CHAPTER VI

General properties of Gibbs and SRB distributions

§6.1 Variational properties of Gibbs distributions

The study of Gibbs distributions can be performed to remarkable depth as we shall hint in this section and in the forthcoming ones. We begin with a structure theorem which can be articulated into various propositions.

P6.1.1 **(6.1.1) Proposition:** (Thermodynamic limit)
Let T be a mixing $(n+1)\times(n+1)$ compatibility matrix and Φ a potential in the space B of the potentials for $\{0,\ldots,n\}_T^{\mathbb{Z}}$, cf. *definition (5.1.1). Define the* partition function *of Φ in the box $\Lambda_N=[1,N]$ as*

e6.1.1

$$Z_N(\Phi)=\sum_{\underline{\sigma}\in\{0,\ldots,n\}_T^{\Lambda_N}}\exp\Big[-\Big(\sum_{R\subset\Lambda_N}\Phi_R(\underline{\sigma}_R)\Big)\Big]. \tag{6.1.1}$$

Then the limit

e6.1.2

$$P(\Phi)=\lim_{N\to\infty}N^{-1}\log Z_N(\Phi) \tag{6.1.2}$$

exists and it defines a function $\Phi\to P(\Phi)$ on B which is convex and Lipshitz continuous:

e6.1.3

$$|P(\Phi)-P(\Psi)|\le\|\Phi-\Psi\|. \tag{6.1.3}$$

The function $P(\Phi)$ is called the pressure *of the potential Φ.*

N6.1.1 *Proof:* Setting[1] $U_\Lambda^\Phi(\underline{\sigma})=\sum_{R\subset\Lambda}\Phi_R(\underline{\sigma}_R)$, cf. (5.1.11) in definition (5.1.3),

---

[1] With the notations used in definition (5.1.3) we should write $U_\Lambda^{0,\Phi}(\sigma)$ instead of $U_\Lambda^{\Phi}(\sigma)$: we prefer to drop the label 0 throughout all this chapter in order not to use an overwhelming notation.

one has

e6.1.4

$$\begin{aligned}&|N^{-1}\log Z_N(\Phi)-N^{-1}\log Z_N(\Psi)|=\\&=\Big|N^{-1}\int_0^1 dt\,\frac{d}{dt}\,\log Z_N(\Psi+t(\Phi-\Psi))\Big|=\\&=\Big|N^{-1}\int_0^1 dt\frac{\sum_{\underline{\sigma}\in\{0,\ldots,n\}_T^{\Lambda_N}}U_{\Lambda_N}^{\Phi-\Psi}(\underline{\sigma})\,e^{-U_{\Lambda_N}^{\Psi+t(\Phi-\Psi)}(\underline{\sigma})}}{\sum_{\underline{\sigma}\in\{0,\ldots,n\}_T^{\Lambda_N}}e^{-U_{\Lambda_N}^{\Psi+t(\Phi-\Psi)}(\underline{\sigma})}}\Big|\le\|\Phi-\Psi\|,\end{aligned}\tag{6.1.4}$$

having used remark (2) to definition (5.1.3), *i.e.*

e6.1.5

$$|U_{\Lambda_N}^{\Phi}(\underline{\sigma})|\le\|\Phi\|N\qquad\text{for all }\Phi\in B.\tag{6.1.5}$$

Equation (6.1.1) and (6.1.5) immediately imply

e6.1.6

$$-\|\Phi\|+\log(n+1)\le N^{-1}\log Z_N(\Phi)\le\|\Phi\|+\log(n+1),\tag{6.1.6}$$

and therefore (6.1.4) and (6.1.6) imply that in order to show existence of the limit $P(\Phi)$ it suffices to consider $\Phi\in B^0$ where B^0 is an arbitrary set of potentials dense in B. Naturally, we shall select the set B^0 of the potentials Φ with finite range $R_\Phi<\infty$, cf. remark (4) to definition (5.1.3).

For simplicity we consider only the case $T_{\sigma\sigma'}=1\ \forall\sigma,\sigma'$.

If the finite range of Φ is $R_\Phi\equiv R$ and if $\underline{\sigma}=(\sigma'_1\ldots\sigma'_N\sigma''_1\ldots\sigma''_M)=(\underline{\sigma}'\underline{\sigma}'')\in\{0,\ldots,n\}^{N+M}$, we get

e6.1.7

$$|U_{\Lambda_{N+M}}^{\Phi}(\underline{\sigma}'\underline{\sigma}'')-U_{\Lambda_N}^{\Phi}(\underline{\sigma}')-U_{\Lambda_M}^{\Phi}(\underline{\sigma}'')|\le R\|\Phi\|,\tag{6.1.7}$$

by using (5.1.18), (5.2.1) and (5.2.3). So that, by (6.1.7) and (6.1.1),

e6.1.8

$$Z_{N+M}(\Phi)=Z_N(\Phi)Z_M(\Phi)e^{oR\|\Phi\|},\tag{6.1.8}$$

with $o\in[-1,1]$. If $N=M=2^k$ and $P_k=2^{-k}\log Z_{2^k}(\Phi)$, (6.1.8) says that

e6.1.9

$$P_k-2^{-(k+1)}R\|\Phi\|\le P_{k+1}\le P_k+2^{-(k+1)}R\|\Phi\|,\tag{6.1.9}$$

hence (6.1.9) imply that the limit as $k\to\infty$ of P_k exists.

Let $N=m2^k+d$, $0\le d<2^k$, $0\le m$; equation (6.1.8), repeatedly applied ($m+1$ times) yields

e6.1.10

$$Z_N(\Phi)=(Z_{2^k}(\Phi))^m Z_d(\Phi)\,e^{o\,(m+1)R\|\Phi\|},\tag{6.1.10}$$

with $o\in[-1,1]$. Then, at fixed k,

e6.1.11

$$\begin{aligned}N^{-1}\log Z_N(\Phi)=\frac{m2^k}{m2^k+d}\,P_k+\frac{m+1}{m2^k+d}\,oR\|\Phi\|\quad+&\\+\frac{1}{m2^k+d}\,\log Z_d(\Phi)\xrightarrow[m\to\infty]{}P_k+2^{-k}oR\|\Phi\|.&\end{aligned}\tag{6.1.11}$$

This shows, by the arbitrariness of k, the existence of the limit (6.1.2) and that it coincides with $\lim_{k\to\infty} P_k$.

Convexity is evident for $N^{-1}\log Z_N(\Phi)$ and hence for $P(\Phi)$. Lipschitz-continuity follows from (6.1.4) and from the existence of the limit (6.1.2). ■

P6.1.2 **(6.1.2) Proposition:** (Tangent plane to the graph of the pressure) *Under the hypotheses of the preceding proposition and if $\mu \in G(\Phi)$, the functional*

e6.1.12
$$\Psi \to \alpha(\Psi) = -\mu(A_\Psi) \tag{6.1.12}$$

is "tangent" to the graph of $\Phi \to P(\Phi)$. This means

e6.1.13
$$P(\Phi+\Psi) \geq P(\Phi) + \alpha(\Psi) \qquad \textit{for all } \Psi \in B. \tag{6.1.13}$$

Vice versa every functional α, tangent to the graph of the function $P(\Phi)$, has the form (6.1.12) with $\mu \in G(\Phi)$ and μ is uniquely determined by α.

Remark: The Gibbs states $\mu \in G(\Phi)$ have therefore the geometric interpretation of "tangent planes" to the graph of $\Phi \to P(\Phi)$. The function $P(\Phi)$ takes the role of *generating function* of the Gibbs states, via the relation

e6.1.14
$$\mu(A_\Psi) = -\frac{d}{d\varepsilon}\, P(\Phi+\varepsilon\Psi)\Big|_{\varepsilon=0}, \tag{6.1.14}$$

that is well defined, by the convexity of the function $\varepsilon \to P(\Phi+\varepsilon\Psi)$, at least where $G(\Phi)$ consists of a single point (and, hence, $\Phi \to P(\Phi)$ has a unique tangent plane).

This explains the importance attributed to the functional P which, because of its meaning in various problems of Statistical Mechanics, is called *pressure.*

Proof: As usual we shall suppose, for simplicity, that $T_{\sigma\sigma'} = 1\ \forall\sigma,\sigma'$.

If $\Lambda_N = [1,N]$ and if we recall definitions (5.1.2) and (5.1.3) (see in particular (5.1.9), (5.1.11) and (5.1.12)) the conditional probabilities of the distributions $\mu \in G(\Phi)$ can be expressed as

e6.1.15
$$p_\Phi(\underline{\sigma}_{\Lambda_N}|\underline{\sigma}_{\Lambda_N^c}) = \frac{\exp\Big(-\sum\limits_{j=1}^{N} A_\Phi(\tau^j\underline{\sigma}) - D_{\Lambda_N,1}(\underline{\sigma})\Big)}{\sum\limits_{\underline{\sigma}'}\exp\Big(-\sum\limits_{j=1}^{N} A_\Phi(\tau^j\underline{\sigma}') - D_{\Lambda_N,1}(\underline{\sigma}')\Big)}, \tag{6.1.15}$$

with $\underline{\sigma} = (\underline{\sigma}_{\Lambda_N}\underline{\sigma}_{\Lambda_N^c})$, $\underline{\sigma}' = (\underline{\sigma}'_{\Lambda_N},\underline{\sigma}_{\Lambda_N^c}) \in \{0,\ldots,n\}^{\mathbb{Z}}$, see also (5.1.8). Here $D_{\Lambda_N,1}(\underline{\sigma})$ is a *boundary term* or, as we called it in definition (5.1.3), a *surface correction.* Define

e6.1.16
$$Z_{\Lambda_N}(\Phi,\underline{\sigma}_{\Lambda_N^c}) = \sum_{\underline{\sigma}'_{\Lambda_N}} \exp\Big(-\sum_{j=1}^{N} A_\Phi(\tau^j\underline{\sigma}') - D_{\Lambda_N,2}(\underline{\sigma}')\Big), \tag{6.1.16}$$

where $D_{\Lambda_N,2}(\underline{\sigma})$ is another surface correction (see again definition (5.1.3)). Given a sequence $u_i \in \mathbb{R}$, let $\overline{p}_i = e^{-u_i}/\sum_j e^{-u_j}$. Then from the inequality

e6.1.17

$$\begin{aligned}\log\sum_i e^{-u_i} &= \max_{\sum_i p_i=1,\, p_i\ge 0}\sum_i p_i(-\log p_i - u_i) = \\ &= \sum_i \overline{p}_i(-\log \overline{p}_i - u_i),\end{aligned}\tag{6.1.17}$$

we deduce

e6.1.18

$$\begin{aligned}\log Z_{\Lambda_N}(\Phi,\underline{\sigma}_{\Lambda^c_N}) = \sum_{\underline{\sigma}'_{\Lambda_N}} & p_\Phi(\underline{\sigma}'_{\Lambda_N}|\underline{\sigma}_{\Lambda^c_N})\Big(-\log p_\Phi(\underline{\sigma}'_{\Lambda_N}|\underline{\sigma}_{\Lambda^c_N})- \\ & -\sum_{j=1}^{N} A_\Phi(\tau^j\underline{\sigma}') - D_{\Lambda_N,2}(\underline{\sigma}')\Big).\end{aligned}\tag{6.1.18}$$

To bound this expression in terms of $P(\Phi)$, we note that (5.1.12), (5.1.13) (5.1.14) and (5.1.16) imply the existence of a simple relation between $Z_N(\Phi)$ defined in (6.1.1) and $Z_{\Lambda_N}(\Phi,\underline{\sigma}_{\Lambda^c_N})$ in (6.1.16); namely

e6.1.19

$$Z_N(\Phi)/Z_{\Lambda_N}(\Phi,\underline{\sigma}_{\Lambda^c_N}) \le \exp\vartheta'\varepsilon_N,\tag{6.1.19}$$

with $\vartheta' \in [-1,1]$ and $\varepsilon_N = \varepsilon_{N,1}$ (see (5.1.14)) such that $\varepsilon_N/N \xrightarrow[N\to\infty]{} 0$.

Furthermore, if $\underline{\sigma} \in \{0,\ldots,n\}^{\mathbb{Z}}$, it is tautological that

e6.1.20

$$p_\Phi(\underline{\sigma}_{\Lambda_N}\,|\,\underline{\sigma}_{\Lambda^c_N}) = \int \mu(d\underline{\sigma}')\, p_\Phi(\underline{\sigma}_{\Lambda_N}|\underline{\sigma}_{\Lambda^c_N})\tag{6.1.20}$$

(because the integrand is constant) but, by (5.2.5) (or, better, by the same computation that leads to (5.2.5) and without making use of the hypothesis $\|\Phi\|_1 < +\infty$) we see that

e6.1.21

$$\frac{p_\Phi(\underline{\sigma}_{\Lambda_N}|\underline{\sigma}'_{\Lambda^c_N})}{p_\Phi(\underline{\sigma}_{\Lambda_N}\,|\,\underline{\sigma}_{\Lambda^c_N})} \le \exp\vartheta''\varepsilon'_N,\tag{6.1.21}$$

N6.1.2 with $\varepsilon'_N/N \xrightarrow[N\to\infty]{} 0$ and $\vartheta'' \in [-1,1]$.[2] Hence combining the tautology (6.1.20) with the bound (6.1.21) we get a bound for $p_\Phi(\underline{\sigma}_{\Lambda_N}|\underline{\sigma}'_{\Lambda^c_N})$

e6.1.22

$$\begin{aligned}p_\Phi(\underline{\sigma}_{\Lambda_N}|\underline{\sigma}_{\Lambda^c_N}) \ge \exp(-\varepsilon'_N)\int p_\Phi(\underline{\sigma}_{\Lambda_N}|\underline{\sigma}'_{\Lambda^c_N})\mu(d\underline{\sigma}') \equiv \\ \equiv \exp(-\varepsilon'_N)\mu(C^{\Lambda_N}_{\underline{\sigma}_{\Lambda_N}}).\end{aligned}\tag{6.1.22}$$

---

[2] ε'_N can be taken to depend only on N.

Then (6.1.19) and (6.1.18) say that

e6.1.23
$$\begin{split}N^{-1}\log Z_N(\Phi)\le N^{-1}\sum_{\underline{\sigma}'_{\Lambda_N}} p_\Phi(\underline{\sigma}'_{\Lambda_N}|\underline{\sigma}_{\Lambda^c_N})\Big(-\log\mu(C^{\Lambda_N}_{\underline{\sigma}'_{\Lambda_N}})-\\ -\sum_{j=1}^{N}A(\tau^j\underline{\sigma}')\Big)+\eta_N,\end{split}\tag{6.1.23}$$

with $\eta_N\xrightarrow[N\to\infty]{}0$, uniformly in $\underline{\sigma}_{\Lambda^c_N}$.

Equation (6.1.23) can be thought of as an inequality between functions of $\underline{\sigma}_{\Lambda^c_N}$ and it can be integrated with respect to μ. We find [3] N6.1.3

e6.1.24
$$\begin{split}N^{-1}\log Z_N(\Phi)\le N^{-1}\sum_{\underline{\sigma}'_{\Lambda_N}}-\mu(C^{\Lambda_N}_{\underline{\sigma}'_{\Lambda_N}})\log\mu(C^{\Lambda_N}_{\underline{\sigma}'_{\Lambda_N}})-\\ -N^{-1}\sum_{j=1}^{N}\mu(A_\Phi)+\overline{\eta}_N,\end{split}\tag{6.1.24}$$

and, taking the limit as $N\to\infty$,

e6.1.25
$$P(\Phi)\le s(\mu)-\mu(A_\Phi),\tag{6.1.25}$$

where $s(\mu)$ is the average entropy of the probability distribution μ with respect to the translation, (cf. definition (3.3.2) and the theorem of the generator given in corollary (3.4.1)).

On the other hand, in general, by (6.1.17) and for an arbitrary τ–invariant probability distribution $m\in\mathcal{M}(\{0,\ldots,n\}^{\mathbb{Z}},\tau)$ one has

e6.1.26
$$N^{-1}\log Z_N(\Phi)>N^{-1}\sum_{\underline{\sigma}_{\Lambda_N}} m(C^{\Lambda_N}_{\underline{\sigma}_{\Lambda_N}})(-\log m(C^{\Lambda_N}_{\underline{\sigma}_{\Lambda_N}})-U^{\Phi}_{\Lambda_N}(\underline{\sigma}_{\Lambda_N})),\tag{6.1.26}$$

and since

e6.1.27
$$\begin{split}\sum_{\underline{\sigma}_{\Lambda_N}} m(C^{N}_{\underline{\sigma}_{\Lambda_N}})U^{\Phi}_{\Lambda_N}(\underline{\sigma}_{\Lambda_N})=\int m(d\underline{\sigma})U^{\Phi}_{\Lambda_N}(\underline{\sigma}_{\Lambda_N})=\\ =\int m(d\underline{\sigma})\Big(\sum_{j=1}^{N}A_\Phi(\tau^j\underline{\sigma})+\Delta_N(\underline{\sigma})\Big)=N(m(A_\Phi)+\eta'_N)\end{split}\tag{6.1.27}$$

with $\eta'_N\xrightarrow[N\to\infty]{}0$, one finds for all $m\in\mathcal{M}(\{0,\ldots,n\}^{\mathbb{Z}},\tau)$

e6.1.28
$$P(\Phi)\ge s(m)-m(A_\Phi).\tag{6.1.28}$$

---

[3] $\overline{\eta}_N$ is the integral of η_N.

By (6.1.25) and (6.1.28) we therefore obtain

e6.1.29
$$P(\Phi) = s(\mu) - \mu(A_\Phi) \qquad \text{for all } \mu \in G(\Phi). \tag{6.1.29}$$

Hence by (6.1.28) and (6.1.29)

e6.1.30
$$\begin{aligned} P(\Phi+\Psi) &\geq s(\mu) - \mu(A_{\Phi+\Psi}) = \\ &= s(\mu) - \mu(A_\Phi) - \mu(A_\Psi) = P(\Phi) - \mu(A_\Psi) \end{aligned} \tag{6.1.30}$$

which means that $\Psi \to \alpha(\Psi) = -\mu(A_\Psi)$ is a tangent functional.

To show the converse part of the proposition we make use of a general property of tangent planes to graphs of continuous and convex functions defined on a separable Banach space: the tangent plane is unique on a set which is at least dense and, furthermore, given $\Phi \in B$ and a functional α_Φ tangent to the graph of $\Phi \to P(\Phi)$ in the point Φ, one can find, for every $k \geq 0$, s_k points $\Phi^{(k)}_1$, $\Phi^{(k)}_2, \ldots, \Phi^{(k)}_{s_k}$, with $\|\Phi^{(k)}_j - \Phi\| < 1/k$, in which the graph of P has a unique tangent plane $\alpha_{\Phi^{(k)}}$ and s_k positive numbers $a^{(k)}_1$, $a^{(k)}_2, \ldots, a^{(k)}_{s_k}$ such that

e6.1.31
$$\begin{aligned} &\sum_{j=1}^{s_k} a^{(k)}_j = 1, \\ &\sum_{j=1}^{s_k} a^{(k)}_j \alpha_{\Phi^{(k)}_j}(\Psi) \xrightarrow[k\to\infty]{} \alpha(\Psi) \qquad \text{for all } \Psi \in B. \end{aligned} \tag{6.1.31}$$

N6.1.4 This is a general property of convex functionals [4] which can be applied to our case in the following way.

We begin by remarking that given a functional α and *assuming* that there is a probability distribution $\mu \in \mathcal{M}(\{0,\ldots,n\}^{\mathbb{Z}}, \tau)$ such that $\alpha(\Psi) = -\mu(A_\Psi)$, for all $\Psi \in B$, then μ is unique. The reason is that by choosing a suitable Ψ we can compute the μ–probability of a given cylinder $C^J_{\widehat{\underline{\sigma}}_J}$ with base on the set J.

The choice of the potential Ψ is such that $\Psi_X = 0$ unless the set X is a translate $\tau^k J$ of J for some $k \in \mathbb{Z}$ and

e6.1.32
$$\Psi_{\tau^k J}(\underline{\sigma}_J) = \delta_{\underline{\sigma}_J \widehat{\underline{\sigma}}_J}. \tag{6.1.32}$$

Then

e6.1.33
$$\mu(A_\Psi) = \sum_{X \ni 0} \frac{1}{|X|} \mu(\Psi_X(\underline{\sigma}_X)) \equiv \mu(C^J_{\widehat{\underline{\sigma}}_J}), \tag{6.1.33}$$

---

[4] In other words every tangent plane in Φ can be obtained as a limit of a sequence of planes that are (finite) convex linear combinations of planes tangent in points where the tangent plane is unique and whose largest distance to Φ is infinitesimal. For discussion and proof cf. p. 450 of [DS58] and p. 329 of [LR68].

that proves the statement, because μ is in it turn determined by its values on the cylinders.

Therefore given a tangent plane α to the graph of P in Φ one may be tempted to define the probability distribution μ, that should generate the plane α, by setting $\mu(A_\Psi) = -\alpha(\Psi)$.

The problem is that we do not know whether the numbers $\mu(C^J_{\widehat{\underline{\sigma}}_J})$, defined in this way in terms of α, and which should be μ–measures of cylinders, satisfy the properties (i), (ii), (iii) of definition (2.3.2), needed to reconstruct from them a probability distribution (cf. proposition (2.3.1)): note that condition (iv), *i.e.* the translation invariance, of definition (2.3.2) is in this case automatically satisfied.

However equation (6.1.31) says that, if $\mu(C^J_{\underline{\sigma}_J})$ has to be defined in terms of α by (6.1.33). Note that the already proved direct part of the proposition and the remark that α determines μ, if μ exists, imply that if there is a unique tangent plane in Φ then $G(\Phi)$ contains a unique point. From this follows that in the points $\Phi^{(k)}_j$ where the tangent plane is unique, one must necessarily have $\alpha_{\Phi^{(k)}_j}(\Psi) = \mu_{\Phi^{(k)}_j}(A_\Psi)$, for all $\Psi \in B$, where $\mu_{\Phi^{(k)}_j}$ is the unique element of $G(\Phi^{(k)}_j)$ This implies that

*e*6.1.34
$$\mu(C^J_{\underline{\sigma}_J}) = \lim_{k\to\infty} \sum_{j=1}^{s_k} a^{(k)}_j \mu_{\Phi^{(k)}_j}(C^J_{\underline{\sigma}_J}). \tag{6.1.34}$$

Since $\sum_j a^{(k)}_j = 1$ and the quantities $\mu_{\Phi^{(k)}_j}(C^J_{\underline{\sigma}_j})$ are values of measures of the cylinders $C^J_{\underline{\sigma}_J}$ with respect to some probability distribution $\mu_{\Phi^{(k)}_j}$ the r.h.s. of (6.1.34) verifies the (i), (ii), (iii) of definition (2.3.2) so that also $\mu(C^J_{\underline{\sigma}_J})$ verifies the same properties. Finally, by proposition (2.3.1), there exists $\mu \in \mathcal{M}(\{0,\ldots,\}^{\mathbb{Z}},\tau)$ such that

*e*6.1.35
$$\alpha(\Psi) = -\mu(A_\Psi) \qquad \text{for all } \Psi \in B. \tag{6.1.35}$$

It remains to check that the distribution μ is in $G(\Phi)$.

For this purpose we remark that the first of (6.1.31) and the explicit definition of conditional probability for a Gibbs distribution (cf. Section §5.1, (5.1.9)), immediately imply

*e*6.1.36
$$\sup_j \sup_{\underline{\sigma}} |p_\Phi(\underline{\sigma}_\Lambda|\underline{\sigma}_{\Lambda^c}) - p_{\Phi^{(k)}_j}(\underline{\sigma}_\Lambda|\underline{\sigma}_{\Lambda^c})| \xrightarrow[k\to\infty]{} 0 \tag{6.1.36}$$

for all $\Lambda \subset \mathbb{Z}$. Furthermore (6.1.34) means that μ is the weak limit $\lim_{N\to\infty} \sum_{j=1}^k a^{(k)}_j \mu_{\Phi^{(k)}_j}$. Hence if $\underline{\sigma}_\Lambda \to f(\underline{\sigma}_\Lambda)$ is a $\mathcal{B}(\Lambda)$–measurable function we have

$$\mu(f) = \lim_{k\to\infty} \sum_{j=1}^{s_k} a^{(k)}_j \int \mu_{\Phi^{(k)}_j}(d\underline{\sigma}) f(\underline{\sigma}_\Lambda) =$$

e6.1.37

$$\begin{aligned}
&= \lim_{k\to\infty} \sum_{j=1}^{s_k} a_j^{(k)} \sum_{\underline{\sigma}_\Lambda} \int \mu_{\Phi_j^{(k)}}(d\underline{\sigma}')\, p_{\Phi_j^{(k)}}(\underline{\sigma}_\Lambda|\underline{\sigma}'_{\Lambda^c}) f(\underline{\sigma}_\Lambda) = \\
&= \lim_{k\to\infty} \sum_{\underline{\sigma}_\Lambda} \int \Big(\sum_{j=1}^{s_k} a_j^{(k)} \mu_{\Phi_j^{(k)}}(d\underline{\sigma}') \Big) p_\Phi(\underline{\sigma}_\Lambda|\underline{\sigma}'_{\Lambda^c}) f(\underline{\sigma}_\Lambda) + \qquad (6.1.37)\\
&+ \lim_{k\to\infty} \sum_{\underline{\sigma}_\Lambda} \int \Big(\sum_{j=1}^{s_k} a_j^{(k)} \mu_{\Phi_j^{(k)}}(d\underline{\sigma}') \Big) \big(p_{\Phi_j^{(k)}}(\underline{\sigma}_\Lambda|\underline{\sigma}'_{\Lambda^c}) - p_\Phi(\underline{\sigma}_\Lambda|\underline{\sigma}'_{\Lambda^c}) \big) f(\underline{\sigma}_\Lambda) = \\
&= \sum_{\underline{\sigma}_\Lambda} \int \mu(d\underline{\sigma}')\, p_\Phi(\underline{\sigma}_\Lambda|\underline{\sigma}'_{\Lambda^c}) f(\underline{\sigma}_\Lambda),
\end{aligned}$$

by (6.1.36) and because $\underline{\sigma} \to p_\Phi(\underline{\sigma}_\Lambda|\underline{\sigma}_{\Lambda^c})$ is a continuous function (hence uniformly continuous). Equality between first and last term in (6.1.37) means that $p_\Phi(\underline{\sigma}_\Lambda|\underline{\sigma}_{\Lambda^c})$ is the conditional probability relative to the region Λ of the distribution μ: thus, by the arbitrariness of Λ, we see that $\mu \in G(\Phi)$. ■

We implicitly obtained also the following results.

C6.1.1 **(6.1.1) Corollary:** (Variational principle for the pressure)
Under the hypotheses of proposition (6.1.1) the set $G(\Phi)$ consists of all probability distributions $\mu \in \mathcal{M}(\{0,\ldots,n\}_T^{\mathbb{Z}}, \tau)$ for which

e6.1.38

$$P(\Phi) = \max_{m\in\mathcal{M}(\{0,\ldots,n\}_T^{\mathbb{Z}},\tau)} (s(m) - m(A_\Phi)) = s(\mu) - \mu(A_\Phi). \qquad (6.1.38)$$

Remark: This is the *variational principle* for the determination of the Gibbs distributions of given potential (Ruelle).

C6.1.2 **(6.1.2) Corollary:** (Decomposition into pure phases)
Under the hypotheses of proposition (6.1.1) the following properties hold.
(i) If $\mu \in G(\Phi)$ then its ergodic decomposition π_μ (cf. *Section §2.4) has support in $G_e(\Phi) = G(\Phi) \cap \mathcal{M}_e$.*
(ii) The set $\mathcal{E}_e(\Phi) = \{$set of the ergodic points of $\{0,\ldots,n\}_T^{\mathbb{Z}}$ which generate a probability distribution in $G_e(\Phi)\}$ (cf. *Section §2.3) is a Borel set and $\mu(\mathcal{E}_e(\Phi)) = 1$, for all $\mu \in G(\Phi)$.*
N6.1.5 *(iii) The extremal points of $G(\Phi)$ are the points of $G_e(\Phi)$.*[5]

Remark: A consequence of this corollary is an interesting interpretation of the pressure $P(\Phi)$ as complexity (cf. definition (3.1.1)). Let in fact D_N be a sequence of functions on $\{0,\ldots,n\}^N$, $N = 1, 2, \ldots$, such that

e6.1.39

$$\lim_{N\to\infty} \sup_{\sigma_1\ldots\sigma_N} N^{-1} |D_N(\sigma_1\ldots\sigma_N)| = 0, \qquad (6.1.39)$$

---

[5] If G is a closed convex set in a linear topological space we say that $\mu \in G$ is extremal if the relation $\mu = \alpha'\mu' + \alpha''\mu''$, with μ', $\mu'' \in G$ and $\alpha' + \alpha'' = 1$, α', $\alpha'' \in (0,1)$, implies $\mu' = \mu''$.

and set

e6.1.40
$$V_N(\sigma_1 \ldots \sigma_N) = U^{\Phi}_{\Lambda_N}(\sigma_1 \ldots \sigma_N) + D_N(\sigma_1 \ldots \sigma_N). \tag{6.1.40}$$

Then, if $\mu \in G(\Phi)$, μ–almost all points $\underline{\sigma}$ of $\{0, \ldots, n\}^{\mathbb{Z}}_T$ have complexity with weight $\underline{V} = \{V_N\}_{N=0}^{\infty}$ given by

e6.1.41
$$s(\underline{\sigma}, \{V_N\}_{N=0}^{\infty}) = P(\Phi) = s(\mu) - \mu(A_\Phi). \tag{6.1.41}$$

This is a consequence of the above corollaries and of the Shannon–McMillan theorem, see also problem [6.1.2].

Proof: The set $G(\Phi)$ is convex and compact, see corollary (5.1.1), and item (i) follows from the ergodic decomposition (proposition (2.4.1)) which implies that $\pi_\mu(\mathcal{M}_e \cap G(\Phi)) = 1$, if $\mu \in G(\Phi)$.
Item (ii) follows immediately from the results of problem [2.4.6].
Item (iii) can be proved by noting that the extremal points of $\mathcal{M}$ are the points in $\mathcal{M}_e$: see remark (5) to proposition (2.4.1).
Therefore assuming existence of a distribution μ, extremal in $G(\Phi)$ but not in $\mathcal{M}$ (*i.e.* not ergodic), one would find $\alpha \in (0, 1)$ and two probability distributions in μ_1 and μ_2 with $\mu_1 \neq \mu_2$ and $\mu_2 \in \mathcal{M} \setminus G(\Phi)$ such that $\mu = \alpha\mu_1 + (1 - \alpha)\mu_2$. Then $P(\Phi) = s(\mu) - \mu(A_\Phi) = \alpha(s(\mu_1)) - \mu_1(A_\Phi)) + (1-\alpha)(s(\mu_2) - \mu_2(A_\Phi)) < P(\Phi)$ by the convexity of the average entropy and by the variational principle in (6.1.38). Hence the extremal points of $G(\Phi)$ would be actually extremal also in $\mathcal{M}$ and therefore ergodic. The proof is therefore complete. ■

Problems for §6.1

Q6.1.1 **[6.1.1]:** Prove (6.1.41). (*Hint:* Consider the function

$$\widehat{f}_N(\underline{\sigma}) = -N^{-1} \log(\mu(C^N_{\underline{\sigma}_{\Lambda_N}}) e^{-V_N(\underline{\sigma}_{\Lambda_N})})$$

and show that as $N \to \infty$ it converges in measure to $s(\mu) - \mu(A_\Phi)$ if $\mu \in G_e(\Phi)$; then make use of the argument given in the remark (2) to proposition (3.2.1))

Q6.1.2 **[6.1.2]:** (*Pressure of the Ising potential*)
Compute the pressure of the potential on $\{0, 1\}^{\mathbb{Z}}$ such that $\Phi_X = 0$ unless X is a set which is a translation of a single point or of two nearest neighbors and

$$\Phi_{\{0\}}(\sigma) = \begin{cases} h & \sigma = 1 \\ -h & \sigma = 0 \end{cases}, \qquad \Phi_{\{0,1\}}(\sigma, \sigma') = J(2\sigma - 1)(2\sigma' - 1),$$

which is called "Ising potential in external field h." (*Hint:* : Remark that Z_N can be written as the scalar product $v \cdot \Xi^N v$, where Ξ is a suitable 2×2 matrix (*transfer matrix*) and $v = (1, 1)$. $P(\Phi)$ is the linked to the eigenvalues of Ξ.)

Q6.1.3 **[6.1.3]:** Compute by means of the variational principle the probability $\mu(C^{00}_{11})$ with respect to the probability distribution of parameters J and h considered in problem [6.1.2].

Q6.1.4 **[6.1.4]:** (*Fisher potential*)
Like problem [6.1.3] for the *Fisher potential* introduced in problem [5.2.1].

Q6.1.5 **[6.1.5]:** (*Exact solubility of Fisher potentials*)
Compute the measure of the cylinders $C^{0\ldots k}_{1\ldots 1}$ for the Fisher potential of the preceding problem, using equation (6.1.33). The fact that an exact computation of the measure of *all* cylinders and of the pressure, entropy and other thermodynamic functions, is possible (and not difficult) makes the Fisher potentials an important tool for the analysis of examples and counterexamples.

Q6.1.6 **[6.1.6]:** (*Pressure singularities and non-uniqueness*)
Check that if $P(\Phi)$ is analytic in Φ for Φ in an open subset Δ in the space of the potentials, $\Delta \subset B$, then the tangent plane to Φ is unique in every point of Δ. (*Hint:* One says that a functional P defined on a Banach space B is analytic at a point Φ if given arbitrarily $k > 0$ and $\Psi_1, \ldots, \Psi_k \in B$ the function $\varepsilon_1 \ldots \varepsilon_k \to P(\Phi + \varepsilon_1\Psi_1 + \ldots + \varepsilon_k\Psi_k)$ is analytic in $\varepsilon_1, \ldots, \varepsilon_k$ in a neighborhood of the origin.)

Q6.1.7 **[6.1.7]:** (*A Fisher potential with phase transition*) Show that if $P(\Phi + \varepsilon\Psi)$ has right derivative different from the left derivative, as function of ε, at $\varepsilon = 0$ then $G(\Phi)$ contains at least two different points (one says that a "*phase transition*" takes place as ε varies).

Q6.1.8 **[6.1.8]:** Show the existence of Fisher potentials with $\|\Phi\|_1 = +\infty$ (cf. problem [6.1.4]) for which P is not an analytic function of the "one-body" component of Φ (*i.e.* of $\Phi_{\{0\}}(1) \equiv -\Phi_0$) and that $P(\Phi)$ has right derivative different from the left derivative at a suitable values of Φ_0. Check that this cannot happen if $\|\Phi\|_1 < +\infty$. (*Hint:* Consider the case $\Phi_n = (1+n)^{-1-\varepsilon}$, with $0 < \varepsilon < 1$, for $n \geq 1$ and Φ_0 arbitrary; then P is not analytic as a function of Φ_0, by explicit computation. This requires having solved problem [6.1.5].)

Bibliographical note to §6.1

The contents of this section are taken from various sources, for instance from Ch. 7, §3,4 of [Ru69], p. 450 of [DS58], [LR68]. See also the bibliographical note to §5.1. The theory of the pressure regarded as a generating function of a Gibbs state is ancient, some examples of "rigorous" applications of this elegant technique are in [Ru72].

§6.2 Applications to Anosov systems. SRB distribution

A classical application of the theory of Gibbs distributions concerns the analysis of the invariant probability distributions associated with an Anosov system (Ω, S) on a compact Riemannian manifold (Sinai).

The definition that we adopted of Anosov system, cf. definition (4.2.1), assumes topological transitivity and implies topological mixing, cf. problem [4.2.10], *i.e.* that, given two open sets G, $F \subset \Omega$ there exists $k_0 > 0$ such that if $k \geq k_0$ then $S^k G \cap F \neq \emptyset$. We shall rely on the latter properties although several results could be suitably extended to cover much more general hyperbolic systems. In agreement with the notations introduced in definition (2.2.2) we shall recall or set the following definitions.

D6.2.1 **(6.2.1) Definition:** (Topological and ergodic distributions)
We denote
(i) $\mathcal{M}(\Omega, S)$ the set of the S–invariant Borel distributions on Ω,
(ii) $\mathcal{M}_e(\Omega, S) \subset \mathcal{M}(\Omega, S)$ the set of the distributions on Ω which are S–ergodic,

(iii) $\mathcal{M}^t_e(\Omega,S)\subset\mathcal{M}_e(\Omega,S)\subset\mathcal{M}(\Omega,S)$ *the set of the* S*–ergodic distributions which give a positive measure to all open sets.*
More generally $\mathcal{M}^{0,t}(\Omega)$ *will be the set of* topological distributions*,* cf. *definition (4.1.2), on* Ω*.*

The aim of the following analysis is to show that Anosov systems admit *"very many" probability distributions in* $\mathcal{M}^t_e(\Omega,S)$ and to show that, among them, one can identify several ones either on the basis of a criterion of "maximum complexity" or by attibuting a special role to the statistical properties of motions whose initial data are chosen randomly with a probability distribution which is absolutely continuous with respect to the volume.

D6.2.2 **(6.2.2) Definition:** (SRB distributions and f–complex distributions)
Given an Anosov map (Ω,S)
(i) A topological probability distribution $\mu\in\mathcal{M}(\Omega,S)$ *is said to have maximum complexity with respect to a continuous function* f *if it maximizes the quantity* $s(\mu)-\mu(f)$*.*
(ii) If μ_0 *is a probability distribution absolutely continuous with respect to the volume measure on* Ω *and if for all continuous functions* F *on* Ω *and for* μ_0*–almost all initial data* $x\in\Omega$ *the limit* $\lim_{N\to\infty}N^{-1}\sum_{j=0}^{N-1}F(S^jx)$ *esists and can be written as*

6.2.0
$$\lim_{N\to\infty}N^{-1}\sum_{j=0}^{N-1}F(S^jx)=\int_\Omega\mu(dy)F(y)\qquad(6.2.1)$$

with $\mu\in\mathcal{M}(\Omega,S)$*, then* μ *is called the SRB distribution of* (Ω,S)*.*

Remark: (1) The results of this section are general results preparing the key results of this Chapter: namely that all Anosov systems admit a SRB distribution, proposition (6.3.3) and corollary (6.3.1), which furthermore is the invariant probability distribution which maximizes complexity with respect to the continuous function $f(x)=\lambda_u(x)$, proposition (6.3.4).
(2) Technically the above properties will be closely related to the results of this section which allow us to identify probability distributions that maximize complexity with suitable Gibbs distributions with short range potentials, see lemma (6.2.1), and with properties of periodic motions, see proposition (6.2.1).

The notion of complexity has been discussed in Section §3.1, where we have given the definitions of complexity of sequences of symbols $\underline{\sigma}$ with respect to families $\{U_N\}_{N\ge0}^\infty$ of weights, *i.e.* of functions defined on strings of length N in $\underline{\sigma}\in\{0,\ldots,n\}^N_T$.
Given a potential $\Phi\in B$, cf. definition (5.1.1), corollary (6.1.1) shows that a distribution μ on $\{0,\ldots,n\}^{\mathbb{Z}}_T$ that maximizes the complexity of μ–almost all the points relatively to a weight $U_N(\underline{\sigma})=\sum_{R\subset\{1,\ldots,N\}}\Phi_R(\underline{\sigma}_R)$ is a Gibbs distribution with potential Φ; such complexity is, then, equal to the pressure

e6.2.2
$$P(\Phi)=\max_{\mu'\in\mathcal{M}_e(\{0,\ldots,n\}^{\mathbb{Z}}_T)}(s(\mu')-\mu'(A_\Phi)).\qquad(6.2.2)$$

In Anosov systems, as seen in Section §4.2, Markovian pavements, which always exist, allow us to describe points of Ω by means of symbolic sequences and their evolution by the shift map.

One is thus led to the search of invariant and topological distributions on Ω via symbolic dynamics. At the same time one would like to characterize the distributions that in the symbolic language become Gibbs distributions in terms of quantities that do not depend explicitly on specific Markovian pavements.

Given the one-to-one correspondence between the S–ergodic topological distributions on Ω and the τ–ergodic topological distributions on $\{1,\ldots,q\}_T^{\mathbb{Z}}$ discussed in proposition (4.1.1) and in the relative remark (2) one is tempted, for instance, to associate with (6.2.2) the "intrinsic", *i.e.* Markovian pavement independent, problem of finding the distributions μ' that maximize

e6.2.3
$$s(\mu') - \mu'(f) \tag{6.2.3}$$

on $\mathcal{M}_e^t(\Omega, S)$ with f a continuous function on Ω.

As we shall see, if f is Hölder continuous, *i.e.* if it verifies a "modest" requirement of regularity beyond the mere continuity, it will be possible to give a quite satisfactory answer to the above questions.

Let (Ω, S) be an Anosov system. Denote $\mathcal{P} = \{P_1, \ldots, P_q\}$ a Markovian pavement constructed with S–rectangles as in Section §4.2: let T be its compatibility matrix (cf. Section §4.1). Then (Ω, S) is mixing and, hence, there exists $a \geq 0$ such that $(T^k)_{\sigma\sigma'} > 0$ for all $k > a$.

Denote by $\underline{\sigma} \to X(\underline{\sigma})$ the Hölder continuous code of $\{1,\ldots,q\}_T^{\mathbb{Z}}$ into Ω discussed in Section §4.1 defined by

e6.2.4
$$x = X(\underline{\sigma}) = \bigcap_{j=-\infty}^{+\infty} S^{-j} P_{\sigma_j}. \tag{6.2.4}$$

The following preliminary result holds.

L6.2.1 **(6.2.1) Lemma:** *Let (Ω, S) be an Anosov system and let $\mathcal{P} = \{P_0, \ldots, P_q\}$ be a Markovian pavement for it with compatibility matrix T.*
(i) If f is Hölder continuous the function $\mu \to s(\mu) - \mu(f)$ admits a maximum $p(f)$ on $\mathcal{M}_e^t(\Omega)$.
*(ii) Let $\nu_N \in \mathcal{M}^{0,t}(\Omega)$ be an arbitrary sequence of topological distributions (*cf.* definition (6.2.1)) on Ω and, given $\underline{\sigma} \in \{1,\ldots,q\}_T^{[0,N]}$, let*

e6.2.5
$$U_N(\sigma_0 \ldots \sigma_N) = \frac{\int_{P^{0\ldots N}_{\sigma_0\cdots\sigma_N}} \sum_{j=0}^N f(S^j x)\nu_N(dx)}{\int_{P^{0\ldots N}_{\sigma_0\cdots\sigma_N}} \nu_N(dx)}, \tag{6.2.5}$$

where $P^{0\ldots N}_{\sigma_0\ldots\sigma_N} = \bigcap_{j=0}^{N} S^{-j} P_{\sigma_j}$. Then the image m of $\mu \in \mathcal{M}_e^t(\Omega, S)$ via the isomorphism mod 0 *defined by (6.2.4) maximizes the complexity with weight $\{U_N\}_{N=1}^\infty$ if and only if μ maximizes the function $\mu \to s(\mu) - \mu(f)$ on $\mathcal{M}_e^t(\Omega, S)$.*

Remark: Note that the sequence ν_N and the Markovian pavement $\mathcal{P}$ used to formulate this theorem are in a certain sense irrelevant: the image m of the probability distribution $\mu \in \mathcal{M}_e^t(\Omega, S)$, that maximizes $s(\mu) - \mu(f)$, makes largest the complexity associated with U_N for any choice of $\mathcal{P}$ or of the sequence ν_N that allows us to construct U_N. Hence it is natural to call $s(\mu) - \mu(f) = p(f)$ the *complexity with weight* f of μ on Ω: this is the interest of the statement (ii).

Proof: Let a be the mixing time of T. Let $\underline{\rho} \in \{1, \ldots, q\}_T^{\mathbb{Z}}$ be a T–compatible sequence, cf. definition (4.1.1), chosen once and for all.

We first need to define a way to merge together two compatible strings into a longer compatible string. This can be done in several "standard ways" provided the compatibility matrix T is mixing.

Suppose that $a \geq 1$ (*i.e.* we do not assume that the compatibility relation is trivial). Associate with each pair of symbols σ, σ' a T–compatible string $(\sigma\eta_0 \ldots \eta_{a-1}\sigma')$ and let $\underline{\vartheta}(\sigma, \sigma') = (\eta_0 \ldots \eta_{a-1})$.

Given such a correspondence ϑ, which we shall call an *interpolation string*, for any two finite T–compatible strings $\underline{\sigma} = (\sigma_1 \ldots \sigma_p)$ and $\underline{\sigma}' = (\sigma_1' \ldots \sigma_q')$ we can form the compatible strings

$$\underline{\sigma}\,\underline{\vartheta}(\sigma_p, \sigma_1')\,\underline{\sigma}' \qquad \text{and} \qquad \underline{\sigma}'\,\underline{\vartheta}(\sigma_q', \sigma_1)\,\underline{\sigma},$$

where, as usual, we indicate with $\underline{\sigma}\,\underline{\vartheta}(\sigma_p, \sigma_1')\,\underline{\sigma}'$ the string starting with $\underline{\sigma}$ continuing with $\underline{\vartheta}(\sigma_p, \sigma_1')$ and ending with $\underline{\sigma}'$. If $\underline{\sigma}$ is semiinfinite only one of the above "mergers" is possible and if both $\underline{\sigma}, \underline{\sigma}'$ are semiinfinite one merger is possible provided $\underline{\sigma}$ is infinite to the left and $\underline{\sigma}'$ is infinite to the right or *vice versa*. We shall simply denote $\underline{\sigma} \circ \underline{\vartheta} \circ \underline{\sigma}'$ the merger of $\underline{\sigma}$ and $\underline{\sigma}'$ with an interpolation string ϑ.

In the continuation of the proof just started we imagine that the set of q^2 interpolating strings $\underline{\vartheta}$ has been fixed once and for all. If $\underline{\sigma}, \underline{\rho} \in \{1, \ldots, q\}_T^{\mathbb{Z}}$ and $N \in \mathbb{N}$ we shall denote by $(\underline{\sigma}^{(N)}\underline{\rho})$ a sequence whose labels for $i \in [-N, N]$ coincide with σ_i and for $i \notin (-N-a, N+a)$ coincide with ρ_i, while for the other i's the sequence $(\underline{\sigma}^{(N)}, \underline{\rho})$ is interpolated by the appropriate interpolating strings $\underline{\vartheta}$.

If $f \in C^\alpha(\Omega)$, as already remarked elsewhere, (cf. remark (3) to definition (4.3.2)), the quantity $\underline{\sigma} \to f(X(\underline{\sigma}))$ can be well developed in terms of cylindrical functions starting from the identity

e6.2.6
$$f(X(\underline{\sigma})) = f(X(\underline{\rho})) + \sum_{N=0}^{\infty} [f(X(\underline{\sigma}^{(N)}\underline{\rho})) - f(X(\underline{\sigma}^{(N-1)}\underline{\rho}))], \tag{6.2.6}$$

where $f(X(\underline{\sigma}^{(-1)}\underline{\rho})) \overset{def}{=} f(X(\underline{\rho}))$. In fact set

e6.2.7
$$\begin{aligned} \Gamma_Y(\underline{\sigma}_Y) = {} & 0 \qquad Y \neq [-N, N] \text{ for some } N \geq 0 \text{ and} \\ \Gamma_Y(\underline{\sigma}_Y) = {} & \big(f(X(\underline{\sigma}^{(N)}\underline{\rho})) - f(X(\underline{\sigma}^{(N-1)}\underline{\rho}))\big), \end{aligned} \tag{6.2.7}$$

where in the second line we suppose that Y is the interval $Y=[-N,N]$ centered at 0. Then we rewrite (6.2.6) as

e6.2.8
$$f(X(\underline{\sigma}))=f(X(\underline{\rho}))+\sum_{Y\ni 0}\Gamma_Y(\underline{\sigma}_Y), \tag{6.2.8}$$

and we note that, if $C_{f,\alpha}$ is the Hölder continuity modulus of $f\in C^\alpha(\Omega)$ while α' is the Hölder continuity exponent of the code X (cf. definition (4.1.3) and (4.1.5)) and C' is the relative modulus, one has

e6.2.9
$$|\Gamma_Y(\underline{\sigma}_Y)|\le C_{f,\alpha}C'^{\alpha}\exp(-\alpha\alpha'(N-1))=\widetilde{C}e^{-b\,\mathrm{diam}(Y)}, \tag{6.2.9}$$

with $\widetilde{C},b>0$. Therefore if we set

e6.2.10
$$\begin{aligned}&\Phi_X(\underline{\sigma}_X)=\Gamma_X(\tau^{-i}\underline{\sigma}_X) &&\text{if } X=\tau^i([-N,N]) \text{ for some } N \text{ and } i,\\ &\Phi_X(\underline{\sigma}_X)=0 &&\text{otherwise,}\end{aligned} \tag{6.2.10}$$

and

e6.2.11
$$A_\Phi(\underline{\sigma})=\sum_{X\ni 0}\frac{\Phi_X(\underline{\sigma}_X)}{|X|}, \tag{6.2.11}$$

we find that Φ is a potential in the space B introduced in definition (5.1.1), because Φ verifies the condition of the problem [5.3.7] of Section §5.3 (*exponential decay of the potential*), so that $||\Phi||_1<\infty$. Hence one deduces

e6.2.12
$$\Big|\sum_{j=0}^{N-1}\Big(f(S^jX(\underline{\sigma}))-f(S^jX(\underline{\rho}))\Big)-\sum_{j=0}^{N-1}A_\Phi(\tau^j\underline{\sigma})\Big|\le C'' \tag{6.2.12}$$

for every $\underline{\sigma}\in\{1,\ldots,q\}^{\mathbb{Z}}_T$ and for all $N>0$, if C'' is suitably chosen. We will choose, if possible, $\underline{\rho}$ such that $\tau^j\underline{\rho}=\underline{\rho}$, *i.e.* $\rho=(\ldots i\,i\,i\ldots)$ for some label $i\in\{0,\ldots,q\}$. However the latter sequence may fail to be compatible: in such a case we choose $\underline{\rho}$ to be a periodic sequence of period a, which exists because of the mixing property of T. Below we suppose, for simplicity, that $\rho=(\ldots i\,i\,i\ldots)$ for some label $i\in\{0,\ldots,q\}$ is a compatible sequence.

From corollary (6.1.2) and (5.1.1), (5.1.17), it follows that there exists an ergodic distribution m on $\{1,\ldots,q\}^{\mathbb{Z}}_T$ which maximizes the complexity with weight $\{U_N\}_{N=1}^\infty$ because by (6.2.5) and (6.2.12)

e6.2.13
$$\Big|U_N(\underline{\sigma})-\sum_{R\subset[0,\ldots,N]}\Phi_R(\underline{\sigma}_R)\Big|\le C''' \quad\text{for all } N, \tag{6.2.13}$$

cf. remark to lemma (6.2.1). Therefore from the general theory of Gibbs distributions, see proposition (5.2.1), it follows that $m\in\mathcal{M}^t_e(\{1,\ldots,q\}^{\mathbb{Z}}_T,\tau)$.

Let now μ' be the distribution isomorphic mod 0, via the isomorphism X, to a given distribution m' arbitrarily chosen in $\mathcal{M}^t_e(\{1,\ldots,q\}^{\mathbb{Z}}_T,\tau)$ (cf.

proposition (4.1.1) and (4.1.9)). Then $\mu' \in \mathcal{M}^t_e(\Omega, S)$ and furthermore $s(\mu') = s(m')$, since entropy is invariant under isomorphisms $\bmod\ 0$, and

e6.2.14

$$\begin{aligned}\mu'(f - f(X(\underline{\rho}))) &= \lim_{N\to\infty} \mu'\Big(N^{-1}\sum_{j=0}^{N-1}[f(S^j(\cdot)) - f(X(\underline{\rho}))]\Big) = \\ &= \lim_{N\to\infty} m'\Big(N^{-1}\sum_{j=0}^{N-1}[f(X(\tau^j\cdot)) - f(X(\underline{\rho}))]\Big) = \qquad (6.2.14)\\ &= \lim_{N\to\infty} m'\Big(N^{-1}\sum_{j=0}^{N-1} A_\Phi(\tau^j\cdot)\Big) = m'(A_\Phi),\end{aligned}$$

and therefore, for all $\mu' \in \mathcal{M}^t_e(\{1,\ldots,q\}^{\mathbb{Z}}_T, \tau)$,

e6.2.15

$$s(\mu') - \mu'(f) = -f(X(\underline{\rho})) + s(m') - m'(A_\Phi), \tag{6.2.15}$$

so that the functional on $\mathcal{M}^t_e(\Omega, S)$ in the l.h.s. of (6.2.15) has a maximum which is reached, by the variational principle in corollary (6.1.1), on the $X-$image of the Gibbs distribution $m \in G_e(\Phi)$. Item (ii) is thus also proved. ■

It is convenient to remark that lemma (6.2.1), or better its proof, also yields an explicit method to study the distributions μ that maximize $s(\mu) - \mu(f)$: indeed in deriving (6.2.10) we have shown that such distributions can be considered as Gibbs distributions with a suitable potential Φ_f which satisfies the hypotheses of problems [5.3.15] and [5.3.7], (*i.e. exponential decay*), if $f \in C^\alpha(\Omega)$. This implies the following result.

C6.2.1 **(6.2.1) Corollary:** (Variational principle for pressure of smooth functions)
If $f \in C^\alpha(\Omega)$ there exists a unique distribution that maximizes $s(\mu) - \mu(f)$ over $\mu \in \mathcal{M}^t_e(\Omega)$. Such a distribution μ_f is mixing and

e6.2.16

$$\mu_f(FS^kG) = \int \mu_f(dx)F(x)G(S^kx) \xrightarrow[k\to\infty]{} \mu_f(F)\mu_f(G), \tag{6.2.16}$$

with exponential rate, for $k \to \infty$, if F, G are Hölder continuous functions on Ω.

Proof: We only have to check the exponential decay, which, however, follows by the same argument seen in (5.4.22). ■

An interesting "intrinsic" method to compute $p(f)$ and μ from the stability properties of the periodic points is given by the following proposition.

P6.2.1 **(6.2.1) Proposition:** (Periodic orbits expansion)
(i) The quantity

e6.2.17

$$p(f) = \max_{\mu\in\mathcal{M}^t_e(\Omega,S)} (s(\mu) - \mu(f)), \tag{6.2.17}$$

with $f \in C^\alpha(\Omega)$, can be computed as

e6.2.18
$$p(f) = \lim_{N\to\infty} N^{-1} \log \sum_{x\in\Omega,\, S^N x = x} e^{-\sum_{j=0}^{N-1} f(S^j x)}. \tag{6.2.18}$$

(ii) If μ_f is the point where the maximum in (6.2.17) is reached, μ_f can be computed as a weak limit as $N \to \infty$ of the distributions on Ω

e6.2.19
$$\mu_N(dy) = \frac{\sum\limits_{S^N x = x} \Big(\exp\big(-\sum\limits_{j=0}^{N-1} f(S^j x)\big)\Big)\delta_x(dy)}{\sum_{S^N x = x}\Big(\exp\big(-\sum_{j=0}^{N-1} f(S^j x)\big)\Big)}, \tag{6.2.19}$$

where δ_x is the Dirac measure on x.

Proof: Let $\mathcal{Q}$ be a Markovian pavement: we shall denote by $\pi_N \equiv \pi_N(\Omega, S)$ the set of the points of Ω periodic with period N and by $\widehat{\pi}_N \equiv \widehat{\pi}_N(\Omega, S)$ the set of the points of $\{1,\ldots,q\}_T^{\mathbb{Z}}$ periodic with period N. Then

e6.2.20
$$X(\widehat{\pi}_N) \subseteq \pi_N, \tag{6.2.20}$$

and to every $x \in \pi_N$ there can correspond at most d elements $\underline{\sigma} \in \widehat{\pi}_N$ such that $X(\underline{\sigma}) = x$, if d is the multiplicity M of the code, cf. definition (4.1.3). But there could exist elements of $\widehat{\pi}_N$ that are not symbolic representations of elements of π_N because the compatible histories of a point of period N might have a period multiple of N (as a consequence of the ambiguity of the coding of points that are on the boundaries of the elements of the Markovian pavement). Hence we can write the inequality

e6.2.21
$$\sum_{x\in\pi_N} e^{-\sum_{j=0}^{N-1} f(S^j x)} \ge d^{-1} \sum_{\underline{\sigma}\in\widehat{\pi}_N} e^{-\sum_{j=0}^{N-1} f(X(\tau^j \underline{\sigma}))}, \tag{6.2.21}$$

which, by (6.2.12), implies

e6.2.22
$$\sum_{x\in\pi_N} e^{-\sum_{j=0}^{N_1} f(S^j x)} \ge d^{-1} \sum_{\underline{\sigma}\in\widehat{\pi}_N} e^{-C''-Nf(X(\underline{\varrho}))-\sum_{j=0}^{N-1} A_\Phi(\tau^j\underline{\sigma})}. \tag{6.2.22}$$

This is a useful information because we can check that

e6.2.23
$$\lim_{N\to\infty} N^{-1} \log \sum_{\underline{\sigma}\in\widehat{\pi}_N} \exp\Big(-\sum_{j=0}^{N-1} A_\Phi(\tau^j \underline{\sigma})\Big) = P(\Phi); \tag{6.2.23}$$

indeed such a limit is certainly $\le P(\Phi)$ by virtue of (6.1.2) and of (5.1.12), (5.1.14) and (5.1.16), and because it is possible to establish, in an obvious way, a one-to-one correspondence between $\widehat{\pi}_N$ and a subset of $\{1,\ldots,q\}_T^N$.

On the other hand, since T is a mixing matrix with mixing time a, every configuration of $\{1,\ldots,q\}_T^{N-a}$ can be continued into a sequence in $\widehat{\pi}_N$, according to a prefixed rule to determine the interpolating strings $\underline{\vartheta}$. If $\underline{\sigma}'$ is any compatible string of length $N-a$, $\underline{\sigma}'=(\sigma'_0,\ldots,\sigma'_{N-a-1})$ we simply extend it to the right by the string $\underline{\vartheta}(\sigma'_{N-a-1},\sigma'_0)$ so that the string $\underline{\sigma}'\circ\underline{\vartheta}$ thus constructed has length N and can be indefinitely repeated to form a period N infinite sequence $\widetilde{\sigma}'$. Thus one sees that, with the notations of Section §6.1, see beginning of the proof of (6.1.2), there is a constant $\widehat{C}>0$ such that

e6.2.24
$$\big|U^{\Phi}_{[0,1,\ldots,N-a-1]}(\underline{\sigma}')-\sum_{j=0}^{N-1}A_\Phi(\tau^j\underline{\widetilde{\sigma}}')\big|\le\widehat{C},\tag{6.2.24}$$

so that

e6.2.25
$$\sum_{\underline{\sigma}'\in\{1,\ldots,n\}^{N-a}}e^{-U^{\Phi}_{[0,1,\ldots,N-a-1]}(\underline{\sigma}')}\le\sum_{\underline{\sigma}'\in\widehat{\pi}_N}e^{\widehat{C}-\sum_{j=0}^{N-1}A_\Phi(\tau^j\underline{\sigma}')},\tag{6.2.25}$$

which gives, together with (6.1.1), the inverse inequality necessary to deduce (6.2.23).

Having established (6.2.23), we note that (6.2.22), (6.2.23) and (6.1.29) imply, if $m\in G(\Phi)$ and μ is its image via X,

e6.2.26
$$\begin{aligned}&\lim_{N\to\infty}N^{-1}\log\sum_{x\in\pi_N}e^{-\sum_{j=0}^{N-1}f(S^jx)}\ge\\&\qquad\ge-f(X(\underline{\rho}))+s(m)-m(A_\Phi)=s(\mu)-\mu(f),\end{aligned}\tag{6.2.26}$$

having also used the general relation (6.2.15) involving a distribution $m'\in\mathcal{M}^t_e\ (\{1,\ldots,q\}_T^{\mathbb{Z}},\tau)$ and its X–image μ'.

To find the inequality opposite to (6.2.26) we proceed in an analogous way. Assuming for simplicity that N is even we associate with every $x\in\pi_N$ several elements $\underline{\sigma}\in\widehat{\pi}_{N+2a}$ such that

e6.2.27
$$d(X(\underline{\sigma}),x)\le\overline{C}e^{-\overline{\alpha}N},\tag{6.2.27}$$

for suitable $\overline{C},\ \overline{\alpha}>0$. The rule that allows us to construct this correspondence is the following. Let $\widehat{\underline{\sigma}}$ be one of the (up to d) elements in $\{1,\ldots,q\}_T^{\mathbb{Z}}$ such that $X(\underline{\sigma})=x$. Set

e6.2.28
$$\sigma_i=\widehat{\sigma}_i\qquad\forall|i|\le N/2,\tag{6.2.28}$$

and define $\sigma_{N/2+1},\ldots,\sigma_{N/2+a}$, $\sigma_{-N/2-1},\ldots,\sigma_{-N/2-a}$ so that

e6.2.29
$$\sigma_{\pm(N/2+a)}=0,\qquad\prod_{j=N/2}^{N/2+a-1}T_{\sigma_j\sigma_{j+1}}T_{\sigma_{-j}\sigma_{-j-1}}=1.\tag{6.2.29}$$

This is possible by the mixing property of T and in up to q^{2a} different ways.

The sequence $\underline{\sigma}$, defined in this way for $i \in [-N/2-a,\ N/2+a]$ is then periodically extended into a periodic sequence with period $N+2a$: note that this is now possible because $\sigma_{-N/2-a} = \sigma_{N/2+a}$ by (6.2.29).

The construction shows that with every $x \in \pi_N$ we can associate (up to q^{2a}) elements $\underline{\sigma} \in \widehat{\pi}_{2N/2+2a}$ and, furthermore, such classes of elements are pairwise disjoint (*i.e.* to different x necessarily there correspond different sequences $\underline{\sigma}$).

Furthermore (6.2.28) implies (6.2.27) by the Hölder continuity of the code, and (6.2.27) in turn means that if $\underline{\sigma}$ is associated with an x by the preceding construction there is a constant $\overline{C}' > 0$ such that

e6.2.30

$$\Big|\sum_{j=0}^{N-1} f(S^j x) - \sum_{j=0}^{N-1} f(X(\tau^j \underline{\sigma}))\Big| \le \overline{C}', \tag{6.2.30}$$

because f is Hölder continuous and, by (6.2.27), there exists $\overline{C}'' > 0$ such that

e6.2.31

$$\sum_{j=0}^{N-1} d(S^j x, S^j X(\underline{\sigma}))^\alpha \le \overline{C}'' \tag{6.2.31}$$

(note that the distance between $S^j x$ and $S^j X(\underline{\sigma})$ begins to grow substantially only when $|j|$ is close to $N/2$, by the construction of $\underline{\sigma}'$). Hence by (6.2.30)

e6.2.32

$$\begin{aligned}\sum_{\underline{\sigma}\in\widehat{\pi}_{N+2a}} & e^{-\sum_{j=0}^{N+2a-1} f(S^j X(\underline{\sigma}))} \ge \\ & \ge e^{-\overline{C}'-2a\|f\|_\infty} \cdot \sum_{x\in\pi_N} e^{-\sum_{j=0}^{N-1} f(S^j x)}.\end{aligned} \tag{6.2.32}$$

This inequality implies, by (6.2.23), the validity of the inequality opposite to (6.2.26), and (6.2.18) is proved. This shows that the limit (6.2.18) exists and equals $P(\Phi) - f(X(\underline{\rho}))$: in particular the latter expression does not depend on the Markovian pavement $\mathcal{Q}$ used, in spite of the fact that both Φ and the code X do depend on $\mathcal{Q}$.

To show the validity of item (ii) call $\overline{\mu}$ a weak limit of the sequence of distributions in (6.2.19). Since the distributions μ_N in (6.2.19) are invariant,[1]
N6.2.1 the limit of any convergent subsequence of μ_N will be S–invariant and hence $\overline{\mu}$ will be S–invariant.

To check that $\overline{\mu} = \mu$ it will suffice to show that $\overline{\mu}$ is absolutely continuous with respect to μ: then the ergodicity of μ will imply $\overline{\mu} = \mu$. The absolute continuity will follow (since μ and $\overline{\mu}$ are regular Borel measures) from an inequality of the type $\overline{\mu}(|F|) \le K\mu(|F|)$, for F that varies in a set which spans densely in $C(\Omega)$. For example for every Hölder continuous F.

---

[1] Observe that if $F(x) = \sum_{j=0}^{N-1} f(x)$ and $S^N x = x$ then $F(Sx) = F(x)$. Moreover the set of points of period N is clearly S-invariant.

Hence we shall fix $1 > \beta > 0$ and show the existence of $K > 0$ such that $\overline{\mu}(F) \le K\mu(F)$ for all $F \ge 0$, $F \in C^{\beta}(\Omega)$. We have

e6.2.33

$$\begin{aligned}
\mu_N(F) &= \frac{\sum_{x\in\pi_N} e^{-\sum_{j=0}^{N-1} f(S^j x)} F(x)}{\sum_{x'\in\pi_N} e^{-\sum_{j=0}^{N-1} f(S^j x')}} \ \le \ \text{by (6.2.22)} \ \le \\
&\le d\, e^{C''} \frac{\sum_{x\in\pi_N} e^{-\sum_{j=0}^{N-1} f(S^j x)} F(x)}{\sum_{\underline{\sigma}\in\widehat{\pi}_N} e^{-Nf(X(\rho))-\sum_{j=0}^{N-1} A_\Phi(\tau^j\underline{\sigma})}} \le \qquad (6.2.33)\\
&\le \ \text{by (6.2.27), (6.2.30), (6.2.32)} \ \le de^{C''+\overline{C}'+2a\|f\|_\infty}\cdot \\
&\cdot \frac{\sum_{\underline{\sigma}\in\widehat{\pi}_{N+2a}} e^{-\sum_{j=0}^{N+2a-1} f(X(\tau^j\underline{\sigma}))}\big(F(X(\underline{\sigma})) + \,\overline{\overline{C}}e^{-\,\overline{\overline{\alpha}}N}\big)}{\sum_{\underline{\sigma}\in\widehat{\pi}_N} e^{Nf(X(\underline{\rho}))-\sum_{j=0}^{N-1} A_\Phi(\tau^j\underline{\sigma})}},
\end{aligned}$$

having also used the Hölder continuity of F to define naturally $\overline{\overline{C}}$ and $\overline{\overline{\alpha}}$.
By (6.2.12), the (6.2.33) can be further bounded above by

e6.2.34

$$\begin{aligned}
\mu_N(F) &\le de^{2C''+\overline{C}'+2a\|f\|_\infty} \sum_{\underline{\sigma}\in\widehat{\pi}_{N+2a}} e^{-(N+2a)f(X(\underline{\rho}))}\cdot \\
&\cdot \frac{e^{-\sum_{j=0}^{N+2a-1} A_\Phi(\tau^j\underline{\sigma})}\big(F(X(\underline{\sigma})) + \,\overline{\overline{C}}e^{-\,\overline{\overline{\alpha}}N}\big)}{\sum_{\underline{\sigma}\in\widehat{\pi}_N} e^{-N\,f(X(\underline{\rho})-\sum_{j=0}^{N-1} A_\Phi(\tau^j\underline{\sigma}))}}, \qquad (6.2.34)
\end{aligned}$$

and, in the limit as $N \to \infty$, we get

e6.2.35

$$\begin{aligned}
&\overline{\mu}(F) \le de^{2C''+\overline{C}'+4a\|f\|_\infty} \lim_{N\to\infty} \\
&\frac{\sum_{\underline{\sigma}\in\widehat{\pi}_{N+2a-1}} e^{-\sum_{j=0}^{N+2a-1} A_\Phi(\tau^j\underline{\sigma})}\big(F(X(\underline{\sigma})) + \,\overline{\overline{C}}e^{-\alpha N}\big)}{\sum_{\underline{\sigma}\in\widehat{\pi}_N} e^{-\sum_{j=0}^{N-1} A_\Phi(\tau^j\underline{\sigma})}}. \qquad (6.2.35)
\end{aligned}$$

However as a consequence of the finiteness of $\|\Phi\|_1$, for all $p \ge 0$ there exists a constant $B_p < \infty$ such that

e6.2.36

$$B_p^{-1} \le \frac{\sum_{\underline{\sigma}\in\widehat{\pi}_N} e^{-\sum_{j=0}^{N-1} A_\Phi(\tau^j\underline{\sigma})}}{\sum_{\underline{\sigma}\in\widehat{\pi}_{N+p}} e^{-\sum_{j=0}^{N-1} A_\Phi(\tau^j\underline{\sigma})}} \le B_p, \qquad (6.2.36)$$

because this ratio can be estimated by the same procedure followed to solve problem [5.3.1] (we leave this to the reader).

Then the limit in (6.2.35) is not larger than $B_{2a}\mu(F)$, for all $F \in C^{\alpha}(\Omega)$, $F \ge 0$, and this implies existence of $K > 0$ such that $\overline{\mu}(F) \le K\mu(F)$, for all $F \in C^{\alpha}(\Omega)$, $F \ge 0$; and the proof is complete. ■

Bibliographical note to §6.2

This section illustrates some aspects of Sinai's theory, [Si72]; see the bibliographical note to §4.1.

§6.3 Periodic orbits, invariant probability distributions and entropy

An interesting and in a certain sense surprising corollary of the results of the preceding section is the following proposition.

P6.3.1 **(6.3.1) Proposition:** (Topological entropy)
Let (Ω, S) be an Anosov system and let $N_m(S) = \{$number of periodic points of period m contained in $\Omega\}$.
(i) The limit

e6.3.1
$$s_0 = \lim_{m\to\infty} m^{-1} \log N_m(S) \tag{6.3.1}$$

exists and has value $s_0 = s(\mu)$, where μ is a probability distribution which maximizes the entropy function $s(\cdot)$ in $\mathcal{M}_e^t(\Omega, S)$.
(ii) If $\mathcal{Q}$ is a Markovian pavement of Ω with compatibility matrix T and if λ_T is the largest eigenvalue of T (simple by Perron–Frobenius' theorem, cf. *problems [2.3.7], [2.3.10] and [2.3.12]), one has*

e6.3.2
$$s_0 = P(0) = \log \lambda_T > 0. \tag{6.3.2}$$

(iii) The closure of the set of periodic points coincides with Ω.

Proof: Property (iii) is true for the subshift $(\{1,\ldots,q\}_T^{\mathbb{Z}}, \tau)$ because T is transitive (see proposition (4.2.4)) and therefore it holds for (Ω, S) because the symbolic code X is continuous and surjective. A proof not based on symbolic dynamics is perhaps more instructive, see problem [4.2.13].

Property (i) holds because the probability distribution μ that maximizes $s(\mu')$ on $\mathcal{M}_e^t(\Omega)$ is such that $s_0 = s(\mu)$ by (6.2.18), while, by proposition (6.2.1), it is the image of the Gibbs distribution on $\{1,\ldots,q\}_T^{\mathbb{Z}}$ with potential $\Phi = 0$ (cf. proof of lemma (6.2.1), in particular (6.2.7) and (6.2.10)). Then

e6.3.3
$$\begin{aligned} P(0) &= \lim_{N\to\infty} N^{-1} \log \sum_{\underline{\sigma}\in\{1,\ldots,q\}_T^N} 1 = \\ &= \lim_{N\to\infty} N^{-1} \log \sum_{\sigma_1,\ldots,\sigma_N} T_{\sigma_1\sigma_2} T_{\sigma_2\sigma_3} \ldots T_{\sigma_{N-1}\sigma_N} = \\ &= \lim_{N\to\infty} N^{-1} \log \sum_{\sigma,\sigma'} (T^{N-1})_{\sigma\sigma'} = \log \lambda_T, \end{aligned} \tag{6.3.3}$$

by Perron–Frobenius' theorem, since T is mixing. Hence also (ii) follows. ■

The problem of the counting of periodic points can be further refined: for instance the following proposition holds (Manning–Ruelle).

P6.3.2 **(6.3.2) Proposition:** (Zeta function)
Let (Ω, S) be an Anosov system.
(i) The function

e6.3.4
$$\zeta(s,f) = \exp\Big[\sum_{n=1}^{\infty}\Big(\frac{e^{-ns}}{n}\sum_{T_x^n=x} e^{-\sum_{j=0}^{N-1} f(S^j x)}\Big)\Big] \tag{6.3.4}$$

is holomorphic in the complex plane in the variable s for $\mathrm{Re}\, s > p(f)$*, if $f \in C^\alpha(\Omega)$, $\alpha > 0$. Such a function can furthermore be extended to a holomorphic function in the half plane Re $s > p(f) - \varepsilon(f)$, with $\varepsilon(f) > 0$ suitable, deprived of the point $s = p(f)$, where a simple pole is present .*
(ii) The function $s \to \zeta(s, 0)$ extends to a meromorphic function in the complex plane s.

We shall not enter into the details of the proof of this proposition that requires an accurate analysis of the multiplicity of the code X associated with a Markovian pavement on Ω of the type of those built in Section §4.2, at least for what concerns the periodic points $x \in \Omega \cap \partial\mathcal{Q}$, [Ma71].

Another question, implicitly solved in the previous sections, is the existence of invariant probability distributions which are absolutely continuous with respect to the volume measure on Ω: we shall only formulate the results in the case in which Ω is two–dimensional, in order to make use of the analysis in Section §4.3. It is however true that both the results of Section §4.3 and those that follow can be extended to Anosov systems with dimension larger than 2.

P6.3.3 **(6.3.3) Proposition:** (SRB distribution)
Under the hypotheses of lemma (6.2.1) there exist two probability distributions μ^+, $\mu^- \in \mathcal{M}_e^t(\Omega, \tau)$ such that for every $F \in C^\alpha(\Omega)$, $\alpha > 0$,

e6.3.5
$$\lim_{N\to\infty} N^{-1}\sum_{j=0}^{N-1} F(S^{\pm j}x) = \int \mu^{\pm}(d\zeta)F(\zeta) \tag{6.3.5}$$

almost everywhere in x with respect to the volume measure μ_0 on Ω.
The dynamical system (Ω, S) admits an ergodic topological probability distribution μ absolutely continuous with respect to the volume measure if and only if $\mu^+ = \mu^-$. In such a case $\mu_0 = \mu^+ = \mu^-$.

Remarks: (1) The above theorem is due to Sinai, Ruelle, and Bowen, [Si68], [Si72], [Si77], [Ru76], [Ru78], [Ru95], [Ru99] and [Bo70]. The difference between this theorem and Birkhoff's theorem, see proposition (2.2.2), is striking. Here the distribution with which the initial data x are sampled *is not invariant, i.e.* x is chosen almost everywhere with respect to μ_0 rather than with respect to an invariant probability distribution.
(2) The above proposition says that the statistical properties of motions whose initial data are chosen randomly with a distribution proportional

to the volume measure on phase space *exist and are independent from the choice of the data*. Such a property is often taken for granted in Statistical Mechanics problems and it is formally called the 0-*th law of statistical mechanics* and $\mu^\pm$ are called the *statistics* of the motions generated by the dynamical system; this *attributes a philosophical primacy to the random choices based on distributions with density with respect to the volume of phase space.*
(3) Note that in general the statistics "toward the future" (μ^+) and the statistics "toward the past" (μ^-) are *different.*
(4) A system (Ω, S) is called *reversible* if there exists an isometry $I : \Omega \to \Omega$ such that

e6.3.6 $$IS = S^{-1}I, \qquad I^2 = 1. \tag{6.3.6}$$

This symmetry, which *ought not be confused with invertibility* which is just the existence of S^{-1}, exists in many concrete applications and it often coincides with the velocity reversal of the constituents of a mechanical system. What might be perhaps surprising is that even in systems in which time reversal symmetry holds one has, in general, $\mu^+ \neq \mu^-$.

Proof: From proposition (4.3.2) it follows that the probability distribution m_0, image via the code X associated with the Markovian pavement $\mathcal{Q}$, of the volume measure μ_0 on Ω, is a probability distribution with conditional probabilities given by (4.3.11). The restriction m_0^+ of m_0 to the σ-algebra $\mathcal{B}(\mathbb{Z}^+)$ has instead conditional probabilities given by (4.3.12).
Let m_u be the Gibbs distribution on $\{1, \ldots, q\}_T^{\mathbb{Z}}$ with potential Φ^u described, according to the notations of (4.3.13), by $\Phi^u_X = 0$ unless X is a set translated of $\{0, \ldots, 2n+1\}$ for some n, and define

e6.3.7 $$\Phi^u_{2n+1}(\sigma_0, \ldots, \sigma_{2n+1}) \equiv \Phi^u_{(0,\ldots,2n+1)}(\sigma_0, \ldots, \sigma_{2n+1}). \tag{6.3.7}$$

Such a probability distribution exists and is unique, ergodic, mixing with exponential rate on the cylinders (because Φ^u verifies (4.3.14) and, therefore, the bound (5.4.19) and its consequence (5.4.22), for the same reason discussed there).
We apply again the methods and ideas employed to study Gibbs distributions in Section §5.2, see in particular the proof of proposition (5.2.1).
The probability distribution m_u^+, restriction of m_u to the σ–algebra $\mathcal{B}(\mathbb{Z}^+)$, and the probability distribution m_0^+, restriction to $\mathcal{B}(\mathbb{Z}^+)$ of the probability distribution m_0 image via X of the volume measure μ_0, are absolutely equivalent. This can be verified by using (4.3.12), (4.3.13), (4.3.14) and (5.2.5) (which is, with the obvious modifications, applicable also to the conditional probability, (4.3.11), of m_0): in fact one has

e6.3.8 $$\begin{aligned}\frac{m_0^+(C^{0\ldots N}_{\sigma_0\ldots\sigma_N})}{m_u^+(C^{0\ldots N}_{\sigma_0\ldots\sigma_N})} &= \frac{\int m_0(\sigma_0\ldots\sigma_N|\widetilde\sigma_{N+1}\ldots,\widetilde\sigma_{-1}\ldots)m_0(d\underline{\widetilde\sigma})}{\int m_u(\sigma_0\ldots\sigma_N|\widetilde\sigma_{N+1}\ldots,\widetilde\sigma_{-1}\ldots)m_u(d\underline{\widetilde\sigma})} \le \qquad (6.3.8)\\ &\le (q^a\exp(8\|\Phi^u\|_1))^2\Delta\frac{\int m_0(\sigma_0\ldots\sigma_N|\widehat\sigma_{N+1}\ldots,\widetilde\sigma_{-1}\ldots)m_0(d\underline{\widetilde\sigma})}{\int m_u(\sigma_0\ldots\sigma_N|\widehat\sigma_{N+1}\ldots,\widetilde\sigma_{-1}\ldots)m_u(d\underline{\widetilde\sigma})},\end{aligned}$$

N6.3.1 where a is the mixing time of T and[1] $\Delta = \max_{\underline{\sigma},\underline{\sigma}'} \left|\frac{\sin\varphi(X(\underline{\sigma}))}{\sin\varphi(X(\underline{\sigma}'))}\right| \le \frac{1}{\min\limits_{x\in\Omega}|\sin\varphi(x)|}$ and $\underline{\widehat{\sigma}}$ is a reference configuration arbitrarily chosen among those that continue $\ldots\underline{\widetilde{\sigma}}_{-1}\ \sigma_0\ldots\sigma_N$ into a configuration $\ldots\underline{\widetilde{\sigma}}_{-1}\ \sigma_0\ldots\sigma_N\ \widehat{\sigma}_{N+1}\ldots$: the $\underline{\widehat{\sigma}}$ is chosen so that it depends exclusively on σ_N and not, in particular, on $\underline{\widetilde{\sigma}}$, see the beginning of the proof of lemma (6.2.1). In (6.3.8) all integrals must be understood to be extended to regions such that the configurations that appear in the conditional probabilities are in $\{1,\ldots,q\}^{\mathbb{Z}}_T$.

Since the sequence $\underline{\widehat{\sigma}}$ is fixed the integrals in (6.3.8) can be immediately performed and (6.3.8) becomes

e6.3.9
$$\frac{m_0^+(C^{0\ldots N}_{\sigma_0\ldots\sigma_N})}{m_u^+(C^{0\ldots N}_{\sigma_0\ldots\sigma_N})} \le K_1 \frac{m_0^+(\sigma_0\ldots\sigma_N|\widehat{\sigma}_{N+1}\ldots)}{m_u^+(\sigma_0\ldots\sigma_N|\widehat{\sigma}_{N+1}\ldots)}, \tag{6.3.9}$$

for a suitable constant K_1. The probability distribution m_0 can be treated as the probability distribution m of Section §5.2 and we can prove for it propositions analogous to propositions (5.3.1), (5.3.2) and corollary (5.3.1) for which the probability distribution m_0^+ turns out to be equivalent (with a bound for the Radon–Nykodim derivative similar to (5.3.10)) to the probability distribution $\widetilde{m}_u^+$, a Gibbs distribution on $\mathbb{Z}^+$ with potential Φ^u. Since also m_u^+ is, by corollary (5.3.1), absolutely continuous with respect to $\widetilde{m}_u^+$ and (5.3.10) holds, the r.h.s. of (6.3.9) can be bounded by a constant K_2. By what said above one deduces that we can choose

e6.3.10
$$K_2 = e^{8\|\Phi^u\|_1 + 4a\|\Phi^u\|}\Delta\, q^{2a} K_1, \tag{6.3.10}$$

and, in conclusion,

e6.3.11
$$\frac{m_0^+(C^{0\ldots N}_{\sigma_0\ldots\sigma_N})}{m_u^+(C^{0\ldots N}_{\sigma_0\ldots\sigma_N})} \le K_2 \text{ and, likewise, } \ge K_2^{-1}. \tag{6.3.11}$$

Hence m_0^+ is absolutely continuous and equivalent to m_u^+ and $\pi_u \overset{def}{=} \frac{dm_0^+}{dm_u^+}$ is a $\mathcal{B}(\mathbb{Z}^+)$–measurable function such that $K_2^{-1} \le \pi_u(\underline{\sigma}) \le K_2$, m_0^+–almost everywhere. In fact with a little extra effort and by using the ideas of Section §5.2 and of the proof of proposition (5.3.2) in particular, it is possible to show that π_u can be chosen to be Hölder continuous on $\{1,\ldots,q\}^{\mathbb{Z}^+}_T$ (*i.e.* $|\pi_u(\underline{\sigma}) - \pi_u(\underline{\sigma}')| \le C\, d(\underline{\sigma},\underline{\sigma}')^\alpha$ for all $\underline{\sigma},\underline{\sigma}' \in \{1,\ldots,q\}^{\mathbb{Z}^+}_T$, , with C, $\alpha > 0$): we leave this derivation to the reader.

If $\underline{\sigma} \to \widehat{F}(\underline{\sigma})$ is a $\mathcal{B}(\mathbb{Z}^+)$–measurable function on $\{1,\ldots,q\}^{\mathbb{Z}}_T$ we shall have m_u^+–almost everywhere and, therefore, m_0^+–almost everywhere and

---

[1] Note that in the course of the proof of proposition (4.3.2) we deduced an explicit expression for the function f that appears in (4.3.12), *i.e.* $f(\underline{\sigma}) = \sin\varphi(X(\underline{\sigma}))$, where $\varphi(X(\underline{\sigma}))$ denotes the angle between the direction of the manifold stable and that of the unstable manifold at the point $x = X(\underline{\sigma})$. Such an expression is here used to get the bound (6.3.8).

m_0–almost everywhere

e6.3.12
$$N^{-1}\sum_{j=0}^{N-1}\widehat F(\tau^j\underline\sigma)\xrightarrow[N\to\infty]{}\int\widehat F(\underline\sigma')m_u^+(d\underline\sigma')=\int\widehat F(\underline\sigma')m_u(d\underline\sigma'),\qquad(6.3.12)$$

by Birkhoff theorem and by the ergodicity of m_u.

Hence, by a density argument, the limit (6.3.12) is valid for all $F\in C(\{1,\ldots,q\}_T^{\mathbb Z})$.[2] N6.3.2

If $F\in C^\alpha(\Omega)$, $\alpha>0$, and μ^+ is the image (isomorphic $\mod 0$ to m_u) via the code X, we shall have that the function $\widehat F(\underline\sigma)=F(X(\underline\sigma))$ is in $C(\{1,\ldots,q\}_T^{\mathbb Z})$ and, by (6.3.12), if $x=X(\underline\sigma)$

e6.3.13
$$\begin{aligned}N^{-1}\sum_{j=0}^{N-1}F(S^jx)=N^{-1}\sum_{j=0}^{N-1}\widehat F(\tau^j\underline\sigma)\xrightarrow[N\to\infty]{}\\ \xrightarrow[N\to\infty]{}\int\widehat F(\underline\sigma')m_u(d\underline\sigma')=\int F(x')\mu^+(dx')\end{aligned}\qquad(6.3.13)$$

μ_0–almost everywhere.

Similar arguments can be given for the averages in the past and lead to the probability distribution μ^-.

The second part of the proposition is a consequence of the first part. ■

Note that in the preceding proof we obtained something more.

Let indeed $\widehat F$, $\widehat G$ be a pair of bounded and $\mathcal B(\mathbb Z^+)$–measurable functions; then, for all $j\ge0$, obviously

e6.3.14
$$\begin{aligned}&\int\widehat F(\tau^j\underline\sigma)\widehat G(\underline\sigma)m_0(d\underline\sigma)=\int\widehat F(\tau^j\underline\sigma)\widehat G(\underline\sigma)m_0^+(d\underline\sigma)=\\ &=\int\widehat F(\tau^j\underline\sigma)\widehat G(\underline\sigma)\pi_u(\underline\sigma)m_u^+(d\underline\sigma)=\int\widehat F(\tau^j\underline\sigma)\widehat G(\underline\sigma)\pi_u(\underline\sigma)m_u(d\underline\sigma),\end{aligned}\qquad(6.3.14)$$

where π_u is the function, Hölder continuous on $\{1,\ldots,q\}_T^{\mathbb Z^+}$, defined after the (6.3.11). Since m_u is mixing we obtain

e6.3.15
$$\lim_{j\to\infty}m_0(\widehat G\tau^j\widehat F)=m_u(\pi_u\widehat G)m_u(\widehat F)=m_0(\widehat G)m_u(\widehat F),\qquad(6.3.15)$$

and such a limit is reached at exponential rate if $\widehat F$ and $\widehat G$ are also Hölder continuous. All this follows easily from (5.4.22) which holds for m_u as noted after (6.3.7). By density this remains true for every $\widehat F$, $\widehat G$ which are Hölder continuous and $\mathcal B(\mathbb Z)$–measurable.

---

[2] Note that the assumption for $\widehat F$ to be $\mathcal B(\mathbb Z^+)$-measurable can be obviously weakened into $\mathcal B([k,+\infty))$-measurable, for some $k\in\mathbb Z$, and such continuous functions are dense in $C(\{1,\ldots,q\}_T^{\mathbb Z})$.

Therefore if $F, G \in C^\alpha(\Omega)$ and if we set $\widehat{F}(\underline{\sigma}) = F(X(\underline{\sigma}))$, $\widehat{G}(\underline{\sigma}) = G(X(\underline{\sigma}))$ we see that this gives the following.

$C6.3.1$ **(6.3.1) Corollary:** (Mixing of the non invariant volume distributions) *If $F,\ G \in C^\alpha(\Omega)$, $\alpha > 0$, one has*

$e6.3.16$
$$\lim_{j\to\pm\infty} \mu_0(GS^jF) = \mu_0(G)\mu_\pm(F) \tag{6.3.16}$$

and the limits are attained with exponential rate.

Finally the following proposition is remarkable.

$P6.3.4$ **(6.3.4) Proposition:** (Pesin formula) *Let (Ω, S) be an Anosov system. With the usual notations of definition (4.3.1) we have*

$e6.3.17$
$$s(\mu^+) - \mu^+(\log\lambda_u) = s(\mu^-) + \mu^-(\log\lambda_s) = 0 \tag{6.3.17}$$

Furthermore μ^+ and μ^- maximize in $\mathcal{M}_e^t(\Omega, S)$, respectively, the functionals $s(\mu) - \mu(\log\lambda_u)$ and $s(\mu) + \mu(\log\lambda_s)$.

Proof (hint): The second statement follows immediately because μ^+ and μ^- are X–images of Gibbs distributions with potentials Φ^u and Φ^s defined via (4.3.5), as shown by proposition (4.3.2) (cf. (4.3.15)) and by the proof of proposition (6.3.3): this statement is then reduced to the variational principle for Gibbs distributions, discussed in the previous sections.

To check (6.3.17) note that, by (6.3.11), if $P^{0\ldots N}_{\sigma_0\ldots\sigma_N} = \bigcap_{j=0}^{N} S^{-j}Q_{\sigma_j}$,

$e6.3.18$
$$K_2^{-1} \le \frac{m_u^+(C^{0\ldots N}_{\sigma_0\ldots\sigma_N})}{\mu_0(P^{0\ldots N}_{\sigma_0\ldots\sigma_N})} \le K_2, \tag{6.3.18}$$

and, therefore, to compute the value of $s(m_u) = s(\mu^+)$ we can make use of Shannon-McMillan's theorem and, by (6.3.18), we deduce that

$e6.3.19$
$$s(\mu^+) = s(m_u) = -\lim_{N\to\infty} N^{-1}\log\mu_0(C^{0\ldots N}_{\sigma_0\ldots\sigma_N}), \tag{6.3.19}$$

where the convergence of $f_N(\underline{\sigma}) = -N^{-1}\log\mu_0(C^{0\ldots N}_{\sigma_0\ldots\sigma_N})$ to $s(\mu^+)$ takes place in $L_1(m_u)$; see lemma (3.2.1).

Let $\underline{\sigma} \in \{1, \ldots, q\}^{\mathbb{Z}}_T$ and $x = X(\underline{\sigma})$. It is possible to see, following the methods used to obtain the estimates of Section §4.3 that there exists K' such that

$e6.3.20$
$$(K')^{-1} \le \frac{\mu_0(P^{0\ldots N}_{\sigma_0\ldots\sigma_N})}{\prod_{j=0}^{N-1}\lambda_u^{-1}(S^jX(\underline{\sigma}))} \le K' \tag{6.3.20}$$

(this is first heuristically shown by interpreting it geometrically with the help of a drawing like Fig. (4.3.1) and Fig. (4.3.4)).

Equation (6.3.20) implies immediately, by Shannon–McMillan's and Birkhoff's theorems, by proposition (6.3.3) and by considering the logarithm divided by N of both sides of (6.3.20), that the entropy $s(\mu^+) = s(m_u)$ is

e6.3.21
$$s(\mu^+) = s(m_u) = \lim_{N\to\infty} -N^{-1} \log \mu_0(P^{0\ldots N}_{\sigma_0\ldots\sigma_N}) = \tag{6.3.21}$$
$$= \lim_{N\to\infty} -N^{-1} \sum_{j=0}^{N-1} \log \lambda_u^{-1}(S^j x) = \int \mu^+(dx) \log \lambda_u^{-1}(x),$$

so that (6.3.17) follows. ■

We have therefore proved also the following corollary.

C6.3.2 **(6.3.2) Corollary:** *Let (Ω, S) be an Anosov system. Suppose the the map S preserves the volume measure μ_0 on Ω then*

e6.3.22
$$s(\mu_0) = -\int \mu_0(dx) \log \lambda_u(x), \tag{6.3.22}$$

i.e. "the entropy of an absolutely continuous invariant distribution is the average value of the logarithm of the contraction coefficient".

Remarks: (1) Note the relation between this theorem and the theorem of Kouchnirenko of Section §3.4.

(2) Hence the dynamical system on $\mathbb{T}^2$ defined by the map $S\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix} = \begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix} \mod 2\pi$ is such that the entropy of the Lebesgue measure μ_0 is

e6.3.23
$$s = \log(3+\sqrt{5})/2 = \log\lambda. \tag{6.3.23}$$

(3) Furthermore μ_0 is the probability distribution that maximizes the functional $s(\mu') - \mu'(\log\lambda)$ in the space of the ergodic probability distributions on $\mathbb{T}^2$. This means that it makes maximal the entropy, since λ is constant.
(4) The number of the S–periodic points is (cf. proposition (6.3.1))

e6.3.24
$$N_m(s) = e^{s\,m+o(m)} = \Big(\frac{3+\sqrt{5}}{2}\Big)^{m+o(m)}. \tag{6.3.24}$$

(5) This means that every Markovian pavement of $(\mathbb{T}^2, S)$ has a compatibility matrix T with the largest eigenvalue equal to $\frac{3+\sqrt{5}}{2} > 2$. See also the problems below.
(6) Hence a Markovian pavement for the system $(\mathbb{T}^2, S)$ consisting of only two S–rectangles cannot exist.

A representation of the SRB distribution similar to the one for the volume measure discussed in corollary (4.3.1) follows from the argument leading to (4.3.23) and from the analysis of Sections 6.2, 6.3 and the following proposition will be useful.

C6.3.3 **(6.3.3) Corollary:** (Fubini's theorem for the SRB distribution) *Let (M, S) be a two-dimensional Anosov map. Given $x_0 \in M$ consider the rectangle $R = W^s_\delta(x_0) \times W^u_\delta(x_0)$ where δ is small enough so that the point $[x, y]$ is uniquely defined (*cf.* Fig.(4.2.1)). Then the SRB probability $\mu(E)$ of a subset $E \subset R$ is given by*

e6.3.25

$$\begin{aligned}\mu(E) = &\int_{W^u_\delta(x_0)} d\sigma_x \int_{W^s_\delta(x_0)} \nu(d\sigma_y)\cdot \\ &\cdot \prod_{i=1}^{\infty} \frac{\lambda_u^{-1}(S^{-i}[y, x])}{\lambda_u^{-1}(S^{-i}x)} \chi_E([y, x])\end{aligned} \tag{6.3.25}$$

where $d\sigma_x$ is the area measure on $W^u_\delta(x_0)$ and $\nu(d\sigma_y)$ is a measure on $W^s_\delta(x_0)$.

The measure ν will in general *not* be absolutely continuous with respect to the arc length $d\sigma_y$. This result in fact is general and it holds in any dimension.

Problems for §6.3

Q6.3.1 **[6.3.1]:** Let (Ω, S) be an Anosov system and let $\mathcal{Q}$ be a pavement of Ω with S–rectangles (cf. definition (4.2.2)) which have the Markov property (4.2.12). If $\mathcal{Q}$ is generating then we could conclude that the compatibility matrix, see Section §4.1, defined by $T_{\sigma\sigma'} = 1$ if $\mathrm{int}(SQ_\sigma) \cap \mathrm{int}(Q_{\sigma'}) \neq \emptyset$ and $T_{\sigma\sigma'} = 0$ otherwise, has the logarithm of the largest eigenvalue equal to the topological entropy by proposition (6.3.1). Suppose that $\mathcal{Q}$ is non-generating: show that the logarithm of the largest eigenvalue is less than (or equal to) the topological entropy. However it cannot differ from the topological entropy by more than the logarithm of the maximum number of connected components of the sets $\mathrm{int}(Q_\sigma) \cap \mathrm{int}(Q_{\sigma'})$. (*Hint:* If $\mathcal{Q}$ is non-generating but it is not trivial then the correspondence between compatible symbolic sequences and points whose trajectories never fall on the boundaries will not be one-to-one but it will have multiplicity.)

Q6.3.2 **[6.3.2]:** Show that corollary (6.3.2) implies that the matrix $S_0 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ and the matrix of problem [4.3.7] have the same largest eigenvalue, *i.e.* $(1 + \sqrt{5})/2$, cf. Fig.(4.3.4).

Q6.3.3 **[6.3.3]:** Check that the matrix S_0 and the matrix of problem [4.3.7] have the same largest eigenvalue by directly computing it.

Bibliographical note to §6.3

This section illustrates some aspects of Sinai's theory, [Si72], [Si77]; see the bibliographical notes to §4.1.

§6.4 Equivalent potentials. Gibbs distributions with transitive vacuum

Let B be the space of potentials for $\{0, \ldots, n\}^{\mathbb{Z}}_T$, where T is a mixing compatibility matrix (cf. definitions (4.1.1)) and (5.1.1).

A natural question is whether it is possible that two potentials Φ and Ψ generate the same Gibbs states and, in the affirmative case, which is the relation between Φ and Ψ.

A fully satisfactory answer does not seem to exist, partly because the space B is not a "natural" space of potentials. It is just a family of conveniently restricted potentials large enough to give rise to a theory endowed of a remarkable variety of phenomena and singularities.

Nevertheless the following propositions will give an idea of what it means that Φ and Ψ produce the same Gibbs states: if this happens we say that the potential Φ and Ψ are *equivalent potentials.*

*P*6.4.1 **(6.4.1) Proposition:** (Equivalent potentials)
Let $B^h_s = \cup_{\kappa > s} B^{(\kappa)}$, where $s \ge 0$, and $B^{(\kappa)} \subset B$ is the space of the potentials $\Phi \in B$ such that

*e*6.4.1
$$\|\Phi\|^{(\kappa)} = \sum_{X \ni 0} e^{\kappa\,\mathrm{diam}(X)} \|\Phi_X\| < +\infty. \tag{6.4.1}$$

The B^h_s is called the space of the potentials that "decay exponentially" on a scale shorter than s^{-1}.[1] *Then* *N*6.4.1
(i) If, given Φ, $\Psi \in B^h_0$, there exists a constant C and a Hölder continuous function $F \in C(\{0, \ldots, n\}^{\mathbb{Z}}_T)$ such that

*e*6.4.2
$$A_\Phi(\underline{\sigma}) = A_\Psi(\underline{\sigma}) + F(\tau\underline{\sigma}) - F(\underline{\sigma}) + C \tag{6.4.2}$$

holds then Φ and Ψ satisfy $G(\Phi) = G(\Psi)$.
(ii) Viceversa if Φ, $\Psi \in B^h_0$ and $G(\Phi) = G(\Psi)$ then there exists a constant C and a Hölder continuous function $F \in C(\{0, \ldots, n\}^{\mathbb{Z}}_T)$ such that (6.4.2) holds.

Proof: Since (6.4.2) implies that for all invariant probability distributions $\mu \in M(\{0, \ldots, n\}^{\mathbb{Z}}_T)$ [2] one has $s(\mu) - \mu(A_\Phi) = s(\mu) - \mu(A_\Psi) + C$, item (i) follows from the variational principle (corollary (6.1.1)). *N*6.4.2

To prove (ii) suppose that Φ, $\Psi \in B^h_0$ and that there exists $C \in \mathbb{R}$ for which

*e*6.4.3
$$\mu(A_\Phi) = \mu(A_\Psi) + C \qquad \text{for all}\, \mu \in M(\{0, \ldots, n\}^{\mathbb{Z}}_T), \tag{6.4.3}$$

We first prove that, under the assumption (6.4.3) there exists $F \in C(\{0, \ldots, n\}^{\mathbb{Z}}_T)$, Hölder continuous, for which (6.4.2) holds.

In fact note that the function $g = A_{\Phi-\Psi} - C$ has zero integral with respect to an arbitrary $\mu \in M(\{0, \ldots, n\}^{\mathbb{Z}}_T)$ and hence, in particular,

*e*6.4.4
$$\sum_{j=0}^{N-1} g(\tau^j \underline{\sigma}) = 0 \tag{6.4.4}$$

---

[1] Note that one has $B^h_0 \subset B$; the superscript h refers to the fact that the function $A_\Phi(\underline{\sigma})$ corresponding to any $\Phi \in B^h_s$, with $s \ge 0$, is Hölder-continuous.

[2] See remark to proposition (2.3.1) for notations.

for every $\underline{\sigma}$ that belongs to the set of the periodic points with period N in
N6.4.3 $\{0,\ldots,n\}^{\mathbb{Z}}_T$.[3]

Furthermore the hypothesis of mixing on T guarantees the existence of $\underline{\sigma}_0$ such that $\cup_{j=0}^{\infty}\{\tau^j\underline{\sigma}_0\}$ is dense in $\{0,\ldots,n\}^{\mathbb{Z}}_T$ (cf. problem [4.1.19]).

We can then set

e6.4.5
$$\begin{aligned}&F(\underline{\sigma}_0)=0,\ F(\tau\underline{\sigma}_0)=g(\underline{\sigma}_0),\ F(\tau^2\underline{\sigma}_0)=g(\underline{\sigma}_0)+g(\tau\underline{\sigma}_0),\\ &\ldots,F(\tau^h\underline{\sigma}_0)=\sum_{j=0}^{h-1}g(\tau^j\underline{\sigma}_0),\ldots\end{aligned}\tag{6.4.5}$$

We now show that the function F thus defined on $\cup_{j=0}^{\infty}\{\tau^j\underline{\sigma}_0\}$ is in fact uniformly Hölder continuous on this set (*i.e.* the Hölder constant and the exponent can be chosen independently the point in $\cup_{j=0}^{\infty}\{\tau^j\underline{\sigma}_0\}$) and *therefore it extends to a continuous function* on all $\{0,\ldots,n\}^{\mathbb{Z}}_T$ that, by construction, verifies $g(\underline{\sigma})=F(\tau\underline{\sigma})-F(\underline{\sigma})$, and this implies (6.4.2).

To see the uniform Hölder continuity we remark that one has $A_{\Phi-\Psi}(\underline{\sigma})=\sum_{X\ni 0}\frac{\Phi_X(\underline{\sigma}_X)-\Psi_X(\underline{\sigma}_X)}{|X|}$ and recall that the definition of distance $d(\underline{\sigma},\underline{\sigma}')$ between two sequences $\underline{\sigma},\underline{\sigma}'$ is $d(\underline{\sigma},\underline{\sigma}')=e^{-\nu}$ if $\sigma_j=\sigma'_j$, for all $|j|\le\nu$ and $\sigma_{\pm(\nu+1)}\ne\sigma'_{\pm(\nu+1)}$ (see definition (4.1.3)). If $\underline{\sigma}$ and $\underline{\sigma}'$ are close then (6.4.1) implies existence of two constants $\kappa'>0$ and $C>0$ such that

e6.4.6
$$|g(\underline{\sigma})-g(\underline{\sigma}')|\le C\ d(\underline{\sigma},\underline{\sigma}')^{\kappa'}.\tag{6.4.6}$$

Then if $\tau^k\underline{\sigma}_0$ and $\tau^h\underline{\sigma}_0$ happen to be very close, *i.e.* if $(\tau^k\underline{\sigma}_0)_i=(\tau^h\underline{\sigma}_0)_i$ for all $|i|\le\nu$, and if $p=k-h\ge 0$ we set $\underline{\sigma}_1=\tau^h\underline{\sigma}_0$ and note that

e6.4.7
$$|F(\tau^k\underline{\sigma}_0)-F(\tau^h\underline{\sigma}_0)|=\Big|\sum_{j=0}^{p-1}g(\tau^j\underline{\sigma}_1)\Big|.\tag{6.4.7}$$

The definitions imply that $(\underline{\sigma}_1)_i=(\underline{\sigma}_1)_{i+p}$, for all $|i|\le\nu$, and hence there exists a configuration $\widehat{\underline{\sigma}}_1\in\{0,\ldots,n\}^{\mathbb{Z}}_T$ which is periodic with period p and such that

e6.4.8
$$(\widehat{\underline{\sigma}}_1)_j=(\underline{\sigma}_1)_j\qquad-\nu\le j\le\nu,\tag{6.4.8}$$

(this is $\widehat{\sigma}_{1j}=\sigma_{1j}$ for $-\nu\le j<-\nu+p$ and then $\widehat{\underline{\sigma}}$ is continued periodically). So that, by (6.4.4) and (6.4.7),

e6.4.9
$$|F(\tau^k\underline{\sigma}_0)-F(\tau^h\underline{\sigma}_0)|\le\sum_{j=0}^{p-1}|g(\tau^j\underline{\sigma}_1)-g(\tau^j\widehat{\underline{\sigma}}_1)|\le\frac{2C_{\kappa'}e^{-\kappa'\nu}}{1-e^{-\kappa'}}.\tag{6.4.9}$$

Therefore F is uniformly Hölder continuous on $\cup_{j=0}^{\infty}\{\tau^j\underline{\sigma}_0\}$ and (6.4.2) is proved under the assumption in (6.4.3).

---

[3] The measure $\mu=N^{-1}\sum_{j=0}^{N-1}\delta_{\tau^j\underline{\sigma}}$ is in $M(\{0,\ldots,n\}^{\mathbb{Z}}_T)$!

To prove the statement in item (ii) we must show that (6.4.3) follows from the assumptions that $\Phi,\Psi\in B_0^h$ and that $G(\Phi)=G(\Psi)$. This means that the tangent planes to the graph of the pressure $P(\Phi)$ in Φ and in Ψ (which are unique, cf. propositions (5.2.1), (6.1.2)) coincide. Since $P(\Phi)$ is convex this means that they are also tangent to P at the points $P(\Phi+t(\Psi-\Phi))$ for $t\in[0,1]$: hence *for all t real*, because the DLR relations are real analytic in the potentials. Let $C=\mu(\Psi-\Phi)$: the variational principle (cf. corollary (6.1.1)) then tells us that

e6.4.10

$$\begin{aligned}&\max_{\nu\in M(\{0,\ldots,n\}_T^{\mathbb{Z}})}\big(s(\nu)-\nu(A_{\Phi+t\,(\Psi-\Phi)})\big)=\\&=s(\mu)-\mu(A_{\Phi+t\,(\Psi-\Phi)})=P(\Phi)-Ct\end{aligned}\tag{6.4.10}$$

Noting that $s(\nu)-\nu(A_{\Phi+t\,(\Psi-\Phi)})\equiv s(\nu)-\nu(A_\Phi)-t\nu(A_{\Psi-\Phi})\le P(\Phi)-Ct$ for all t real and for all $\nu\in M(\{0,\ldots,n\}_T^{\mathbb{Z}})$ can only be if $\nu(A_{\Psi-\Phi})\equiv C$ we get $A_\Phi-A_\Psi=C+G$ with $\nu(G)\equiv 0$ *for all* $\nu\in M(\{0,\ldots,n\}_T^{\mathbb{Z}})$. Hence we see that (6.4.3) holds. ■

Remark: We can define on B_0^h an equivalence relation setting $\Phi\approx\Psi$ if there exist $F\in C(\{0,\ldots,n\}_T^{\mathbb{Z}})$ and a constant $C\in\mathbb{R}$ for which (6.4.2) holds. A norm on the equivalence classes defined by this relation is given by

e6.4.11

$$\|\Phi\|=\inf_{C,F}\sup_{\underline{\sigma}}|A_\Phi(\underline{\sigma})+C+F(\tau\underline{\sigma})-F(\underline{\sigma})|,\tag{6.4.11}$$

It would be natural to call "space of the potentials" the space obtained by taking the closure B' of the classes of B_0^h with respect to the norm (6.4.11).

But this new space B' of (equivalence classes of) potentials is difficult to study and it is not clear not only whether $\Phi\ne\Psi$ implies $G(\Phi)\ne G(\Psi)$ for Φ, $\Psi\in B'$, but even whether we could extend the notion of Gibbs distribution to the elements Φ of the completion of the classes of B_0^h.

Nevertheless the idea of choosing one representative for every class of equivalent potentials can, in several important cases, be realized. For example we consider the following case.

C6.4.1 **(6.4.1) Corollary:** (Potentials in systems with transitive vacuum)
If $T_{0\sigma}=T_{\sigma 0}=1$ for all $\sigma=0,\ldots,n$, ("transitivity of T on the vacuum") then every $\Psi\in B^h_{\log 2}$ is equivalent to a potential $\Phi\in B_0^h$ such that $\Phi_X(\underline{\sigma}_X)=0$ if $\sigma_j=0$ for some $j\in X$; such a potential is simply defined by

e6.4.12

$$\begin{aligned}&\Phi_X(\underline{\sigma}_X)=0\qquad \text{if } \sigma_j=0 \text{ for some } j\in X,\\&\Phi_X(\underline{\sigma}_X)=\sum_{Y\supset X}\Psi_Y(\underline{\sigma}_X\underline{0}),\end{aligned}\tag{6.4.12}$$

N6.4.4 *where $(\underline{\sigma}_X\underline{0})_j=0$ if $j\in Y/X$ and $(\underline{\sigma}_X\underline{0})_j=\sigma_j$ $(\ne 0)$ if $j\in X$.*[4]

---

[4] The notation we are using here is the same used in Section §5.1.

Remark: In this context the symbol 0 is called here "vacuum": however its only property is that transitions from $\sigma = 0$ to any σ' and vice versa are possible. Therefore if another symbol enjoys this property we could equally well take it as a reference symbol, call it "vacuum", and draw analogous conclusions.

Proof: It is immediate to check that if $\Psi \in B^h_{\log 2}$ then $\Phi \in B^h_0$, by using the definition in (6.4.12).

Taking into account proposition (5.1.6), (5.1.11) and (5.1.16) one finds for all $\mu \in M(\{0,\ldots,n\}^{\mathbb{Z}}_T)$

e6.4.13

$$\begin{aligned}\mu(A_\Psi) &= \lim_{N\to\infty} N^{-1}\mu\Big(\sum_{j=0}^{N-1} A_\Psi(\tau^j \underline{\sigma})\Big) = \lim_{N\to\infty} N^{-1}\mu\Big(\sum_{X\subset[0,\ldots,N]} \Psi_X(\underline{\sigma}_X)\Big) =\\ &= \lim_{N\to\infty} N^{-1}\mu\Big(\sum_{X\subset[0,\ldots,N]} \Phi_X(\underline{\sigma}_X) + N\sum_{X\ni 0}\frac{\Psi_X(\underline{0})}{|X|}\Big) =\\ &= \lim_{N\to\infty} N^{-1}\mu\Big(\sum_{j=0}^{N-1} A_\Phi(\tau^j\underline{\sigma}) + N\sum_{X\ni 0}\frac{\Psi_X(\underline{0})}{|X|}\Big) = \qquad (6.4.13)\\ &= \mu(A_\Phi) + \sum_{X\ni 0}\frac{\Psi_X(\underline{0})}{|X|},\end{aligned}$$

hence we are in the situation of item (i) of proposition (6.4.1). ■

It is then convenient to set the following definition.

D6.4.1 **(6.4.1) Definition:** (Gibbs states with transitive vacuum, particle potentials)
Let T be a mixing compatibility matrix with labels in $\{0,\ldots,n\}$ which is transitive on the vacuum, i.e. such that $T_{0\sigma} = T_{\sigma 0} = 1$ for all $\sigma = 0,\ldots,n$. Denote by $\widetilde{B} \subset B$ the space of the potentials for $\{0,\ldots,n\}^{\mathbb{Z}}_T$ such that

e6.4.14

$$\Phi_X(\underline{\sigma}_X) = 0 \qquad \textit{if} \quad \sigma_j = 0 \textit{ for some } j \in X. \qquad (6.4.14)$$

We shall say that $\{0,\ldots,n\}^{\mathbb{Z}}_T$ is the space of the configurations of n species of particles, denoted by the label $\sigma = 1,\ldots,n$, on $\mathbb{Z}$ subjected to a hard core *condition described by T (i.e. we interpret every 0 of the matrix T as a condition of incompatibility between two particles, that forbids the possibility of configurations in which they appear as nearest neighbours).*

The potentials $\Phi \in \widetilde{B}$ will be said particle potentials for n species of particles and a vacuum *and the Gibbs states on $\{0,\ldots,n\}^{\mathbb{Z}}_T$ associated with potentials $\Phi \in \widetilde{B}$ will be said* Gibbs states for n species of particles and a vacuum.

Remarks: The wording in the preceding definition is useful because it allows us to give a suggestive physical interpretation to the Gibbs states on $\{0,\ldots,n\}^{\mathbb{Z}}_T$ with T transitive on the vacuum. The configuration $\underline{\sigma} = \underline{0}$, with $\sigma_i \equiv 0$, is called, for obvious reasons, the *vacuum*.

The following proposition holds (Griffiths–Ruelle), [GR71].

P6.4.2 **(6.4.2) Proposition:** (Different particle potentials generate different Gibbs states)
If $\{0,\ldots,n\}_T^{\mathbb{Z}}$ *is defined by a mixing matrix* T *which is transitive on the vacuum and if* $\widetilde{B}$ *is the set of the* particles potentials *on* $\{0,\ldots,n\}_T^{\mathbb{Z}}$*, then* $\Phi,\ \Psi\in\widetilde{B}$ *and* $\Phi\neq\Psi$ *imply* $G(\Phi)\cap G(\Psi)=\emptyset$*.*

Remark: If we consider the graph of $\Phi\to P(\Phi)$ for $\Phi\in\widetilde{B}$ we see that the above proposition, combined with the results of propositions (6.1.1) and (6.1.2), means that the graph of $P(\Phi)$ on $\widetilde{B}$ is *strictly convex*, *i.e.* it is convex and in every point it has a different tangent plane (one checks that the propositions (6.1.1) and (6.1.2) and the remark following proposition (6.1.2) remain valid if, under the present hypotheses on T, we replace B with $\widetilde{B}\subset B$).

Proof: Let $\Lambda_0\subset\mathbb{Z}$ be a fixed finite region and let $\Lambda_N=[-N,N]\supset\Lambda_0$.
The DLR equations, see (5.1.9), imply that if $\mu\in G(\Phi)$ one has

e6.4.15
$$\mu(C^{\Lambda_N}_{\underline{\sigma}_{\Lambda_0}\underline{0}})=\int\frac{e^{-\sum_{X\subset\Lambda_0}\Phi_X(\underline{\sigma}_X)-W(\underline{\sigma}_{\Lambda_0}\underline{0}|\underline{\sigma}_{\Lambda_N^c})}}{Z_{\Lambda_N}(\Phi;\underline{\sigma}_{\Lambda_N^c})}\,\mu(d\underline{\sigma}_{\Lambda_N^c}),\tag{6.4.15}$$

where $\mu(d\underline{\sigma}_{\Lambda_N^c})$ denotes the restriction to $\mathcal{B}(\Lambda_N^c)$ of the measure μ, the configuration $(\underline{\sigma}_{\Lambda_0}\underline{0})\in\{0,\ldots,n\}_T^{\Lambda_N}$ is defined to be $\sigma'_j=0$ if $j\notin\Lambda_0$ and $\sigma'_j=\sigma_j$ if $j\in\Lambda_0$,

e6.4.16
$$W(\underline{\sigma}'_{\Lambda_N}|\underline{\sigma}_{\Lambda_N^c})=\sum_{\substack{R\cap\Lambda_N\neq\emptyset\\ R\cap\Lambda_N^c\neq\emptyset}}\Phi_R(\underline{\sigma}'_R),\tag{6.4.16}$$

with $\underline{\sigma}'=((\underline{\sigma}_{\Lambda_0}\underline{0})\underline{\sigma}_{\Lambda_N^c})$ and

e6.4.17
$$Z_{\Lambda_N}(\Phi;\underline{\sigma}_{\Lambda_N^c})=\sum_{\underline{\sigma}_{\Lambda_N}}e^{-\sum\limits_{X\subset\Lambda_N}\Phi_X(\underline{\sigma}_X)-W(\underline{\sigma}_{\Lambda_N}|\underline{\sigma}_{\Lambda_N^c})},\tag{6.4.17}$$

according to the notations of Section §6.1.
The condition $\|\Phi\|<+\infty$ shows that if Λ_0 is fixed

e6.4.18
$$\lim_{N\to\infty}W(\underline{\sigma}_{\Lambda_0}\underline{0}|\underline{\sigma}_{\Lambda_N^c})=0,\tag{6.4.18}$$

uniformly in $\underline{\sigma}_{\Lambda_0}$, $\underline{\sigma}_{\Lambda^c}$ and hence

e6.4.19
$$\frac{\mu(C^{\Lambda_N}_{\underline{\sigma}_{\Lambda_0}\underline{0}})}{\mu(C^{\Lambda_N}_{\underline{0}})}\xrightarrow[N\to\infty]{}e^{-\sum_{X\subset\Lambda_0}\Phi_X(\underline{\sigma}_X)}.\tag{6.4.19}$$

This means that $\mu \in G(\Phi)$ determines uniquely $\overline{U}_{\Lambda_0}(\underline{\sigma}_{\Lambda_0}) = \sum_{X\subset\Lambda_0} \Phi_X(\underline{\sigma}_X)$
N6.4.5 for all $\sigma_{\Lambda_0} \in \{0,\ldots,n\}^{\Lambda_0}_T$.[5] In turn $\overline{U}_{\Lambda_0}$ determines Φ recursively as

e6.4.20
$$\Phi_X(\underline{\sigma}_X) = \sum_{R\subset X} (-1)^{|R|}\,\overline{U}_R(\underline{\sigma}_R), \tag{6.4.20}$$

therefore $\mu \in G(\Phi)$ determines Φ uniquely if $\Phi \in \widetilde{B}$. ■

We conclude this section by discussing an interesting property related to the exponential mixing properties of the SRB distributions. If (Ω, S) is an Anosov system with a SRB distribution μ and f is a Hölder continuous function the limit as $N \to \infty$ of the dispersion of the variable $N^{-\frac{1}{2}}\sum_{k=0}^{N-1} f(S^k x)$ will be finite if f has zero SRB average. It will be given by the exponentially rapidly converging series $D = \sum_{k=-\infty}^{\infty} \langle f(S^k\cdot)f(\cdot)\rangle_\mu$ where $\langle\cdot\rangle_\mu$ denotes average with respect to μ. Clearly $D \ge 0$ and a natural question is whether this can vanish. The following proposition gives a necessary and sufficient condition (Livsic–Sinai).

P6.4.3 **(6.4.3) Proposition:** (A vanishing dispersion condition)
Let (Ω, S) be an Anosov system and let f be a Hölder continuous function with vanishing SRB average, $\int f d\mu = 0$, and with vanishing dispersion, $D \overset{def}{=} \sum_{k=-\infty}^{\infty} \langle f(S^k\cdot)f(\cdot)\rangle_\mu = 0$. Let $A(x)$ be the energy function for the SRB distribution μ, i.e. the energy function A, see definition (5.1.1), associated with the potential Φ for the SRB distribution. Then
(i) A and $A+f$ are equivalent energy functions, i.e. the same Gibbs distribution μ maximizes $s(\mu') - \mu'(A)$ and $s(\mu') - \mu'(A+f)$ over $\mu' \in \mathcal{M}^t_e(\Omega, S)$, cf. *corollary (6.2.1).*
(ii) $f(x) = F(x) - F(Sx)$ for a suitable F which is Hölder continuous on Ω.
Property (ii) implies that a necessary and sufficient condition in order to have $D > 0$ is that the average of f along a periodic orbit does not vanish.

See Appendix (6.4).

Appendix 6.4: *Vanishing dispersion conditions*

The following lemma shows that $D = 0$ has interesting implications and prepares the proof of proposition (6.4.3)

L6.4.1 **(6.4.1) Lemma:** (Vanishing dispersion: a L_2 condition)
Let (Ω, S) be an Anosov system. A necessary and sufficient condition in order that a Hölder continuous function f with zero SRB average ($\int d\mu f = 0$) has a zero dispersion $D \overset{def}{=} \sum_{k=-\infty}^{\infty} \int f(S^k x) f(x)\mu(dx)$ in the SRB distribution μ for (Ω, S) is that there is a function $F(x)$ defined on a set $V_0 \subset \Omega$ such that

eA6.4.1
$$f(x) = F(x) - F(Sx), \qquad \textit{for } x \in V_0 \textit{ and } \mu(V_0) = 1 \tag{A6.4.1}$$

and there exist constants $C, a > 0$ and $V \subset \Omega\times\Omega, \mu\times\mu(V) = 1$ such that

---

[5] With $\{0,\ldots,n\}^{\Lambda_0}_T$ we mean (naturally) the set of the elements of $\{0,\ldots,n\}^{\Lambda_0}$ that can be extended in at least one way to an element of $\{0,\ldots,n\}^{\mathbb{Z}}_T$.

eA6.4.2
$$|F(x)-F(y)|<C|x-y|^\alpha \qquad \text{for all } (x,y)\in V \tag{A6.4.2}$$

Hence F is bounded μ–almost everywhere.

Remark: From (A6.4.2) we *cannot* (yet) conclude that F is Hölder continuous because $V\neq\Omega\times\Omega$.

Proof: We discuss the case of 2–dimensional Anosov maps, for simplicity. Consider the sum $F_N(x)\overset{def}{=}\sum_{k=0}^{N-1}f(S^kx)$. If $D=0$ one finds that $\int d\mu\,(\sum_{k=0}^{N-1}f(S^kx))^2$ is bounded uniformly in N. Therefore the sequence $\sum_{k=0}^{N-1}f(S^kx)$ is bounded in $L_2(\mu)$. Furthermore for all g smooth $\lim_{N\to\infty}\int\sum_{j=0}^{N-1}f(S^jx)g(x)\mu(dx)$ exists (by mixing) so that the limit $F(x)=\lim_{N\to\infty}\sum_{k=0}^{\infty}f(S^kx)$ exists weakly in $L_2(\mu)$ and (therefore) $f(x)=F(x)-F(Sx)$ μ–almost everywhere. The difficulty is in the proof of the weak kind of Hölder continuity of F claimed in (A6.4.2).

The same arguments as above can be made for time tending to $-\infty$: thus we can define a function $F_-(x)=\sum_{k=-1}^{-\infty}f(S^kx)$. The function $F_-(x)+F(x)$ is however invariant: and therefore it must be constant. The constant must be 0 because f has zero average.

Let $\xi_0\in\Omega$ and let $R=W^u_\gamma(\xi_0)\times W^s_\gamma(\xi_0)$ be an S–rectangle, in the sense of definition (4.2.2), with center ξ_0 and pair of axes $C=W^u_\gamma(\xi_0)$ and $D=W^s_\gamma(\xi_0)$ with the notation of Section §4.2, cf. Fig.(4.2.1). The generic point in R has the form $[\varphi,\varphi']$ with $\varphi\in C,\varphi'\in D$, where $[\varphi,\varphi']=W^s_\gamma(\varphi)\cap W^u_\gamma(\varphi')$. We can represent, by corollary (6.3.3), the integration of a function G on R as

eA6.4.3
$$\int_R G(x)\mu(dx)=\int_C d\sigma_\varphi\int_D\nu(d\varphi')\rho([\varphi,\varphi'])G([\varphi,\varphi']) \tag{A6.4.3}$$

where $d\sigma_\varphi$ is the arc length on C and ν is a suitably defined finite measure on the arc D and $\rho(\varphi,\varphi')=\prod_{k=0}^{\infty}\frac{\lambda_u^{-1}(S^k[\varphi,\varphi'])}{\lambda_u^{-1}(S^k\varphi)}$; note that $\rho(\varphi,\varphi')$ is Hölder continuous in $(\varphi,\varphi')\in C\times D$ and bounded away from $0,\infty$.

The key remark is that the function on $V\times V$ defined by $F_N([x,y])$ is weakly convergent to the limit $F([x,y])$ in $L_2(\mu)$. Indeed let $\Gamma(x,y)$ be a bounded test function. Then [6] N6.4.6

eA6.4.4
$$\begin{aligned}\int_R\Gamma(x,y)F_N([x,y])\mu(dx)\mu(dy)=&\int_{C\times C}d\sigma_{\varphi_0}d\sigma_{\varphi_1}\int_{D\times D}\nu(d\varphi'_0)\nu(d\varphi'_1)\cdot\\ &\cdot\rho(\varphi_0,\varphi'_0)\rho(\varphi_1,\varphi'_1)\Gamma([\varphi_0,\varphi'_0],[\varphi_1,\varphi'_1])F_N([\varphi_0,\varphi'_1]),\end{aligned} \tag{A6.4.4}$$

which, multiplying and dividing by $\rho(\varphi_0,\varphi'_1)$ can be written $\int_R F_N(z)\widetilde\Gamma(z)\,\mu(dz)$ where, if $z=[\varphi_0,\varphi'_1]$,

eA6.4.5
$$\widetilde\Gamma(z)=\int_{C\times D}d\sigma_{\varphi_1}\nu(d\varphi'_0)\Gamma([\varphi_0,\varphi'_0],[\varphi_1,\varphi'_1])\frac{\rho(\varphi_0,\varphi'_0)\rho(\varphi_1,\varphi'_1)}{\rho(\varphi_0,\varphi'_1)} \tag{A6.4.5}$$

is a bounded function.

Therefore $\int\Gamma(x,y)F_N([x,y])\mu(dx)\mu(dy)\xrightarrow[N\to\infty]{}\int_R\mu(dz)F(z)\widetilde\Gamma(z)$ because F_N converges weakly to F in $L_2(\mu)$. By going backwards this means that $F_N([x,y])$ is weakly convergent to $F([x,y])$ in $L_2(\mu\times\mu)$.

The function $\sum_{k=0}^{N-1}(f(S^kx)-f(S^k[x,y]))$ is bounded by

eA6.4.6
$$\Big|\sum_{k=0}^{N-1}(f(S^kx)-f(S^k[y,x]))\Big|\le C|x-y|^\alpha, \tag{A6.4.6}$$

---

[6] We use that $[x,y]=[\varphi_0,\varphi'_1]$ if $x=[\varphi_0,\varphi'_0]$ and $y=[\varphi_1,\varphi'_1]$.

if C is a suitable constant and α is the Holder continuity exponent of f times the logarithm of the constant λ bounding the minimum expansion rate. This is so because the points x and $[x,y]$ are on the same stable manifold and therefore $S^k x$ and $S^k[x,y]$ get closer and closer exponentially fast with their distance bounded proportionally to λ^{-k}. Hence for a suitable $C>0$

eA6.4.7
$$|F(x)-F[x,y]| \le C|x-y|^\alpha \tag{A6.4.7}$$

holds almost everywhere with respect to $\mu\times\mu$. Likewise $|F_-(y)-F_-([x,y])| \le C|x-y|^\alpha$ almost everywhere with respect to $\mu\times\mu$. Since $F_-(x)=-F(x)$ holds μ–almost everywhere we get

eA6.4.8
$$|F(x)-F(y)| \le 2^{1-\alpha}C|x-y|^\alpha \qquad \text{for all } (x,y)\in V \tag{A6.4.8}$$

where V is such that $\mu\times\mu(V)=\mu\times\mu(R)=1$.

Hence for μ–almost all x one has $|F(x)-F(y)| \le 2^{1-\alpha}\,C\,\mathrm{diam}(R)^\alpha$ for μ-almost all y: and for μ-almost all y one has $|F(y)| \le 2^{1-\alpha}\,C\,\mathrm{diam}(R)^\alpha$ so that the function F is bounded uniformly μ-almost everywhere. ∎

If b is a μ-almost everywhere bound on $|F(x)|$ the above lemma implies that

eA6.4.9
$$e^{-2b}\mu(dx) \le \mu_N(dx) = \mu(dx)e^{\sum_{j=-N}^{N} f(S^j x)} \le e^{2b}\mu(dx) \tag{A6.4.9}$$

because $\sum_{j=-N}^{N} f(S^j x) = F(S^N x) - F(S^{-N}x)$. On the other hand it follows also that the limit as $N\to\infty$ of $c_N\mu_N$ with c_N being a normalization factor ($e^{-2b}\le c_N \le e^{2b}$) is the Gibbs distribution with energy function $A(x)+f(x)$ if $A(x)$ is the potential function for μ, cf. (6.2.11). Indeed by proposition (5.2.1) we know that the Gibbs distributions associated with A and with $A+f$ are ergodic: they coincide with μ and with the limit as $N\to\infty$ of μ_N, respectively, hence (A6.4.9) implies that the two distributions are absolutely continuous with respect to each other hence they coincide. Therefore the following result holds.

C6.4.2 **(6.4.2) Corollary:** *Let (Ω,S) be an Anosov system and let f be a Hölder continuous function with vanishing SRB average, $\int f d\mu = 0$, and vanishing dispersion $D=\sum_{j=-\infty}^{\infty}\mu\big(f(S^j\cdot)f(\cdot)\big)=0$. If $A(x)$ is the potential function for μ then $A(x)+f(x)$ is also a potential function for μ.*

The problem posed before Lemma (6.4.1) above is solved by the following proposition

Proof of proposition (6.4.3): The argument preceding corollary (6.4.2) shows that $|\sum_{j=-N}^{N} f(S^j x)| < 2b$ everywhere (because f is continuous). Therefore by the same argument one sees that given any Gibbs distribution μ_0 with potential function A_0 the potentials A_0 and A_0+f are equivalent. Therefore we can apply (i) of proposition (6.4.1) to conclude that there exists F continuous and a constant C such that $f(x)=F(x)-F(Sx)+C$; in our case $C=0$ because $\int f d\mu = 0$ and (i) and (ii) are checked. Item (iii) follows from (ii). ∎

Problems for §6.4

Q6.4.1 **[6.4.1]:** Consider the potential on $\{0,1\}^{\mathbb{Z}}$

$$\Phi_X(\underline{\sigma}_X)=0 \qquad |X|\ge 2\ , \qquad \Phi_{\{0\}}(\underline{\sigma})=h\sigma.$$

Find an element Φ' of B equivalent to it (*i.e.* such that $G(\Phi)=G(\Phi')$) but with potential with non vanishing many-body components and which: (a) decay exponentially ($\Phi'\in B_0^h$) (see proposition (6.4.1) and definition (6.4.1)), (b) is such that $\|\Phi'\|_1=+\infty$, (c) does not contain potentials with more than three bodies (*here a, b, c are meant as mutually excluding properties*).

Q6.4.2 **[6.4.2]:** Does it exist $\Phi\in B$ that is not equivalent to some $\Phi'\in\widetilde{B}$? (see proposition (6.4.1)).

Q6.4.3 **[6.4.3]:** Check that in the theory of Gibbs measures the function $A_\Phi = \sum_{X\ni 0} \frac{\Phi_X(\underline{\sigma}_X)}{|X|}$ can be replaced without substantial modifications by $A'_\Phi(\underline{\sigma}) = \sum_{X\ni 0,\, X\geq 0} \Phi_X(\underline{\sigma}_X)$, ($X \geq 0$ means $X \subset \mathbb{Z}^+$).

Q6.4.4 **[6.4.4]:** Making use of the results in problem [6.4.3] show that the potential $\Phi_X = 0$ unless $|X| = 3$ and $\Phi_{\{0,j,j+1\}}(\sigma_0, \sigma_j, \sigma_{j+1}) = (\sigma_j - \sigma_{j+1})\,\Phi_j$ and zero otherwise, is equivalent to zero if $\sum_j |\Phi_j| < +\infty$.

Q6.4.5 **[6.4.5]:** Consider the group of the $(n+1)^{\rm th}$ roots of unity: $r_0,\ r_1, \ldots, r_n$. Show that every potential for $\{0,\ldots,n\}^{\mathbb{Z}}$, $\Phi \in B^h_s$, with s large enough, is equivalent to a potential $\widetilde{\Phi}$ (*spin potential*) such that $\widetilde{\Phi}_X(\underline{\sigma}_X) = \Phi_X \cdot\ {\rm Re}\,\Big(\prod_{j\in X} r_{\sigma_j}\Big)$ with $\Phi_X \in \mathbb{R}$ and $\sum_{X\ni 0} e^{\kappa {\rm diam}(X)} |\Phi_X| < +\infty$, with $\kappa > 0$ suitable.

Bibliographical note to §6.4

The contents of this section are based on [GR71], and on the interesting proposition (6.4.3) (cf. p. 78, theorem 5.7 and the bibliographical note at p. 96 in [Ru78]). This theorem can be extended to potentials more general than those verifying (6.4.1) (it suffices to push the precision of the estimates, giving up proving properties as strong as (6.4.6) or (6.4.9) but looking for the minimal conditions sufficient to obtain continuity of F). The vanishing dispersion analysis is taken, at Ruelle's suggestion, from [Ru78].