, \ldots, \sigma_{h'})$ and "activity" $I(J, \boldsymbol{\sigma}_J)$ defined

in Eq.(1.4). Of course this is only an analogy as $I(J, \boldsymbol{\sigma}_J)$, although real numbers (under the restrictive temporary assumption) might be < 0.

A simple bound can be set on $|I(J, \boldsymbol{\sigma}_J)|$ by remarking that the cone property implies a n-independent lower bound $\frac{1}{\delta} > 0$ on the scalar products appearing in the denominators in the definition Eq.(1.4): actually $\frac{1}{\delta} > 0$ can be chosen as a lower bound for the absolute value of the product of any pair or unit vectors in Γ. If γ is an upper bound to the matrices spectral gaps, see definition 4, Eq.(1.4) implies. *also defining* η:

$$\sum_{\boldsymbol{\sigma}_J} |I(J, \boldsymbol{\sigma}_J)| \leq ((d-1)\delta\gamma)^{h'-h+1} \stackrel{def}{=} \eta^{h'-h+1} \tag{2.2}$$

Remarks: (1) As shown by the last bound, forgetting that the activities may be be negative and treating $I(J, \boldsymbol{\sigma}_J)$ as weight (or activity) of the polymer $(J, \boldsymbol{\sigma}_J)$ in the formal probability distribution of the polymers appearing in Eq.(1.3), the number of J's should be very small if η (*i.e.* γ) is small: the "preferred state" is no J at all; but the $I(J, \boldsymbol{\sigma}_J)$ may be negative and cannot be regarded as probability weights.
(2) Hence it will be necessary to evaluate the "averages" (like Eq.(1.3)) with respect to the "distribution" in which a configuration of polymers with bases $J_1 < \ldots < J_s$ and spin structures $\boldsymbol{\sigma}_{J_1}, \ldots, \boldsymbol{\sigma}_{J_s}$ has weight $\prod_{i=1}^{s} I(J_i, \boldsymbol{\sigma}_{J_i})$: this must be done *algebraically*, *i.e.* without profiting of positivity properties. It is natural to have recourse to *the cluster expansion*: which is a method that was designed, in statistical mechanics, precisely for such tasks.

The general theory of the cluster expansion for polymers, see [5, Ch.7] (it is recalled from scratch, *for completeness*, in Appendix B), yields a formal expression for $\Omega(n, N)$ as

$$\Omega(n, N) = \exp \sum_{\mathbf{Y}} \varphi^T(\mathbf{Y}) \zeta(\mathbf{Y}), \qquad \zeta(\mathbf{Y}) = \prod_{J \in \mathbf{Y}} I(J, \boldsymbol{\sigma}_J) \tag{2.3}$$

where the summation runs over all polymer configurations $\mathbf{Y}$ consisting of $Y_1, \ldots, Y_s$, with $Y_i = (J_i, \boldsymbol{\sigma}_{J_i})$, in which the constraint of no overlap on the polymers base intervals J_i (implied by the $*$ in Eq.(1.4)) is *dropped* and $\varphi^T(\mathbf{Y})$ are suitable (real) *combinatorial coefficients*, see Appendix B, Eq.(B.8).

The coefficients $\varphi^T(\mathbf{Y})$ have the important property of being *translation invariant* under a simoultaneous translation of the polymers in $\mathbf{Y}$ and of *vanishing unless the intervals J which are bases of the polymers in $\mathbf{Y}$ overlap* in the sense that $\cup_{J \in \mathbf{Y}} J$ is a connected interval, called a *cluster* (imagine to draw the J_i's as continuous segments joining their extremes).

Abridge $\varphi^T(\mathbf{Y})\zeta(\mathbf{Y})$ into $\widehat{\varphi}(\mathbf{Y})$: the sum $\sum_{\mathbf{Y},J_i\subset[n,N]}\widehat{\varphi}^T(\mathbf{Y})$ over the clusters $\mathbf{Y}$ will be an absolutely convergent series, summing to $\log\Omega(n,N)$, if it will be shown that, for a suitable choice of $r(J)$, it is

$$\mu \stackrel{def}{=} \sup_{J,\boldsymbol{\sigma}_J} \frac{|\zeta(J,\boldsymbol{\sigma}_J)|}{r(J)} \exp \sum_{S,\boldsymbol{\sigma}_S}^{*} r(S) < 1 \tag{2.4}$$

where the sum is over the polymers $(S,\boldsymbol{\sigma}_S)$ overlapping with J, *i.e.* $S\cap J\neq\emptyset$, as recalled in a self contained proof, in Appendix B.

This is a non trivial property: certainly the overlap condition strongly reduces the number of addends in the series for $\Omega(n,N)$ and helps together with the Eq.(2.4), which implies that $\zeta(\mathbf{Y})$ is exponentially small with the size of the interval covered by the clusters of the polymers bases in $\mathbf{Y}$. However the help is not sufficient and important combinatorial cancellations have to be taken into account: they are exhibited through the *remark that the coefficients* $\varphi^T(\mathbf{Y})\zeta(\mathbf{Y})$ *satisfy an identity*, see Eq.(B.12), reducible to an identity known in Physics as *Kirkwood-Salsburg equations*, [4].

Take $r(S)\stackrel{def}{=}\eta^{\frac{1}{2}|S|}$, see Eq.(2.2); if $\eta=(d-1)\delta\gamma$ is small enough:

$$\mu \leq \sup_{|J|=j\geq 1} \frac{\eta^j}{\eta^{\frac{1}{2}j}} \exp j\Big(\sum_{\ell=0}^{\infty}(\ell+1)\eta^{\frac{1}{2}(\ell+1)}\Big) = \eta^{\frac{1}{2}} \exp \frac{\eta^{\frac{1}{2}}}{(1-\eta^{\frac{1}{2}})^2} < 1 \tag{2.5}$$

Setting, see Eq.(1.3),

$$h(n,N)\stackrel{def}{=}\sum_{s\geq 0}\sum_{\substack{J_1<\ldots<J_s\\ \boldsymbol{\sigma}_{J_1},\ldots,\boldsymbol{\sigma}_{J_s}}} |n,\sigma_n\rangle \frac{\prod_{i=1}^s I(J_i,\boldsymbol{\sigma}_{J_i})}{\Omega(n,N)} \tag{2.6}$$

this, as remarked in Sec.1, can be interpreted as $\sum_{\sigma=0}^{d-1}P_{n,\sigma,N}|n,\sigma\rangle$ with

$$P_{n,\sigma,N}\stackrel{def}{=}\frac{\Omega_\sigma(n,N)}{\Omega(n,N)}\stackrel{def}{=}\sum_{s\geq 0}\sum_{\substack{J_1<\ldots<J_s\\ \boldsymbol{\sigma}_{J_1},\ldots,\boldsymbol{\sigma}_{J_s}}}^{*\sigma} \frac{\prod_{i=1}^s I(J_i,\boldsymbol{\sigma}_{J_i})}{\Omega(n,N)} \tag{2.7}$$

where the $*\sigma$ indicates that the sum is restricted to polymers configurations with $\sigma_n=\sigma$: *i.e.* with $n\in J_1$ and $\sigma_n=\sigma$ if $\sigma\geq 1$ or with $n\notin J_1$ if $\sigma_n=0$.

Hence if the activities $I(J,\boldsymbol{\sigma}_J)$ were non negative $P_{n,\sigma,N}$ would be the probability of finding a configuration with spin σ at site n, and $P_{n,0,N}+\sum_{\sigma=1}^{d-1}P_{n,\sigma,N}\equiv 1$. This relation is a purely algebraic property and is identically satisfied (as long as the addends are well defined).

For the $P_{n,\sigma,N}$ with $\sigma \geq 1$ the cluster expansion (see Eq.(B.15)) yields

$$P_{n,\sigma,N} = \sum_{J \ni n, \boldsymbol{\sigma}_J, \sigma_n = \sigma} \langle\, D_{J,\boldsymbol{\sigma}_J} \widehat{\varphi}^T \,\rangle, \qquad |P_{n,\sigma,N}| \leq \frac{\eta^{\frac{1}{2}}}{(1-\mu)(1-\eta^{\frac{1}{2}})} \tag{2.8}$$

where $\langle\, D_{J,\boldsymbol{\sigma}_J} \widehat{\varphi}^T \,\rangle \stackrel{def}{=} \sum_{\mathbf{Y}} \widehat{\varphi}^T((J,\boldsymbol{\sigma}_J) \cup \mathbf{Y})$; the bound is obtained by using the mentioned overlap property that the $\widehat{\varphi}^T(\mathbf{Y})$ vanish unless the bases of the polymers in $\mathbf{Y}$ form a connected interval, see below.

The overlap property shows that the sum $\langle\, D_{J,\boldsymbol{\sigma}_J} \widehat{\varphi}^T \,\rangle$ is restricted to polymers that contain the site n: hence the union of the bases of the polymers contributing terms that depend on the boundary vector v must cover the whole $[n,N]$[1] and therefore contribute a quantity that is exponentially small as $N \to \infty$, as mentioned above.

Explicit bounds, see Eq.(B.13), (B.14), if Eq.(2.4) holds, give $|P_{n,\sigma,N}| \leq \sum_{j=1}^{\infty} \sum_{m=0}^{\infty} \eta^{\frac{j}{2}} \mu^m$ where j is the length of the polymer J containing n. Therefore $P_{n,\sigma,N}$ depends on the boundary condition vector v: but the dependence disappears in the limit $N \to \infty$. Hence the limits $P_{n,\sigma}$ of $P_{n,\sigma,N}$ as $N \to \infty$ exist and

$$|P(n,\sigma)| \leq \begin{cases} \frac{\eta^{\frac{1}{2}}}{(1-\mu)(1-\eta^{\frac{1}{2}})} & \text{if } \sigma \geq 1 \\ 1 + \frac{\eta^{\frac{1}{2}}}{(1-\mu)(1+\eta^{\frac{1}{2}})}(d-1) & \text{if } \sigma = 0 \end{cases} \tag{2.9}$$

where the bound for $\sigma = 0$ reflects the *algebraic* identity (due to the probabilistic interpretation) $P(n,0,N) + \sum_{\sigma=1}^{d-1} P(n,\sigma,N) \equiv 1$ (which holds whether or not the activities $I(J,\boldsymbol{\sigma}_J)$ are ≥ 0, provided convergence holds).

Hence the limit $h(n) \stackrel{def}{=} \lim_{N\to\infty} h(n,N)$ exists and by Eq.(1.3) is

$$\begin{aligned} h(n) &= \lim_{N\to\infty} \frac{T_n \dots T_N\, v}{\overline{\Lambda}_v(n,N)} = \sum_{\sigma=0}^{d-1} P_{n,\sigma} |n,\sigma\rangle, \qquad \text{with} \\ \overline{\Lambda}_v(n,N) &\stackrel{def}{=} \Omega(n,N) \prod_{j=n}^{N} \Big(\lambda_{j,0} \langle j,0|j+1,0\rangle\Big) \\ B^{-1} &\leq ||h(n)|| \leq B, \end{aligned} \tag{2.10}$$

with $B = (1 + 2d \frac{\eta^{\frac{1}{2}}}{(1-\mu)(1-\eta^{\frac{1}{2}})})$, if η is small enough so that $\mu, \eta < 1$ (together with Eq.(2.5) this is the condition on the spectral gap that determines the

[1] The site n must be contained since J must contain n because $\sigma_n \geq 1$; and the site N must be contained as otherwise the φ^T does not depend on v.

parameter γ). Notice that the convergence of the cluster expansion also implies $\Omega(n,N) > 0$ and upper and lower bounds on $\Omega(n,N)$:

$$\begin{aligned} |\log&\Omega(n,N)| \le |\sum_{\mathbf{Y}} \widehat{\varphi}^T(\mathbf{Y})| \le \sum_{\gamma\subset[n,N]} \sum_{\mathbf{Y}} |\widehat{\varphi}^T(\gamma\cup\mathbf{Y})| \\ &\le \sum_{\gamma\subset[n,N]} r(\gamma) \sum_{m=1}^{\infty} I_m \le \sum_{k=1}^{\infty} \frac{(N-n+1)}{1-\mu} \eta^{\frac{1}{2}k} \le \frac{(N-n+1)\sqrt{\eta}}{(1-\mu)(1-\sqrt{\eta})}, \end{aligned} \tag{2.11}$$

of course not uniform in N. Thus exhibiting the important cancellation that occurs in the ratios $\frac{\Omega_\sigma(n,N)}{\Omega(n,N)}$, above, thus estimated uniformly in N.

Defining

$$\begin{aligned} \Lambda(n,p) &\stackrel{def}{=} \lim_{N\to\infty} \frac{\Omega(p,N) \prod_{j=p}^{N} \Big(\lambda_{j,0} \langle j,0|j+1,0\rangle\Big)}{\Omega(n,N) \prod_{j=n}^{N} \Big(\lambda_{j,0} \langle j,0|j+1,0\rangle\Big)} \\ &= \Big(\prod_{j=n}^{p-1} \lambda_{j,0} \langle j,0|j+1,0\rangle\Big)^{-1} \lim_{N\to\infty} \frac{\Omega(p,N)}{\Omega(n,N)} \end{aligned} \tag{2.12}$$

Λ can be evaluated again by the cluster expansion, which also allows us to see the cancellation that shows the v-independence of the last limit, as:

$$\lim_{N\to\infty} \frac{\Omega(p,N)}{\Omega(n,N)} = \exp \sum_{\mathbf{Y}}^{*} \widehat{\varphi}^T(\mathbf{Y}) = \exp \sum_{q=n}^{p-1} \Phi(q) \tag{2.13}$$

where $\mathbf{Y} = (J_1, \boldsymbol{\sigma}_1, \dots, J_s, \boldsymbol{\sigma}_s)$ and the $*$ means that the cluster $(\cup J_i)$ overlaps with $[n,p]$ and $\Phi(q)$ is the sum $\sum_{\mathbf{Y}, J_1\ge q, J_1\ni q} \widehat{\varphi}^T(\mathbf{Y})$ over all polymer configurations which are to the right of q and q is the first point of J_1: in Eq.(2.13) numerator and denominator are exponentials of sums of many terms which are common (see Eq.(2.3)), hence cancel, except those relative to $\mathbf{Y}$'s with bases touching $[n,p]$. Such polymers are independent of v unless their bases touch also N: hence their contributions to $\Phi(q)$ tend to 0 as $N\to\infty$ at fixed n,p.

The positivity of Λ is due to the reality of $\widetilde{\varphi}^T$ and to the positivity of the scalar products and of $\lambda_{j,0}$ in Eq.(2.12) (by the cone property).

Remarks: (1) It is important to stress that $\Phi(q)$ depends only on the matrices T_j with $j \ge q$: $\Phi(q) = F(T_q, T_{q+1}, \dots)$ and each $\Phi(q)$ is given by an absolutely convergent series because of the bounds on $\widehat{\varphi}^T(\mathbf{Y})$ in Eq.(B.13),(B.14). Furthermore the function F *is independent of* q (*i.e.* it is translation invariant as a function of the sequence of matrices), because of the translation invariance of $\varphi^T(\mathbf{Y})$.

(2) The uniform convergence of the series defining $\Phi(q)$, Eq.(2.13), implies that all limits points as $p \to \infty$ of $\Omega(n,p)^{\frac{1}{p}}$ are n-independent.
(3) Uniformity of the limits defining $P_{n,\sigma}$ also holds if the matrix elements of the matrices T_j are varied keeping the cone property independent of the variations and their norms bounded by B_0: this is due to the uniformity of the bounds on $I(J,\boldsymbol{\sigma}_J)$ only depending on the inclination ε of Γ' in Γ, on the opening angles ϑ, ϑ' between the cones and on the bounds on the matrices norms.
(4) The activities $I(J,\boldsymbol{\sigma}_J)$, which so far have been real valued, are holomorphic functions of small complex variations of the matrices entries (since it is temporarily being supposed that the eigenvalues are real and keep pairwise a positive minimal distance) still satisfying the same bounds possibly with slightly different constants: hence $h(n), \Lambda(n,p)$ are analytic in the entries.
(5) The $h(n)$ form, essentially by definition once convergence has been established, a *covariant family* of vectors in the sense of theorem 1, Eq.(1.6), as implied by Eq.(2.10). In fact $h(n)$ and likewise $\Lambda(n,p), \Omega(n,N)$ are given by convergent expansions (see the series in Eq.(2.8), for instance) in the activities $I(J,\boldsymbol{\sigma}_J)$: the latter are given by simple algrebraic expressions, see Eq.(1.4), which are analytic in the matrix elements of the T_n.
(6) If the matrix elements of the T_j are chosen randomly with respect to a distribution ρ on $\prod_{j=-\infty}^{\infty} R^d \times R^d$ which is invariant under translations and has samples satisfying the (Γ, Γ')-cone property the limit

$$\lim_{p\to\infty} \frac{1}{p} \log \Lambda(n,p) = \lim_{p\to\infty} \frac{1}{p} \Big(\sum_{j=n}^{p} \log(\lambda_{j,0} \langle j,0|j+1,0\rangle) + \sum_{j=n}^{p-1} \Phi(j) \Big) \quad (2.14)$$

exists, consequence of the ergodic theorem, because the addends in the sums are translates to the right *of the same function of the random matrices* (*i.e.* in the language of ergodic theory they are "Birkhoff averages").
(7) Furthermore n-independence of the limit points, remarked in Eq.(2.14), implies that the limits in Eq.(2.14) are constants of motion under the right translations. Hence if ρ is also ergodic the limits are constant almost everywhere and the maximum Lyapunov exponent λ_+ is the integral with respect to ρ of $\log(\lambda_{1,0}\langle 1,0|2,0\rangle + \Phi(1)$: hence it is analytic in the parameters on which the matrices may depend.

Finally the *assumption that T_j have spectrum consisting of pairwise uniformly separated eigenvalues and have a large gap has to be removed*: lemma 2 shows that the matrices $T'_j = T_{jp+1} T_{jp+2} \cdots T_{(j+1)p}$ to which we want to

apply the analysis, see remark after Lemma 2, have a spectral representation as $T'_j = \lambda'_{0,j}\left(|j,0\rangle\langle j,0| + \sum_{\sigma,\sigma'\geq 1}(\Theta_j)_{\sigma,\sigma'}|j,\sigma\rangle\langle j,\sigma'|\right)$, where the bases $|j,\sigma\rangle$,$\langle j,\sigma'|$ consist of the vectors $|j,0\rangle$ and $\langle j,0|$ and correspondingly $d-1$ other orthogonal vectors, repectively to $|j,0\rangle$ and $\langle j,0|$, arbitrarily chosen. The matrix elements are bounded above by $c\alpha^p, \alpha < 1$.

Since the basis vectors $|j,\sigma\rangle, \langle j,\sigma|$ for $\sigma > 0$ need not be eigenvectors of T_j the basis can be taken real: hence the matrix Θ_j can be taken real.

Therefore it appears that this case simply leads to more complicated formulae in which at each site j are now associated *two spins* (σ_j, σ'_j), instead of just 1, and the products like $\prod_{j=h}^{h'} \frac{\lambda_{j,\sigma_j}}{\lambda_{0,j}} \frac{\langle j,\sigma_j|j+1,\sigma_{j+1}\rangle}{\langle j,0|j+1,0\rangle}$, appearing in Eq.(1.4) must be replaced by the product $\prod_{j=h}^{h'}(\Theta_j)_{\sigma_j,\sigma'_j} \frac{\langle j,\sigma_j|j+1,\sigma'_{j+1}\rangle}{\langle j,0|j+1,0\rangle}$.

This amounts at more values of the spins associated with each site: from $d-1$ to $(d-1)^2+1$. There is no need to perform a full spectral decompostion and therefore to worry about degeneracies, complex eigenvalues and eigenvectors, [2] eigenvalues crossing and the like. The only property needed is that the Θ_j be as small as necessary for the cluster expansion and to be analytic in the matrices entries: this is achieved, as mentioned, by taking p large enough.

Theorem 1 is thus proved aside from the analyticity statement. About analyticity we prove the stronger statement in theorem 2. Suppose first that the T_h depend on a parameter z as in the assumption of theorem 2 *and* the largest eigenvalue of T_h is separated, uniformly in h, by a large enough gap from the rest of the spectrum: then the property remains valid in a complex region within a distance $0 < \nu' \leq \nu$ of the real interval $z \in \Delta$.

The spectral decomposition yields that the largest eigenvalue $\lambda_{0,h}$ and the relative eigenvectors $\langle h,0|, |h,0\rangle$ are analytic in z and $|\langle h,0|h\pm 1\rangle|$ is bounded below by a positive δ and, finally the matrices $\Theta = \frac{1}{2\pi i}\oint \frac{\zeta d\zeta}{\zeta - T)}$ with $T = T_h$ and the integral on a contour surrounding the spectra of the T_h (or $T(x)$) but excluding $\lambda_{0,h}$ has also analytic matrix elements uniformly bounded.

Therefore the $I(J, \boldsymbol{\sigma}_J)$ are holomorphic in $|z| < \nu'$ and satisfy essentially the same bounds sufficient for the cluster expansion convergence, *i.e.* Eq.(2.5) with suitable δ, γ: and η is small for large spectral gap.

Finally when the spectral gap of the T_h is not small it is, nevertheless,

[2] *I.e* only the largest eigenvalue, which is separated from the rest of the spectrum by an h-independent factor < 1 (related to the α in lemma 1), and the relative eigenvector are needed.

smaller than a prefixed γ for the family of matrices $T'_h = T_h \cdots T_{h+q-1}$ if q is large enough: the analyticity holds, therefore, as above for the matrices T'_h and consequently for T_h.

This concludes a proof of the Theorems.

3 Comments

Neither the theorem in [2], nor its extensions in [8], have been used, the ergodic theorem being sufficient in the simple cases considered. The general and deep result in [2] does not give analyticity: for analyticity a more restricted class of matrices has to be considered, *e.g.* the class considered in [11] or here.

A simple application of theorem 1 is to matrices $(T_n)_{\sigma,\sigma'} > 0$ with $1 \leq \frac{\max_{\sigma,\sigma'}(T_n)_{\sigma,\sigma'}}{\min_{\sigma,\sigma'}(T_n)_{\sigma,\sigma'}} \leq C < \infty$: they have the cone property with Γ the cone of the vectors with components ≥ 0 and some, n-independent, cone Γ', [2, Sec.3]. This can be immediately applied to obtain free energy analyticity for a $1D$ spin glass with short range interaction as remarked in [11, p.69]: indeed the positivity of $(T_n)_{\sigma,\sigma'} > 0$ turns the problem into one in Statistical Mechanics with interaction $J_{\sigma\sigma'} = \log(T_n)_{\sigma,\sigma'}$, [10, p.121].

The analysis is fully constructive for what concerns the contents of Theorem 1. In fact lemma 1 can be replaced by the solution of finitely many eigenvalue problems, like the determination of the largest eigenvalue of the matrices $T_1, \ldots, T_{N_0}$ or their products with N_0 that can be computed *a priori*, if the approximation needed is given. N_0 is directly related to the maximum size of the polymers necessary to achieve a desired approximation: it is *a priori* determinable through the value of α appearing in lemma 2, the estimate Eq.(2.4) and the cluster expansion estimates.

Constructivity is only lost, as usual, in the application of the ergodic theorem, as there is no control on which is the set of matrices for which the limits like (2.14), exist.

Nevertheless in the case of sequences of matrices randomly chosen with respect to a τ-ergodic measure the determination of the *maximal Lyapunov exponent* (or minimal, depending on the cone property holding)) can be again expressed constructively, as the integral of the function appearing inside the sum in Eq.(2.14) setting $j = 1$, in which the first term is explictly known while the second, *i.e.* $\Phi(1)$, can be expressed to any prefixed accuracy by the cluster expansion.

The theorems are not optimal: for instance invertibility of the matrices

T_j, absent in [11, 8], is used only in item (3) of theorem 1: for the remaining statements it is not needed.

The point of this work has been to show how the cluster expansion technique can be of great help in problems that can be cast into a statistical mechanics context: after all it has been among the major achievements in equilibrium statistical mechanics of the XX century. Its use is limited to special problems but when applicable (as here) it gives a complete and constructive solution.

Appendices

A Algebraic properties of cones

Call ε, "inclination", the minimum angle between pairs of vectors in Γ and Γ'; ϑ, "opening", the maximum angle betwee pairs of vectors in Γ and ϑ' the maximum angle between vectors in Γ': $\pi > \vartheta > \vartheta' > 0, \varepsilon > 0$.
proof of Lemma 1: (following [11]) Let T be a $d \times d$ matrix and let Γ, Γ' be proper, convex, closed cones (with apex at the origin O) in R^d. Suppose that $T\,\Gamma \subset \Gamma'$ and a relative inclination $\varepsilon > 0$ of Γ to Γ'.

Let $\Gamma^* = \{w | \langle w|v\rangle \geq 0, \forall v \in \Gamma\}$. Then, fixed $0 \neq v_0 \in \Gamma$ and $0 \neq w_0 \in \Gamma^*$ the maps

$$v \to \frac{Tv}{\langle T^* w_0 | v\rangle}, \quad \text{and} \quad w \to \frac{T^* w}{\langle w | T v_0\rangle} \tag{A.1}$$

map continuously the convex compact sets $\{v|\langle w_0|v\rangle = 1\}$ and, respectively, $\{w|\langle w|v_0\rangle = 1\}$ strictly into themselves. Hence $\exists\ a \in \Gamma$ and $a^* \in \Gamma^*$ which are fixed points of the maps, respectively; hence, if $b \stackrel{def}{=} \frac{a}{||a||}$ and $b^* \stackrel{def}{=} \frac{a^*}{||a^*||}$

$$\frac{Tb}{||Tb||} = b, \qquad \frac{T^* b^*}{||T^* b^*||} = b^* \tag{A.2}$$

Let $\overline{T}\xi \stackrel{def}{=} \frac{T\xi}{||Tb||}$ and let

$$K \stackrel{def}{=} \{\xi | \langle b^*|\xi\rangle = 0, \text{ and } b + \xi \in \Gamma\} \tag{A.3}$$

Since $\frac{Tb}{||Tb||} = b$ and $\frac{T^* b^*}{||T^* b^*||} = b^*$ the set K is mapped into itself by $\overline{T}$ (*e.g.* $\langle b^*|T\xi\rangle = 0$ and $b + \frac{T\xi}{||Tb||} = \frac{T(b+\xi)}{||Tb||} \in \Gamma$) and since the cone Γ is shrunk by T the set K is mapped into $\overline{T}K \subset \alpha K$ with $\alpha < 1$ (determined by the inclination and opening angles, see Sec.2).

Hence $\overline{T}^n(b+\xi) = b + \overline{T}^n\xi$ and $||\overline{T}^n\xi|| \leq \alpha^n$. For any $v \in \Gamma, v \neq 0$, there is a $\nu \neq 0$ such that $v = \nu b + \xi$, $\xi \in K$, so that

$$\frac{T^n(\nu b + \xi)}{||T^n(\nu b + \xi)||} \equiv \frac{\overline{T}^n(\nu b + \xi)}{||\overline{T}^n(\nu b + \xi)||} \equiv \frac{(\nu b + O(\alpha^n))}{||\nu b + O(\alpha^n)||} \xrightarrow[n\to+\infty]{} \lambda b \tag{A.4}$$

because T and $\overline{T}$ are proportional: notice that the above analysis implies that the largest eigenvalue λ_0 of T is positive and that it is the unique eigenvalue of T with maximum modulus.

proof of Lemma 2: (following [11]) Let $T' \stackrel{def}{=} T_1 T_2 \cdots T_p$ and let $\boldsymbol{\vartheta}, \boldsymbol{\vartheta}^*$ be the, respective, normalized eigenvectors with maximum modulus eigenvalue $\lambda' > 0$ for T' and $(T')^*$ (existing by lemma 1). Define

$$\begin{aligned} &\boldsymbol{\vartheta}_p \stackrel{def}{=} \boldsymbol{\vartheta},\ \boldsymbol{\vartheta}_{p-1} = \frac{T_p \boldsymbol{\vartheta}_p}{||T_p \boldsymbol{\vartheta}_p||}, \ldots, \boldsymbol{\vartheta}_0 = \frac{T_1 \boldsymbol{\vartheta}_1}{||T_1 \boldsymbol{\vartheta}_1||} = \boldsymbol{\vartheta} \\ &\boldsymbol{\vartheta}_0^* \stackrel{def}{=} \boldsymbol{\vartheta}^*,\ \boldsymbol{\vartheta}_1^* = \frac{T_1^* \boldsymbol{\vartheta}_0^*}{||T_1^* \boldsymbol{\vartheta}_0^*||}, \ldots, \boldsymbol{\vartheta}_p^* = \frac{T_p^* \boldsymbol{\vartheta}_{p-1}^*}{||T_p^* \boldsymbol{\vartheta}_{p-1}^*||} = \boldsymbol{\vartheta}^* \end{aligned} \tag{A.5}$$

By the argument in the proof of lemma 1 the action of T_j on the plane orthogonal to $\boldsymbol{\vartheta}_j^*$ maps it on the plane $\boldsymbol{\vartheta}_{j+1}^*$ and contracts by at least $\alpha < 1$. Therefore $T_1 T_2 \cdots T_p$ contracts by at least α^p in the space orthogonal to $\boldsymbol{\vartheta}^*$, proving Lemma 2.

B Cluster expansion: a rehearsal

This section follows [5, Ch.7] (in turn based on [9, 6]) and it is here only for the purpose of making the paper self-contained for the reader. Cluster expansion is an algorithm to compute the logarithm of a sum

$$\Xi = \sum_{\mathbf{J}}^{*} \zeta(\mathbf{J}) \equiv \sum_{\mathbf{J}}^{*} \prod_{i} \zeta(J_i)^{n_i} \tag{B.1}$$

where: (1) $\mathbf{J} = (J_1^{n_1}, \ldots, J_{\mathcal{N}}^{n_{\mathcal{N}}})$ with J_i's subsets in a box Λ on a d-dimensional lattice (here $d = 1$) called *polymers* and $n_i \geq 0$ are integers defining the "multiplicity" of each (or "counting" how many times each set is counted) hence $\mathcal{N} = 2^{|\Lambda|}$. The sets J could be decorated by associating to each site $k \in J$ a "spin", *i.e.* a variable assuming $d-1$ values. However in the following the decorations will not be mentioned as they would only make the notations heavier. In the applications in Sec.2 the decorations will be necessary and the formulae of this section (which correspond to the case $d = 2$, *i.e.* all spins 1) are directly usable simply by imagining that each J is in fact a pair $Y = (J, \boldsymbol{\sigma}_J)$ where $\boldsymbol{\sigma}_J = (\sigma_j)_{j \in J}$ and $\sigma_j = 1, \ldots, d-1$.
(2) $\zeta(\mathbf{J}) = \prod \zeta(J_i)^{n_i}$ with $\zeta(J)$ (small) constants called *activities*, $\zeta(\emptyset) \stackrel{def}{=} 1$.
(3) the $*$ means that the sum runs over the $\mathbf{J}$'s in which no two of the $J_i \in \mathbf{J}$ with multiplicity $n_i > 0$ overlap in the sense that they contain pairs of points

at distance ≤ 1 on the lattice. If $\widetilde{\mathbf{J}}$ denotes the sets in $\mathbf{J}$ which have positive multiplicity then the $*$ indicates that the sum is restricted to $\mathbf{J} \equiv \widetilde{\mathbf{J}}$ in which no two of the J's intersect.

In applications $\zeta(J) \neq 0$ only for a few of the possible subsets of Λ. For instance in the present case Λ is the interval $[n, N]$ and the "polymers" are just the subintervals.

The Ξ can certainly be written as $\exp(\sum_{\mathbf{J}} \varphi^T(\mathbf{J})\zeta(\mathbf{J}))$ by expanding the $\log \Xi$ in powers of the $\zeta(J)$: of course the sum in the exponential will involve $\mathbf{J}$ with J's which can overlap or that can be counted many times. The $\varphi^T(\mathbf{J})$ are suitable combinatorial coefficients.

For instance if Λ is just one point $\Xi = 1 + z$ can be written as the exponential of $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} z^k$. If Λ consists of two points, say 1 and 2 then the polymers are $\emptyset, 1, 2, 12$ and $\Xi = 1 + z_1 + z_2 + z_{12}$ is the exponential of $\sum_{k_1+k_2+k_3>0} \frac{(-1)^{k_1+k_2+k_3+1}(k_1+k_2+k_3-1)!}{k_1!k_2!k_3!} z_1^{k_1} z_2^{k_2} z_{12}^{k_3}$.

The cluster expansion is the general form of the above examples. It is of interest, for instance, if $\sum_{\mathbf{J}}^{\&} |\varphi^T(\mathbf{J})||\zeta(\mathbf{J})| < +\infty$ where the & means that the sum is restricted to $\mathbf{J}$'s which contain any fixed point $x \in \Lambda$ (*i.e.* with $x \in \cup_{J \in \widetilde{\mathbf{J}}} J$). It is therefore necessary to determine conditions that imply the mentioned convergence.

The first step is to define $\mathbf{J}+\mathbf{J}'$ simply as $J_1^{n_1+n_1'}, \ldots, J_{\mathcal{N}}^{n_{\mathcal{N}}+n_{\mathcal{N}}'}$, *i.e.* as the family of polymers with multiplicities equal to the sum of the corresponding ones in $\mathbf{J}$ and $\mathbf{J}'$. Let

$$\begin{aligned} \mathcal{F} =& \text{set of functions } F(\mathbf{J}) \\ \mathcal{F}_0 =& \text{set of functions } F(\mathbf{J}) \text{ with } F(\emptyset) = 0 \\ \mathcal{F}_1 =& \text{set of functions } F(\mathbf{J}) \text{ with } F(\emptyset) = 1 \\ \mathbf{1}(\mathbf{J}) =& \begin{cases} 0 & \text{if } \mathbf{J} \neq \emptyset \\ 1 & \text{if } \mathbf{J} = \emptyset \end{cases} \\ f \in& \mathcal{F}_1 \longleftrightarrow \widetilde{f} \stackrel{def}{=} f - \mathbf{1} \in \mathcal{F}_0 \end{aligned} \tag{B.2}$$

and remark that $f \in \mathcal{F}_1$ can be written $f = \mathbf{1} + \widetilde{f}$ with $\widetilde{f} \in \mathcal{F}_0$.

Then if $f * g(\mathbf{J}) \stackrel{def}{=} \sum_{\mathbf{J}_1+\mathbf{J}_2=\mathbf{J}} f(\mathbf{J}_1)g(\mathbf{J}_2)$, for $f, g \in \mathcal{F}$ define

$$\begin{aligned} &\mathrm{Exp} f(\mathbf{J}) = \sum_{k=0}^{\infty} \frac{f^{*k}(\mathbf{J})}{k!}, \quad f \in \mathcal{F}_0 \\ &\mathrm{Log} f(\mathbf{J}) = \sum_{k=1}^{\infty} \frac{(-1)^k \widetilde{f}^{*k}(\mathbf{J})}{k}, \quad f = \mathbf{1} + \widetilde{f} \in \mathcal{F}_1 \\ &f^{*-1} = \sum_{k=1}^{\infty} (-1)^k \widetilde{f}^{*k}, \qquad f = \mathbf{1} + \widetilde{f} \in \mathcal{F}_1 \\ &\langle\, f \,\rangle = \sum_{\mathbf{J} \subset \Lambda} f(\mathbf{J}), \quad f \in \mathcal{F} \end{aligned} \tag{B.3}$$

here all sums over k are finite sums for f in the corresponding domains.

A key remark is

$$\begin{aligned} &\mathrm{Log}\,(\mathrm{Exp}(f)) = f \qquad \forall\ f \in \mathcal{F}_0, \qquad \mathrm{Exp}\,(\mathrm{Log}(f)) = f \qquad \forall\ f \in \mathcal{F}_1 \\ &f^{*-1} * f = \mathbf{1}, \quad \forall\ f = \mathbf{1} + \widetilde{f} \in \mathcal{F}_1, \qquad \langle\, f * g \,\rangle = \langle\, f \,\rangle \langle\, g \,\rangle \end{aligned} \tag{B.4}$$

If $\chi(\mathbf{J}) = \prod \overline{\chi}(J_i)^{n_i}$ is a *multiplitive function* $\chi \in \mathcal{F}$ then $\langle\, f * g\chi \,\rangle = \langle\, f\chi \,\rangle \langle\, g\chi \,\rangle$ so that if $\varphi \in \mathcal{F}_1$ and $\overline{\chi}(J) = \zeta(J)$

$$\langle\, f \cdot \zeta \,\rangle = \langle\, \mathrm{Exp}(\mathrm{Log}(\mathrm{f} \cdot \zeta) \,\rangle = \exp \langle\, (\mathrm{Log} f \cdot \zeta) \,\rangle \tag{B.5}$$

Therefore call $\mathbf{J}$ compatible if $n_i = 0, 1$ (*i.e.* $\mathbf{J} = \widetilde{\mathbf{J}}$) and the elements of $\widetilde{\mathbf{J}}$ are not connected then if

$$\varphi(\mathbf{J}) = \begin{cases} 0 & \text{if } \mathbf{J} \text{ is not compatible} \\ 1 & \text{otherwise} \end{cases} \tag{B.6}$$

then $\varphi \in \mathcal{F}_1$ and $\varphi^T = \mathrm{Log}\varphi \in \mathcal{F}_0$ makes sense and

$$\Xi = \langle\, \varphi \cdot \zeta \,\rangle = \exp \langle\, \varphi^T \cdot \zeta \,\rangle \equiv \exp \sum_{\mathbf{J}} \varphi^T(\mathbf{J})\, \zeta(\mathbf{J}) \tag{B.7}$$

which is the exponential of a power series in the $\zeta(J)$ variables.

Calculating $\varphi^T(\mathbf{J})$ requires computing the sum of finitely many quantities: if $\mathbf{J}$ is represented as a set of "points" or nodes" and if G is the graph obtained by joining all pairs of polymers in $\mathbf{J}$ which are "incompatible" (regarding as different, and incompatible with each other, the n_i copies of J_i) it is, (*e.g.* see [6, Eq.(4.21)]),

$$\varphi^T(\mathbf{J}) = \frac{1}{\prod n_i!} \sum_{C \subset G}^{*} (-1)^{\#\ of\ edges\ in\ C} \tag{B.8}$$

where the $*$ means that the sum is restricted to the subgraphs of G which visit all polymers in G: their number is huge, growing faster than any power in the number of polymers so that convergence occurs because of cancellations due to the relation in Eq.(B.12).

The series in Eq.(B.7) is certainly convergent for $\zeta(J)$'s small enough: *however the radius of convergence might be very small and* Λ *dependent.*

Define the *differentiation operation* as

$$(D_\Gamma \Psi)(\mathbf{H}) \stackrel{def}{=} \Psi(\Gamma + \mathbf{H}) \frac{(\Gamma + \mathbf{H})!}{\mathbf{H}!} \tag{B.9}$$

with $\Gamma! = \prod_{i=1}^{s} n_i!$. The name is attributed because of the validity of the following rules:

$$\begin{aligned} &D_\gamma(\Psi_1 * \Psi_2) = (D_\gamma \Psi_1) * \Psi_2 + \Psi_1 * (D_\gamma \Psi_2), \\ &D_\gamma \mathrm{Exp}\Psi = D_\gamma \Psi * \mathrm{Exp}\Psi, \end{aligned} \tag{B.10}$$

A direct check of the above relations can be reduced to the case in which $\Gamma = n\gamma$, *i.e.* to the case in which there is only one polymer species γ, and the check is left to the reader. The first relation above, *Leibniz rule*, can be seen as a consequence the combinatorial identity $\sum_{p_1+p_2=n} \binom{q_1}{p_1}\binom{q_2}{p_2} = \binom{q_1+q_2}{n}$ for all n, q_1, q_2 with $n \le q_1 + q_2$.

The definitions lead to the derivation of the expression for $\varphi^T(\Gamma)$ in (B.8): which not only is quite explicit but also implies immediately that $\varphi^T(\Gamma)$ vanishes for nonconnected Γ's.

To determine sufficient conditions for the convergence which are independent on the size of Λ let $\widehat{\varphi}(\mathbf{Y}) \stackrel{def}{=} \varphi(\mathbf{Y})\zeta(\mathbf{Y})$ and $\Delta_{\mathbf{J}}(\mathbf{Y}) \stackrel{def}{=} \widehat{\varphi}^{*-1} * D_{\mathbf{J}}\widehat{\varphi})(\mathbf{Y})$. Then if γ is a polymer, and $\mathbf{J}, \mathbf{Y}$ are polymer configurations

$$\begin{aligned} \Delta_{\gamma+\mathbf{J}}(\mathbf{Y}) =& \sum_{\mathbf{Y}_1+\mathbf{Y}_2=\mathbf{Y}} \widehat{\varphi}^{*-1} * (\mathbf{Y}_1)\varphi(\gamma + \mathbf{J} + \mathbf{Y}_2)\zeta(\gamma + \mathbf{J} + \mathbf{Y}_2) \\ =& \zeta(\gamma) \sum_{\mathbf{Y}_1+\mathbf{Y}_2=\mathbf{Y}} \widehat{\varphi}^{*-1}(\mathbf{Y}_1) * \varphi(\gamma + \mathbf{J} + \mathbf{Y}_2)\zeta(\mathbf{J} + \mathbf{Y}_2) \end{aligned} \tag{B.11}$$

Here no factorials appear because $\varphi(\mathbf{J})$ vanishes unless $\mathbf{J} = \widetilde{\mathbf{J}}$.

Remark that $\varphi(\gamma + \mathbf{J} + \mathbf{Y}_2) = \varphi(\mathbf{J} + \mathbf{Y}_2) \prod_{\gamma' \in \mathbf{Y}_2} (1 + \chi(\gamma, \gamma'))$ with $\chi(\gamma, \gamma') = 0$ if γ, γ' do not overlap and $\chi(\gamma, \gamma') = -1$ otherwise, so that $\varphi(\gamma + \mathbf{J} + \mathbf{Y}_2) = \varphi(\mathbf{J} + \mathbf{Y}_2) \sum^*_{\mathbf{S} \subset \mathbf{Y}_2} (-1)^{|\mathbf{S}|}$, with $|\mathbf{S}|$ = number of polymers in $\mathbf{S} = (s_1, s_2, \ldots)$ and $*$ means that the s_i overlap with γ, for all i. Hence

setting $\mathbf{Y}_2 = \mathbf{S} + \mathbf{H}$

$$\begin{aligned}\Delta_{\gamma+\mathbf{J}}(\mathbf{Y}) =& \zeta(\gamma) \sum_{\mathbf{S}\subset\mathbf{Y}}^{*} \sum_{\mathbf{Y}_1+\mathbf{H}=\mathbf{Y}-\mathbf{S}} \widehat{\varphi}^{-1}(\mathbf{Y}_1)\varphi(\mathbf{J}+\mathbf{S}+\mathbf{H})\zeta(\mathbf{J}+\mathbf{S}+\mathbf{H}) \\ =& \zeta(\gamma) \sum_{\mathbf{S}\subset\mathbf{Y}}^{*} (-1)^{|\mathbf{S}|}\Delta_{\mathbf{J}+\mathbf{S}}(\mathbf{Y}-\mathbf{S})\end{aligned} \tag{B.12}$$

Let $r(\gamma) \geq |\zeta(\gamma)|$ and $r(\mathbf{X}) = \prod_{\gamma\in\mathbf{X}} r(\gamma)$; then

$$I_m \stackrel{def}{=} \sup_{1\leq n\leq m} \sup_{|\mathbf{J}|=n} \sum_{\mathbf{Y},|\mathbf{Y}|=m-n} \frac{|\Delta_{\mathbf{J}}(\mathbf{Y})|}{r(\mathbf{J})} \tag{B.13}$$

and I_1 is then $I_1 = \sup \frac{|\zeta(\gamma)|}{r(\gamma)}$ and recursively $I_{m+1} \leq \mu^m I_1$ where

$$\mu \stackrel{def}{=} \sup_{\gamma} \frac{|\zeta(\gamma)|}{r(\gamma)} \exp \sum_{J}^{*} r(J) \tag{B.14}$$

where here J is a single polymer (intersecting γ): see [5, Eq. 7.1.28] for more details on the algebra. Therefore $I_{m+1} \leq \mu^m I_1$, if $\mu < 1$.

The latter property $\mu < 1$ holds in various applications, notably in the present work, to bound $\Omega(n, N)$ as well as a few more quantities.

The method has several other applications, see [6], [5, Ch.7]. Here the polymers J will be $\boldsymbol{\sigma}_J$ corresponding to intervals J (on the lattice $[1, N]$) with the associated spin structures $\boldsymbol{\sigma}_J$. We shall make use of Eq.(B.7) and, by Eq.(B.7) and the third in (B.10), of

$$\begin{aligned}P(J) \stackrel{def}{=} \frac{\sum_{\mathbf{H}\ni J} \zeta(\mathbf{H})}{\Xi} = \frac{\langle\, D_J\varphi\zeta \,\rangle}{\langle\, \varphi\zeta \,\rangle} = \langle\, \widehat{\varphi}^{*-1} * D_J\widehat{\varphi} \,\rangle \\ = \langle\, \widehat{\varphi}^{*-1} * D_J Exp(\widehat{\varphi}^T) \,\rangle = \langle\, D_J\widehat{\varphi}^T \,\rangle, \qquad \widehat{\varphi} \equiv \varphi\zeta\end{aligned} \tag{B.15}$$

In an ensemble in which the polymer configurations $\mathbf{J}$ in Λ are given a weight proportional to $\prod_{\gamma\in\mathbf{J}} \zeta(\gamma)$ this would be the probability of finding a configuration of polymers containing the polymer J if $\zeta(\gamma) \geq 0$. Hence the complementary sum $P'(J) \stackrel{def}{=} \frac{\sum_{\mathbf{H}\in J} \zeta(\mathbf{H})}{\Xi}$ will be such that $P(J) + P'(J) = 1$.

C Dynamical systems application

Let $\mathcal{F}$ be a smooth compact manifold and τ a smooth, smoothly invertible, map on $\mathcal{F}$ (take smooth to mean C^∞, for simplicity). At each point $x \in \mathcal{F}$

there are proper closed convex cones $\Gamma(x) \supset \Gamma'(x)$, with apex at x in a linear space $E(x)$ of dimension d smoothly dependent on x (and call its adjoint $E(x)^*$). The cones are also supposed to depend smoothly on x.

Definition: *The minimum angle between vectors on the boundary of* $\Gamma(x)$ *and on that of* $\Gamma'(x)$ *will be called inclination* $\varepsilon(x)$*; while the maximum angle between vectors in* $\Gamma(x)$ *will be called* $\vartheta(x)$*, likewise define* $\vartheta'(x)$*.*

Let $T(x)$, $x \in \mathcal{F}$, be an invertible mapping of $E(x)$ onto $E(\tau x)$, and $T(x)$ maps $\Gamma(x)$ into $\Gamma'(\tau x) \subset \Gamma(\tau x)$ with $\Gamma'(x)/\{x\} \subset \Gamma(x)^0$ and $T^{\pm}(x), \Gamma(x), \Gamma'(x), \varepsilon(x), \vartheta(x), \vartheta'(x)$ be smooth, $\pi > \vartheta(x) > \varepsilon(x), \vartheta'(x) > 0$.

Making use of Lemma 1, 2 in Appendix A it will not be restrictive to suppose that $T(x)$ is "almost diagonalizable" in the sense that there exist $\lambda_0(x)$, $|x,0\rangle \in E(\tau x), \langle x,0| \in E(x)^*$ smoothly dependent on x and $\Theta(x)$ with norm such that $||\Theta(x)||/\lambda_0(x)$ is smaller than a prefixed quantity (x-uniformly):

$$T(x) = \lambda_0(x)|x,0\rangle\langle x,0| + \Theta(x) \tag{C.1}$$

Then setting $T_h \stackrel{def}{=} T(\tau^{-h}x)$ and repeating the proof of theorem 1 leads to

Theorem 3: *Let* $T(x)$ *be as above. Let* $x \to v(x) \in \Gamma'(x), ||v(x)|| \equiv 1$ *be a measurable functio (not necessarily continuous), it is*
(1) There are continuous functions $x \to b(x) \in \Gamma(x)$, $x \to \overline{\Lambda}(x,p)$ *and* $x \to \Lambda(x,p), p = 0,1,\ldots$, *such that, for all* $p > 0, x \in \mathcal{F}, v$, *exist the limits*

$$\begin{aligned} b(x) &= \lim_{N\to\infty} \frac{T(x)\cdots T(\tau^{-(N-1)}x)v(\tau^{-(N-1)}x)}{\overline{\Lambda}_v(x,N)} \\ \Lambda(x,p) &= \lim_{N\to\infty} \frac{\overline{\Lambda}_v(\tau^{-p}x,N)}{\overline{\Lambda}_v(x,N)} > 0, \end{aligned} \tag{C.2}$$

(2) The vectors $b(x)$ *are eigenvectors for products of* $T(\tau^{-j}x)$ *in the sense*

$$b(\tau^{-p}x) = \Lambda(x,p)\, T^{-1}(\tau^{-(p-1)}x)\ldots T^{-1}(x)\, b(x) \tag{C.3}$$

(3) $b(x), \Lambda(x,p)$ *are* v*-independent and continuous in* $x \in \mathcal{F}$ *and* $\exists B$ *such that* $B^{-1} < ||b(x)|| < B$*; if* $T(x)$ *is the Jacobian of* τ^{-1} *the unit vector* $\frac{b(x)}{||b(x)||}$ *will be called the unstable unit vector, or unstable direction, at* x*.*
(4) The upper and lower limit values $\ell^{\pm}(x)$ *of* $\frac{1}{p}\log\Lambda(x,p)$ *as* $p \to \infty$ *are constant along trajectories, i.e.* k*-independent if evaluated at* $\tau^{-k}x$*.*

The continuity is an extra property due to the continuity of the tems appearing in the cluster expansion. Analiticity of $b(x), \Lambda(x,p)$ can be obtained as in the case of theorem 2 under natural analogous assumptions.

If x is chosen randomly with respect to an invariant measure ρ then the limits in item (4) are a.e. equal (as in the case of Sec.2: via the cluster expansion, they are represented as "Birhoff averages". If ρ is ergodic the limits not only exist but are x-independent a.e. and $b(x)$ identifies the unstable direction at x while $\ell = \ell^+ = \ell^-$ is the maximum Lyapunov exponent.

References

[1] M.E. Fisher. Theory of condensation and the critical point. *Physics Physica Fyzika*, 3:255–283, 1967.

[2] H. Furstenberg and H. Kesten. Products of random matrices. *Annals of Mathematical Statistics*, 31:457–469, 1960.

[3] G. Gallavotti. Zeta functions and basic sets. *Rendiconti Accademia dei Lincei*, LXI:309–317, 1976 (Italian)
(English: http://ipparco.roma1.infn.it/pagine/1967-1979).

[4] G. Gallavotti. *Statistical Mechanics. A short treatise*. Springer Verlag, Berlin, 2000.

[5] G. Gallavotti, F. Bonetto, and G. Gentile. *Aspects of the ergodic, qualitative and statistical theory of motion*. Springer Verlag, Berlin, 2004.

[6] G. Gallavotti, A. Martin-Löf, and S. Miracle-Solé. Some problems connected with the phase separation in the Ising model at low temperature. *in Lecture Notes in Physics (ed. A. Lenard)*, 20:162–204, 1973.

[7] C. Liverani. Decay of correlations. *Annals of Mathematics*, 142:239–301, 1995.

[8] M. S. Raghunathan. A proof of Oseledec's multiplicative ergodic theorem. *Israel Journal of Mathematics*, 32:356–362, 1979.

[9] D. Ruelle. Cluster property of the correlation functions of classical gases. *Reviews of Modern Physics*, 36:580–584, 1964.

[10] D. Ruelle. *Thermodynamic formalism*. Addison Wesley, Reading, 1978.

[11] D. Ruelle. Analyticity properties of the characteristic exponents of random matrix products. *Advances in Mathematics*, 32:68–80, 1979.