ON THE ZERO-TEMPERATURE OR VANISHING VISCOSITY LIMIT FOR CERTAIN MARKOV PROCESSES ARISING FROM LAGRANGIAN DYNAMICS

NALINI ANANTHARAMAN

Abstract. We study the zero-temperature limit for Gibbs measures associated to Frenkel-Kontorova models on $(\mathbb{R}^d)^{\mathbb{Z}}/\mathbb{Z}^d$. We prove that equilibrium states concentrate on configurations of minimal energy, and, in addition, must satisfy a variational principle involving metric entropy and Lyapunov exponents, a bit like in the Ruelle-Pesin inequality. Then we transpose the result to certain continuous-time stationary stochastic processes associated to the viscous Hamilton-Jacobi equation. As the viscosity vanishes, the invariant measure of the process concentrates on the so-called "Mather set" of classical mechanics, and must, in addition, minimize the gap in the Ruelle-Pesin inequality.

In statistical mechanics, Gibbs measures are probability measures on the configuration space, describing states of thermodynamical equilibrium. One of the major problems is to study the dependence of equilibrium states on the temperature (or other parameters): a lack of analyticity in this dependence is interpreted as the occurrence of a phase transition, and the existence of several Gibbs measures at a given temperature, as the coexistence of several phases.

In Part I of this paper, we are interested in the behaviour of Gibbs measures as temperature goes to zero, in the model where the particles of the system lie on the 1-dimensional lattice $\mathbb{Z}$. This is not the favourite situation in statistical mechanics: in this case, and if the energy of interaction between particles satisfies reasonable assumptions, there is usually no phase transition. But even then, there is, to my knowledge, no general result describing completely the behaviour of Gibbs measures at zero temperature: for instance, the existence or not of a limit of the equilibrium state. It is intuitive to think, and possible to prove, that such a limit must minimize the mean energy, but there are examples where it is not enough to conclude, as there may be several states of minimal mean energy ([Si82]).

This paper deals with the case where the state of each particle is represented by an element of $\mathbb{R}^d$, so that a configuration of the system is described by a sequence $\gamma = (\gamma_k)_{k\in\mathbb{Z}} \in (\mathbb{R}^d)^{\mathbb{Z}}$. We work in the Markovian case: the potential of interaction is of the form $\bar{L}(\gamma) = L(\gamma_0, \gamma_1)$. Such models are sometimes called Frenkel-Kontorova models. In the paper, the function $L : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ will be of class C^3, and satisfy the following assumptions :

(Periodicity) $L(x + s, y + s) = L(x, y)$, for all $s \in \mathbb{Z}^d$.

(Superlinear growth) $\frac{L(x,y)}{||x-y||} \underset{||x-y||\to\infty}{\longrightarrow} +\infty$

('Twist property') For all $x \in \mathbb{R}^d$, $y \mapsto \partial_1 L(x, y)$ is a diffeomorphism of $\mathbb{R}^d$.

Received by the editors April 2003.

Given the periodicity property of L, the convenient configuration space to work with is the quotient space $W = (\mathbb{R}^d)^{\mathbb{Z}}/\mathbb{Z}^d$. We will denote σ the shift transformation on W, acting on sequences by shifting them to the left.

For each parameter $\beta > 0$ (representing the inverse of the temperature), we will construct in Section 2 a σ-invariant probability measure μ_β on W, called "the Gibbs measure associated to the potential $\bar{L}$, at temperature $1/\beta$".

We will then prove the following theorem:

Theorem 0.0.1. *Let μ_β be the Gibbs measure associated to the potential $\bar{L}$, at temperature $1/\beta$. Let μ_∞ be a limit point of the family $(\mu_\beta)_{\beta>0}$ as β tends to infinity. Then μ_∞ minimizes the mean energy $\int_W \bar{L} d\mu$ over the set of all σ-invariant probability measures on W.*

Moreover, under suitable assumptions (A1), (A2), (A3), μ_∞ maximizes the functional

$$\mu \mapsto h_\sigma(\mu) - \frac{1}{2} \lim_{n\to\infty} \int_W \frac{1}{n} \log[{}_nA''(\bar{\gamma})] d\mu(\bar{\gamma})$$

over all energy-minimizing measures.

In the theorem, $h_\sigma(\mu)$ stands for the metric entropy of the measure μ under the action of the shift σ. This functional is one of the fundamental objects in ergodic theory; its definition is given in Section 1.

We denote $A''(\gamma)$ the hessian matrix of the formal sum

$$A(\gamma) = \sum_{k\in\mathbb{Z}} L(\gamma_k, \gamma_{k+1}).$$

It is an infinite matrix, tridiagonal by blocks of size d. The notation ${}_nA''(\gamma)$ stands for the $nd \times nd$ submatrix corresponding to $k \in [1, n]$ and, in Theorem 0.0.1, $[{}_nA''(\gamma)]$ stands for the determinant of that matrix.

The assumptions (A1), (A2), (A3) are technical assumptions concerning the behaviour of the energy near its minima. They will be stated in Section 1.

We note that our result does not necessarily imply the existence of a limit for the family $(\mu_\beta)_{\beta\to+\infty}$, since the functional that μ_∞ must maximize is affine.

Part I is organized as follows:

– in Section 1, we introduce notations, and give a more precise statement of Theorem 0.0.1 with its assumptions (A1), (A2), (A3).

– in Section 2, we define the Gibbs measure μ_β and give some of its properties.

– in Section 3, we prove Theorem 0.0.1.

– Section 4 serves as a transition with Part II. We explain briefly the connection between Frenkel-Kontorova models and symplectic twist diffeomorphisms of $\mathbb{T}^d \times \mathbb{R}^d$. In this context, the quantity $\lim_n \frac{1}{n} \log[{}_nA''(\bar{\gamma})]$ has a nice interpretation in terms of Lyapunov exponents. To draw an explicit link with Part II, we consider the example

$$L(\gamma_0, \gamma_1) = \frac{||\gamma_1 - \gamma_0||^2}{2} - V(\gamma_0) - \langle \omega, \gamma_1 - \gamma_0 \rangle,$$

where $\omega \in \mathbb{R}^d$ and V is a Z^d-periodic potential of class C^3.

In Part II, we find that our result reads in an interesting way when transposed to the field of lagrangian mechanics. In that part, we replace the configuration space $W = (\mathbb{R}^d)^{\mathbb{Z}}/\mathbb{Z}^d$ by the space of continuous bi-infinite paths on the d-torus,

$W = C(\mathbb{R}, \mathbb{R}^d)/\mathbb{Z}^d = C(\mathbb{R}, \mathbb{T}^d)$, and the function L by a lagrangian of the form

$$\mathcal{L}_\omega(x, v) = \frac{\| v \|^2}{2} - V(x) - \langle \omega, v \rangle \tag{0.0.1}$$

on $\mathbb{R}^d \times \mathbb{R}^d$, V being a $\mathbb{Z}^d$-periodic function of class C^3, ω an element of $\mathbb{R}^d$, and $\| \,.\, \|$ the norm arising from the usual euclidean structure $\langle ., . \rangle$ on $\mathbb{R}^d$.

For $\beta > 0$, we consider the "twisted" Schrödinger operators:

$$\mathcal{H}_\beta^\omega = e^{-\beta\langle \omega, x\rangle} \circ \left(\frac{\Delta}{2\beta^2} + V(x) \right) \circ e^{\beta\langle \omega, x\rangle}, \tag{0.0.2}$$

where Δ stands for the Laplace operator on $\mathbb{R}^d$. (For $\beta = \frac{i}{\hbar}$, $\hbar$ being the Planck constant, the operator $\mathcal{H}_\beta^\omega$ would be the quantization of the classical Hamiltonian

$$H_\omega(x, p) = \frac{||p + \omega||^2}{2} + V(x)$$

associated to the Lagrangian $\mathcal{L}_\omega$; but this is quite a different problem.)

Let ψ_β, ψ_β^* be the positive $\mathbb{Z}^d$-periodic eigenfunctions of, respectively, $\mathcal{H}_\beta^\omega$ and its adjoint $\mathcal{H}_\beta^{\omega *}$, associated to their common largest eigenvalue (this statement will be given a precise meaning in Section 5). The focus of our attention is the measure

$$\mu_\beta^0 = \frac{\psi_\beta(x)\psi_\beta^*(x)dx}{\int_{\mathbb{T}^d} \psi_\beta(y)\psi_\beta^*(y)dy},$$

which is the invariant measure for the Markov process generated by the twisted Schrödinger operators (Section 5). This process seems to be called $P(\phi)_1$-process in quantum field theory ([Si79]).

We study the behaviour of the family (μ_β^0) as $\beta \longrightarrow +\infty$; $1/\beta$ will now play the role of a viscosity coefficient, or of the diffusion coefficient of the stochastic process. We first prove that every limit point μ_∞^0 of the family $(\mu_\beta^0)_{\beta \to +\infty}$ can be lifted to the tangent bundle $\mathbb{T}^d \times \mathbb{R}^d$ to a probability measure μ_∞, invariant under the Euler-Lagrange flow of $\mathcal{L}_\omega$, and which minimizes the integral of the lagrangian. Such measures play a central role in J. Mather's theory in lagrangian dynamical systems: they are called "action-minimizing measures" (see [Ma91], and the work of Mañe on the subject, [Mn92], [Mn96] and [Mn97]). It is shown in the paper [Go02] (Section 8) how the measures μ_β^0, for $\beta > 0$, may be seen as action-minimizing measures in the world of stochastic dynamics.

Since there may be several action-minimizing measures, we seek additional conditions satisfied by the limits of (μ_β^0) as $\beta \longrightarrow +\infty$.

One way to state the result is as follows :

Theorem 0.0.2. *Let μ_∞^0 be a limit point of the family (μ_β^0) as $\beta \to +\infty$.*

Then there exists a probability measure μ_∞ on $\mathbb{T}^d \times \mathbb{R}^d$, which is invariant under the Euler-Lagrange flow, action-minimizing in the sense of J. Mather, and whose projection on $\mathbb{T}^d$ is μ_∞^0.

Moreover, under suitable assumptions (A1), (A2), (A3), μ_∞ maximizes the functional

$$h_\phi(\mu) - \frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} (\sum_{i=1}^d \lambda_i^+(x, v)) d\mu(x, v)$$

over the set of all action-minimizing measures.

The assumptions (A1), (A2), (A3) are technical assumptions concerning the behaviour of the action near its global minima. They will be stated in Section 5.

Here $h_\phi(\mu)$ stands for the metric entropy of the invariant probability measure μ on $\mathbb{T}^d \times \mathbb{R}^d$, under the action of the Euler-Lagrange flow $\phi = (\phi_t)_{t\in\mathbb{R}}$; and the $\lambda_i^+(x, v)$ are the d first (nonnegative) Lyapunov exponents of (x, v), under the action of ϕ. The definition of Lyapunov exponents will be given in Section 4. Note that, for a smooth transformation ϕ of a compact manifold of dimension n, the Ruelle inequality always holds:

$$h_\phi(\mu) \leq \frac{1}{2} \int \sum_{i=1}^{n} |\lambda_i(x)| d\mu(x)$$

where the sum runs over *all* Lyapunov exponents (this is Theorem S.2.13 of [KH95], applied to both ϕ and ϕ^{-1}). In Theorem 0.0.2, if we knew that μ was supported on a smooth invariant Lagrangian graph (hence, of dimension d), we could interprete the result as: "μ minimizes the gap in Ruelle's inequality". As we shall explain in Section 5, the fact that μ is action-minimizing in the sense of Mather is a weak form of the property of being carried by a smooth invariant Lagrangian graph.

There are alternative ways of formulating the result. For instance, $u_\beta = -\frac{\log \psi_\beta^*}{\beta}$ is solution of the viscous Hamilton-Jacobi equation:

$$-\frac{\Delta u}{2\beta} + H_\omega(x, d_x u) = \lambda_\beta,$$

whereas $v_\beta = -\frac{\log \psi_\beta^*}{\beta}$ is solution of the same equation for the time-reversed system:

$$-\frac{\Delta v}{2\beta} + H_\omega(x, -d_x v) = \lambda_\beta.$$

The constant λ_β is the eigenvalue of $\mathcal{H}_\beta^\omega$ and $\mathcal{H}_\beta^{\omega *}$ associated to the eigenfunctions ψ_β, ψ_β^*. We see that β appears here in the role of the inverse of a viscosity coefficient. The measure $\psi_\beta(x)\psi_\beta^*(x)dx$ may thus be written in the form $e^{-\beta(u_\beta(x)+v_\beta(x))}dx$. In dimension $d = 1$, the problem may also be formulated in terms of the Burgers equation, which is the equation satisfied by du_β (or dv_β), obtained by differentiating the Hamilton-Jacobi equation; in the paper [Si91], the asymptotic behaviour of the viscous Burgers equation (as time tends to infinity, for a fixed viscosity coefficient) was studied via the definition of Gibbs measures on path spaces; our construction, in Section 5, of the Markov process associated to the Schrödinger equation, is similar.

Let us mention that in dimension $d = 1$, the convergence of the functions $u_\beta, v_\beta (\beta \to +\infty)$ was proved in [Bes02] for a time-dependent Lagrangian, and that the result proved therein implies ours. However, the approach relies very much on low-dimensional considerations and cannot be extended to higher dimension in an obvious way. Besides, in low dimension, the entropy does not come into play.

The application of Theorem 0.0.2 in the case $\omega = 0$ yields an already known result about the tunnelling effect in semi-classical mechanics ([He88], Section 4.4):

Corollary 0.0.3. *Let $\mathcal{H}_\hbar = \hbar^2 \frac{\Delta}{2} + V$, and let $\psi_\hbar$ be the unique $\mathbb{Z}^d$-periodic positive eigenfunction, corresponding to the largest eigenvalue of $\mathcal{H}_\hbar$ in $L^2(\mathbb{T}^d)$.*

Then, in the semi-classical limit $\hbar \longrightarrow 0$, the probability measure

$$\frac{\psi_\hbar^2(x)dx}{\int_{\mathbb{T}^d} \psi_\hbar^2(y)dy}$$

concentrates on the maxima of V.

Assume furthermore that the system satisfies (A1). If we consider the expansion of V in orthonormal coordinates near a maximum x_0, in the form

$$V(x_0 + y) = V(x_0) - \frac{1}{2} \sum |a_i(x_0)|^2 y_i^2 + O(y^3),$$

then the measure $\psi_\hbar^2(x)dx$ concentrates on those x_0s for which the quantity

$$\sum |a_i(x_0)|$$

is minimal.

Part II is organized as follows:

– in Section 5, we give more details about Mather theory and the notion of viscosity solutions of Hamilton-Jacobi equations. We explain the spectral properties of the twisted Schrödinger operator (0.0.2). We show how it generates a Markov process of invariant distribution μ_β^0, and, finally we state Assumptions (A1), (A2), (A3).

– in Section 6, we show how to adapt the proof of Theorem 0.0.1 to the new situation. We also check that Assumptions (A2), (A3) are always satisfied in the case $\omega = 0$, that is, we prove Corollary 0.0.3.

Part 1. **Statistical mechanics**

1. Introduction and statement of results

In this part, we consider a model where particles lie on the "1-dimensional lattice" $\mathbb{Z}$, and the state of each particle is described by an element of $\mathbb{R}^d$. Thus,a configuration of the whole system is described by an element of $(\mathbb{R}^d)^{\mathbb{Z}}$. A function

$$\bar{L} : (\mathbb{R}^d)^{\mathbb{Z}} \longrightarrow \mathbb{R},$$

called the potential of interaction, is used to describe the energy of interaction between particles. This is done the following way: given a configuration $\gamma = (\gamma_k)_{k\in\mathbb{Z}} \in (\mathbb{R}^d)^{\mathbb{Z}}$, the energy of interaction associated to a finite subsequence $(\gamma_k)_{m\le k\le n}$ is by definition

$$A(\gamma_{|[m,n]}) = \sum_{k=m}^{n-1} \bar{L}(\sigma^k \gamma)$$

where σ denotes the shift acting to the left:

$$(\sigma\gamma)_k = \gamma_{k+1}$$

We will restrict our attention to potentials $\bar{L}$ depending only on the two first coordinates (nearest neighbour interactions): in other words $\bar{L}(\gamma) = L(\gamma_0, \gamma_1)$, where now L is a function from $\mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}$.

Moreover, L will be of class C^3, and have the following properties:

(Periodicity) $L(x + s, y + s) = L(x, y)$, for all $s \in \mathbb{Z}^d$.

(Superlinear growth) $\frac{L(x,y)}{||x-y||} \underset{||x-y||\longrightarrow\infty}{\longrightarrow} +\infty$

('Twist property') For all $x \in \mathbb{R}^d$, $y \mapsto \partial_1 L(x, y)$ is a diffeomorphism of $\mathbb{R}^d$.

A model which assigns an energy

$$A(\gamma_{|[m,n]}) = \sum_{k=m}^{n-1} L(\gamma_k, \gamma_{k+1})$$

to any finite segment $m \leq k \leq n$ of a configuration $\gamma = (\gamma_k)$, is usually called a Frenkel-Kontorova model. Its "stationary configurations" are, by definition, the configurations γ which, for all $m < n$, are critical points of A with respect to variations of γ_k, $m < k < n$. In other words,

$$\partial_2 L(\gamma_{k-1}, \gamma_k) + \partial_1 L(\gamma_k, \gamma_{k+1}) = 0$$

for all k.

Given the periodicity property of L, the convenient configuration space to work with is the quotient space $W = (\mathbb{R}^d)^{\mathbb{Z}}/\mathbb{Z}^d$, the action of $\mathbb{Z}^d$ on $(\mathbb{R}^d)^{\mathbb{Z}}$ being defined by

$$(s.\gamma)_k = \gamma_k + s$$

for all $s \in \mathbb{Z}^d$, for all $\gamma \in (\mathbb{R}^d)^{\mathbb{Z}}$, for all $k \in \mathbb{Z}$.

Notations : We shall denote $\bar{\gamma} \in W$ the equivalence class of $\gamma \in (\mathbb{R}^d)^{\mathbb{Z}}$ under this action. An element of W will always be denoted in the form $\bar{\gamma}$, meaning thereby that it is the equivalence class of some $\gamma \in (\mathbb{R}^d)^{\mathbb{Z}}$.

Similarly, for any subset $I \subset \mathbb{Z}$, we shall introduce the quotient space $W_I = (\mathbb{R}^d)^I/\mathbb{Z}^d$, with the action of $\mathbb{Z}^d$ defined as above, and we shall denote $\bar{\gamma} \in W_I$ the equivalence class of $\gamma \in (\mathbb{R}^d)^I$.

The shift σ, defined previously on $(\mathbb{R}^d)^{\mathbb{Z}}$, can be defined on the quotient space W; the same holds for the potential $\bar{L}$. We keep the same notation for the shift σ and the potential $\bar{L}$ defined on W. More generally, when some functions or transformations originally defined on $(\mathbb{R}^d)^{\mathbb{Z}}$ can go to the quotient space W, we keep the same notation.

We also introduce the projections $\pi_I : (\mathbb{R}^d)^{\mathbb{Z}} \longrightarrow (\mathbb{R}^d)^I$, which go to the quotient spaces:

$$\pi_I : W \longrightarrow W_I$$

When $I = [0, k]$, we shall write W_k, π_k instead of W_I, π_I. In particular, $W_0 \simeq \mathbb{T}^d$, $W_1 \simeq \mathbb{T}^d \times \mathbb{R}^d$.

The topology used on $(\mathbb{R}^d)^{\mathbb{Z}}$ is the product topology, and the topology on W is the quotient topology. It is defined by the distance

$$d_W(\bar{\gamma}, \bar{\xi}) = d_{\mathbb{T}^d}(\overline{\gamma_0}, \overline{\xi_0}) + \sum_{k \in \mathbb{Z}} \frac{1}{2^{|k|}} \min\left(|\,||\gamma_{k+1} - \gamma_k|| - ||\xi_{k+1} - \xi_k||\,|, 1\right).$$

If I is an interval of $\mathbb{Z}$ containing 0, we define similarly a distance d_{W_I} on W_I; the $\sum$ now runs over all ks such that $k \in I, k+1 \in I$.

We can now introduce our Gibbs measures μ_β. As we shall prove in the next section, for all $\beta > 0$, we can find $\mathbb{Z}^d$-periodic, positive continuous functions ψ_β, ψ^*_β, and a real number λ_β, such that

$$\int_{\mathbb{R}^d} e^{-\beta L(x,y)} \psi_\beta(y) dy = e^{\lambda_\beta} \psi_\beta(x) \tag{1.0.3}$$

and

$$\int_{\mathbb{R}^d} e^{-\beta L(y,x)} \psi^*_\beta(y) dy = e^{\lambda_\beta} \psi^*_\beta(x), \tag{1.0.4}$$

for all x. Actually, the real number λ_β and (up to a multiplicative factor) the functions ψ_β, ψ^*_β are characterized by these properties.

We normalize the functions ψ_β, ψ^*_β so that $\int_{[0,1)^d} \psi_\beta(x)\psi^*_\beta(x)dx = 1$.

The measure $\psi_\beta(x)\psi^*_\beta(x)dx$ appears as the unique $\mathbb{Z}^d$-periodic invariant measure for the Markov process with transition probabilities

$$P(x, dy) = \frac{\psi_\beta(y)}{e^{\lambda_\beta}\psi_\beta(x)} e^{-\beta L(x,y)} dy.$$

The stationary Markovian process on $(\mathbb{R}^d)^{\mathbb{N}}$, of initial distribution $\psi_\beta(x)\psi^*_\beta(x)dx$ and transition probabilities $P(x, dy)$, is realized by the following measure μ_β on $(\mathbb{R}^d)^{\mathbb{N}}$:

$$\begin{aligned}(1.0.5)\quad & \mu_\beta(\{\gamma, \gamma_0 \in A_0, \gamma_1 \in A_1, ..., \gamma_n \in A_n\}) \\ & \qquad = e^{-n\lambda_\beta} \int_{A_0\times A_1\times\ldots\times A_n} \psi^*_\beta(\gamma_0)\psi_\beta(\gamma_n) e^{-\beta\sum_{k=0}^{n-1} L(\gamma_k,\gamma_{k+1})} d\gamma_0 d\gamma_1 ... d\gamma_n\end{aligned}$$

for all $n \geq 0$, for all Borel sets $A_0, A_1, ..., A_n$. By invariance of the initial distribution, the measure μ_β is invariant under the 1-sided shift on $(\mathbb{R}^d)^{\mathbb{N}}$, so that it can be extended to a σ-invariant measure on $(\mathbb{R}^d)^{\mathbb{Z}}$, that we still denote μ_β.

Actually, the periodicity properties of ψ_β, ψ^*_β and L imply that this measure is invariant under the action of $\mathbb{Z}^d$ on $(\mathbb{R}^d)^{\mathbb{Z}}$. Also, the measure of the fundamental domain $(\mathbb{R}^d)^{\mathbb{Z}^*_-} \times [0,1)^d \times (\mathbb{R}^d)^{\mathbb{Z}^*_+}$ is 1, due to our normalization of ψ_β, ψ^*_β. Identifying the quotient space W to this fundamental domain, we obtain a probability measure (still denoted μ_β) on W, which is σ-invariant, and which we call *the Gibbs measure* for the potential L, at temperature $1/\beta$.

Note that, if we replace the potential $L(x,y)$ by $L(x,y) - u(y) + u(x) + c$, where u is a continuous $\mathbb{Z}^d$-periodic function and c a constant, then the eigenfunctions $\psi_\beta(x)$ and $\psi^*_\beta(x)$ are replaced respectively by $\psi_\beta(x)e^{-\beta u(x)}$ and $\psi^*_\beta(x)e^{\beta u(x)}$, and λ_β is replaced by $\lambda_\beta - \beta c$; the Gibbs measure μ_β is unchanged. According to the usual terminology, we say that two potentials $L(x,y)$ and $\tilde{L}(x,y)$ are cohomologous if there exists a continuous $\mathbb{Z}^d$-periodic function u such that $\tilde{L}(x,y) = L(x,y) - u(y) + u(x)$, and we write $L \sim \tilde{L}$.

Remark 1.0.1. For $n > 0$, we denote $d\mu_\beta\left(\bar{\gamma}|\pi_{[n+1,+\infty)}(\bar{\gamma}), \pi_{(-\infty,0]}(\bar{\gamma})\right)$ the conditional law of $\bar{\gamma}$ knowing $\pi_{[n+1,+\infty)}(\bar{\gamma})$ and $\pi_{(-\infty,0]}(\bar{\gamma})$. What is usually called the "Gibbs property" is a property about the form of conditional measures (see for instance [Ru78], Chapter 1.5):

$$d\mu_\beta\left(\bar{\gamma}|\pi_{[n+1,+\infty)}(\bar{\gamma}), \pi_{(-\infty,0]}(\bar{\gamma})\right) = \frac{\sum_{s\in\mathbb{Z}^d} e^{-\beta(\sum_{k=0}^{n-1} L(\gamma_k,\gamma_{k+1}) + L(\gamma_n,\gamma_{n+1}+s))} d\gamma_1 ... d\gamma_n}{Z^\beta_n(\pi_{[n+1,+\infty)}(\bar{\gamma}), \pi_{(-\infty,0]}(\bar{\gamma}))}.$$

To write this formula we have identified W with the fundamental domain $(\mathbb{R}^d)^{\mathbb{Z}^*_-} \times [0,1)^d \times (\mathbb{R}^d)^{\mathbb{Z}^*_+}$; the term $Z^\beta_n(\pi_{[n+1,+\infty)}(\bar{\gamma}), \pi_{(-\infty,0]}(\bar{\gamma}))$ is a normalization factor.

It is not too hard to check that the measure μ_β constructed above has this property. Moreover, it is proved in [Ru78], Chapter 5.9 (however, in the simpler situation when the configuration space is discrete) that this property actually characterizes the measure. We will not go further into this problem here, as we are not going to use the Gibbs property in this form.

Our aim is now to investigate the existence of a limit for the Gibbs measure μ_β, as $\beta \longrightarrow +\infty$.

We shall say that a sequence $(\mu_n)_{n\in\mathbb{N}}$ of probability measures on W converges to a measure μ if, for every finite interval $I \subset \mathbb{Z}$, for every bounded continuous function f on W_I,

$$\int f(\pi_I \bar{\gamma}) d\mu_n(\bar{\gamma}) \underset{n\to\infty}{\longrightarrow} \int f(\pi_I \bar{\gamma}) d\mu(\bar{\gamma}).$$

We shall prove in Section 3 (Lemma 3.1.5) that, from every sequence $(\mu_{\beta_k})_{k\in\mathbb{N}}$ of Gibbs measures, one can extract a subsequence which converges to a probability measure μ_∞. We ask which measures μ_∞ can be obtained this way.

First, the measure μ_∞, like the μ_βs, has to be invariant under the action of the shift σ. Then, intuition tells us that the measure μ_∞ has to be carried by configurations "minimizing the energy":

Definition 1.0.2. A configuration $\bar{\gamma}$ is called *energy-minimizing* if, for all $m < n$, for all $s \in \mathbb{Z}^d$, for all $(\gamma'_{m+1}, ..., \gamma'_{n-1}) \in (\mathbb{R}^d)^{n-m-1}$,

$$\begin{aligned} L(\gamma_m, \gamma_{m+1}) + L(\gamma_{m+1}, \gamma_{m+2}) + ... + L(\gamma_{n-1}, \gamma_n) \\ \leq L(\gamma_m, \gamma'_{m+1}) + L(\gamma'_{m+1}, \gamma'_{m+2}) + ... + L(\gamma'_{n-1}, \gamma_n + s) \end{aligned}$$

In Section 3 (Lemma 3.1.6), we shall prove that limits of Gibbs measures are carried by energy-minimizing configurations. We shall also prove a theorem, due to Mather in the context of lagrangian dynamical systems ([Ma91]), which says that a σ-invariant probability measure μ on W is carried by energy-minimizing configurations if an only if it minimizes the mean energy $\int \bar{L} d\mu$ amongst all σ-invariant probability measures.

Definition 1.0.3. A σ-invariant probability measure μ on W, achieving the infimum of the mean energy $\int \bar{L} d\mu$ over the set of σ-invariant probability measures, is called an *energy-minimizing measure*.

We introduce the set

$$\mathcal{M} = \overline{\cup_\mu supp\mu} \subset W,$$

(the union runs over energy-minimizing measures), and call it the *Mather set*, in reference to the work of J. Mather in the theory of lagrangian dynamical systems (see Part II).

We will thus show that every limit point of μ_β ($\beta \to \infty$) is an energy-minimizing measure. This fact, known by many, already appears in a paper by Sinai ([Si82]). However, as Sinai's paper precisely shows, there may be several energy-minimizing measures.

Thus, we need a selection principle, telling us which energy-minimizing measures can be obtained as limits of Gibbs measures. The main result of this paper, Theorem 0.0.1, selects an affine subset (possibly not reduced to one point) in the set of energy-minimizing measures.

We now give the assumptions of the theorem, and define the objects entering its statement:

Assumptions : Let $m \leq n$ and $\xi_m, \xi_n \in \mathbb{R}^d$; we introduce the notation

$$(\mathbb{R}^d)^{[m,n],(\xi_m,\xi_n)} = \{(\gamma_k)_{m\leq k\leq n} \in (\mathbb{R}^d)^{[m,n]}, \gamma_m = \xi_m, \gamma_n = \xi_n\}.$$

Recall that we have defined the energy of a sequence $(\gamma_k)_{0\leq k\leq n}$ as $A(\gamma_{|[0,n]}) = \sum_{k=0}^{n-1} L(\gamma_k,\gamma_{k+1})$.

Assumption (A1) For all n, for all endpoints $\xi_0,\xi_n \in \mathbb{R}^d$, the minima of the energy A in the set $(\mathbb{R}^d)^{[0,n],\xi_0,\xi_n}$ are non-degenerate (we mean thereby that the hessian matrix of A at each minimum is non-degenerate). Besides, the number of minimizers is bounded, independently of n,ξ_0,ξ_n.

In order to simplify the writing of the proof, we will assume that there is only one minimizer, for all n,ξ_0,ξ_n.

Assumption (A2) There exists $\varepsilon_0 > 0$ such that, for all $0 \leq \varepsilon \leq \varepsilon_0$, there exists a sequence $(c_n) \in [0,1]^{\mathbb{N}}$ satisfying:
– $\lim_n \frac{\log c_n}{n} = 0$,
and :
– for all $n > 0$, for all $\gamma_0,\gamma_n \in \mathbb{R}^d$ such that

$$\| \gamma_0 - \xi_0 \| \leq c_n\varepsilon$$

$$\| \gamma_n - \xi_n \| \leq c_n\varepsilon$$

for some energy-minimizing configuration $\xi \in (\mathbb{R}^d)^{\mathbb{Z}}$, then there exists a minimizer γ of $A : (\mathbb{R}^d)^{[0,n],\gamma_0,\gamma_n} \longrightarrow \mathbb{R}$ such that $\| \gamma_k - \xi_k \| \leq \varepsilon$ for all $0 \leq k \leq n$.

Change of gage: We will prove in Section 3 (Proposition 3.1.3) that there exists a $\mathbb{Z}^d$-periodic, Lipschitz function u, such that the potential $\tilde{L}(\gamma_0,\gamma_1) = L(\gamma_0,\gamma_1) - u(\gamma_1) + u(\gamma_0) + c$ is nonnegative, and vanishes on the Mather set. As we already mentioned, replacing L by a potential $\tilde{L} \sim L + c$ does not change the definition of the Gibbs measure. In all the definitions given above, we can replace L by a new energy $\tilde{L}$, without changing the definition of energy-minimizing configurations, Mather set, etc... The fact that u is not smooth is not really a problem, since we only need to differentiate the energy functional A on the spaces $(\mathbb{R}^d)^{[0,n],\xi_0,\xi_n}$, that is, for fixed boundary conditions. *Thus, by a change of gage, we may and will assume in the rest of the paper that L is nonnegative, and vanishes on the Mather set.*

After performing this change of gage, we introduce the function

$$h_n(x,y) = \inf_{(\mathbb{R}^d)^{[0,n],(x,y)}} A,$$

defined on $\mathbb{R}^d \times \mathbb{R}^d$.

Assumption (A3) There exists a sequence $B_n \geq 0$ satisfying $\lim_n \frac{\log B_n}{n} = 0$, such that for all n

$$\sup_{\gamma_0} \beta^{d/2} \int_{\mathbb{R}^d} e^{-\beta h_n(\gamma_0,\gamma_n)} d\gamma_n \leq B_n.$$

Assumptions (A1) and (A2) seem merely technical, and it is probably possible to get rid of the second part of (A1) (about the number of minimizers). As to (A3), it says something about the behaviour of the function h_n near its minima, *uniformly in n*. Although these assumptions are not easy to interprete, we can at least check (A2) and (A3) in the case when L is of the form $L(\gamma_0,\gamma_1) = \frac{||\gamma_1-\gamma_0||^2}{2} - V(\gamma_0)$, where V is $\mathbb{Z}^d$-periodic and of class C^3 (Lemma 6.4.2). However, it would be nice to have another set of assumptions which, if not easier to check on examples, would be more conceptual and related to familiar notions of the theory of dynamical systems. In Section 5, we will formulate a conjecture about other possible assumptions.

Metric entropy: Let us now recall the definition of the metric entropy, which comes into play in Theorem 0.0.1. Consider a probability space $(X, \mathcal{B}, \mu)$, and a measurable transformation $T : X \longrightarrow X$ preserving the probability measure μ (meaning that $\mu(T^{-1}A) = \mu(A)$ for every $A \in \mathcal{B}$). One defines the metric entropy of μ with respect to the action of T, denoted $h_T(\mu)$, as follows:

For any partition P of X into a finite number of measurable sets, $X = \sqcup_{i=1}^{k} P_i$, one first defines the entropy of μ with respect to T and the partition P, as

$$h_T(\mu, P) =$$

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{\alpha_0,\dots,\alpha_{n-1}\in\{1,\dots,k\}} -\mu(P_{\alpha_0} \cap T^{-1}P_{\alpha_1} ... T^{-n+1}P_{\alpha_{n-1}}) \log \mu(P_{\alpha_0} \cap T^{-1}P_{\alpha_1} ... T^{-n+1}P_{\alpha_{n-1}}).$$

The existence of the limit can be proved by a sub-additivity argument ([KH95], Chapter 4.3).

Then, $h_T(\mu)$ is defined as

$$h_T(\mu) = \sup_P h_T(\mu, P);$$

the supremum is taken over the set of all finite measurable partitions of X.

In this part, we shall apply this definition to $X = W$ and $T = \sigma$, and μ will be μ_β, μ_∞, or any σ-invariant measure.

Hessian of the energy: The definition of an energy-minimizing configuration implies in particular that if $\bar\gamma$ is such a configuration, then, for all $m < n$, $(\gamma_{m+1}, ..., \gamma_{n-1})$ is a global minimum of the function

$$A(\gamma_m, \gamma'_{m+1}, ..., \gamma'_{n-1}, \gamma_n) = L(\gamma_m, \gamma'_{m+1}) + L(\gamma'_{m+1}, \gamma'_{m+2}) + ... + L(\gamma'_{n-1}, \gamma_n),$$

defined on $(\mathbb{R}^d)^{[m,n],(\gamma_m,\gamma_n)}$.

Let us consider the Hilbert space $l^2(\mathbb{Z}, \mathbb{R}^d) = \{(\gamma_k) \in (\mathbb{R}^d)^{\mathbb{Z}}, \sum_{k\in\mathbb{Z}} ||\gamma_k||^2 < +\infty\}$. Consider the hessian matrix $A''(\bar\gamma) \in \mathcal{L}(l^2(\mathbb{Z}, \mathbb{R}^d))$ of the formal sum

$$A(\bar\gamma) = \sum_{k\in\mathbb{Z}} L(\gamma_k, \gamma_{k+1}).$$

It is an infinite symmetric matrix which can be decomposed into $d \times d$ blocks :

$$A''_{ii} = \partial^2_{22} L(\gamma_{i-1}, \gamma_i) + \partial^2_{11} L(\gamma_i, \gamma_{i+1})$$

$$A''_{i,i+1} = \partial_{21} L(\gamma_i, \gamma_{i+1})$$

and $A''_{i,j} = 0$ for $|j - i| > 1$. This way, the $nd \times nd$ submatrix ${}_nA''(\bar\gamma)$, corresponding to indices $1 \le i, j \le n$, is the hessian matrix of the function A on $(\mathbb{R}^d)^{[0,n+1],(\gamma_0,\gamma_{n+1})}$.

We can now rewrite the statement of Theorem 0.0.1:

Theorem 1.0.4. *Let μ_∞ be a limit point of (μ_β) as $\beta \to +\infty$. Then, μ_∞ is an energy-minimizing measure.*

Moreover, under assumptions (A1), (A2) and (A3), we have

$$h_\sigma(\mu) - \frac{1}{2}\int_W \lim_n \frac{1}{n} \log[{}_nA''(\bar\gamma)] d\mu(\bar\gamma) \le h_\sigma(\mu_\infty) - \frac{1}{2}\int_W \lim_n \frac{1}{n} \log[{}_nA''(\bar\gamma)] d\mu_\infty(\bar\gamma)$$

for any energy-minimizing measure μ.

The proof includes a proof of the existence of the limit $\lim_n \frac{1}{n}\log[{}_nA''(\bar{\gamma})] \in \mathbb{R}$, for every energy-minimizing measure μ, for μ-almost every $\bar{\gamma}$. The metric entropy $h_\sigma(\mu)$, by definition, belongs to $[0,+\infty]$, but we shall see in Section 4 that it is finite in our situation.

Note that the functional

$$\mu \mapsto h_\sigma(\mu) - \frac{1}{2}\int_W \lim_n \frac{1}{n}\log[{}_nA''(\bar{\gamma})]d\mu(\bar{\gamma})$$

is affine, so that our theorem still does not necessarily imply that μ_β converges.

2. Definition of Gibbs measures, and some of their properties.

In this part, we prove the existence and uniqueness of ψ_β, ψ_β^* and λ characterized by (1.0.3), (1.0.4), and we construct the Gibbs measures.

We identify functions on $\mathbb{T}^d$ and $\mathbb{Z}^d$-periodic functions. We also identify Radon measures on $\mathbb{T}^d$ and $\mathbb{Z}^d$-invariant Radon measures on $\mathbb{R}^d$.

We introduce an operator P_β^+, acting on the space of $\mathbb{Z}^d$-periodic continuous functions as follows: if f is such a function, then $P_\beta^+ f$ is defined by :

$$(P_\beta^+ f)(x) = \int_{\mathbb{R}^d} e^{-\beta L(x,y)} f(y)dy,$$

for all $x \in \mathbb{R}^d$. If the continuous function f is nonnegative and does not vanish identically, then $P_\beta^+ f$ is positive.

By duality, P_β also acts on the set of Radon measures on the torus; we define the dual action P_β^{+*} on the set of measures by

$$\int_{\mathbb{T}^d} f d(P_\beta^{+*}\mu) = \int_{\mathbb{T}^d} P_\beta^+ f d\mu,$$

for every continuous function f on the torus, for every measure μ.

We also introduce the operator P_β^- (the adjoint of P_β^+ in $L^2(\mathbb{T}^d)$):

$$(P_\beta^- f)(x) = \int_{\mathbb{R}^d} e^{-\beta L(y,x)} f(y)dy;$$

we let it act on the space of $\mathbb{Z}^d$-periodic continuous functions. We denote P_β^{-*} the dual action on measures.

It is immediate that, for all ν, $P_\beta^{-*}\nu$ has density

$$(D\nu)(x) = \int_{\mathbb{R}^d} e^{-\beta L(x,y)} d\nu(y),$$

whereas $P_\beta^{+*}\nu$ has density

$$(D^*\nu)(x) = \int_{\mathbb{R}^d} e^{-\beta L(y,x)} d\nu(y).$$

(To define the integral on $\mathbb{R}^d$, one considers measures on $\mathbb{T}^d$ as $\mathbb{Z}^d$-invariant Radon measures on $\mathbb{R}^d$.) The operators D, D^* go from the space of measures on the torus to the space of continuous $\mathbb{Z}^d$-periodic functions. In particular, we note that if ν has a density f with respect to Lebesgue measure, then $P_\beta^{+*}\nu$ has density

$$g(x) = \int_{\mathbb{R}^d} e^{-\beta L(y,x)} f(y)dy.$$

In other words, $P_\beta^{+*}(f(x)dx) = (P_\beta^- f)(x)dx$.

We now consider two transformations M^+ and M^- (we forget the dependence on β in the notations), acting on the set of probability measures on the torus the following way :

$$M^+\mu = \frac{P_\beta^{+*}\mu}{\int P_\beta^+ 1 d\mu} \text{ and } M^-\mu = \frac{P_\beta^{-*}\mu}{\int P_\beta^- 1 d\mu}$$

They act continuously on the convex, compact set of probability measures on the torus, endowed with the weak topology. The Schauder fixed point theorem implies that M^+ and M^- both have fixed points.

This exactly means that there exist probability measures μ_β, μ_β^*, and real numbers $\lambda_\beta, \lambda_\beta^*$ such that :

$$P_\beta^{+*}\mu_\beta^* = e^{\lambda_\beta^*}\mu_\beta^* \text{ and } P_\beta^{-*}\mu_\beta = e^{\lambda_\beta}\mu_\beta.$$

The reader will readily check that we have the commutation relations $P_\beta^+ D = DP_\beta^{-*}$ and $P_\beta^- D^* = D^* P_\beta^{+*}$ on the space of measures.

Thus, if $P_\beta^{+*}\mu_\beta^* = e^{\lambda_\beta^*}\mu_\beta^*$, then $D^*\mu_\beta^*$ is an eigenfunction of P_β^- for the eigenvalue $e^{\lambda_\beta^*}$; we denote $\psi_\beta^* = D^*\mu_\beta^*$. Similarly, $D\mu_\beta$ is an eigenfunction of P_β^+ for the eigenvalue e^{λ_β}; we denote it ψ_β.

We can also write:

$$P_\beta^{-*}(\psi_\beta(x)dx) = e^{\lambda_\beta}\psi_\beta(x)dx \text{ and } P_\beta^{+*}(\psi_\beta^*(x)dx) = e^{\lambda_\beta^*}\psi_\beta^*(x)dx.$$

Note that ψ_β and ψ_β^* are positive continuous functions, and that e^{λ_β} (respectively $e^{\lambda_\beta^*}$) is a simple eigenvalue for P_β^+ (respectively P_β^-) in the space of L^2, $\mathbb{Z}^d$-periodic functions. To see this, first note that a $\mathbb{Z}^d$-periodic, L^2 eigenfunction is necessarily continuous. Then, consider a $\mathbb{Z}^d$-periodic continuous function ψ, satisfying

$$P_\beta^+\psi = e^{\lambda_\beta}\psi.$$

Let $\lambda = \sup \psi/\psi_\beta$. Then the function $\lambda\psi_\beta - \psi$ is nonnegative, and by continuity vanishes at one point at least. Besides, it satisfies

$$(\lambda\psi_\beta - \psi)(x) = e^{-\lambda_\beta}\int_{\mathbb{R}^d} e^{-\beta L(x,y)}(\lambda\psi_\beta - \psi)(y)dy$$

for all x. Thus, if $(\lambda\psi_\beta - \psi)(x) = 0$ for some x, then we must have $(\lambda\psi_\beta - \psi)(y) = 0$ for all y; in other words ψ_β and ψ are proportional.

We have proved the beginning of the following proposition:

Proposition 2.0.5. *The eigenvalue e^{λ_β} (respectively $e^{\lambda_\beta^*}$) is a simple eigenvalue for P_β^+ (respectively P_β^-) in $L^2(\mathbb{T}^d)$. Besides, e^{λ_β} is the spectral radius of P_β^+ in $L^2(\mathbb{T}^d)$.*

To prove the last assertion, note that the operator

$$N : f \mapsto \frac{1}{e^{\lambda_\beta}\psi_\beta}P_\beta^+(f\psi_\beta),$$

is stochastic: it fixes the constant function 1. We also say that it is "normalized".

The dual operator N^* on the space of measures fixes the measure $\psi_\beta(x)\psi_\beta^*(x)dx$. It follows from the Cauchy-Schwarz inequality that the norm of N in $L^2(\mathbb{T}^d, \psi_\beta(x)\psi_\beta^*(x)dx)$ is 1, so that its spectral radius is also 1. This now implies that the spectral radius of P_β^+ in $L^2(\mathbb{T}^d)$ is e^{λ_β}.

We know, by the definition of $\psi_\beta^*(x)$, that μ_β^* is proportional to $\psi_\beta^*(x)dx$; and Proposition 2.0.5 implies that μ_β^* is, up to a multiplicative factor, the unique measure such that $P_\beta^{+*}\mu_\beta^* = e^{\lambda_\beta^*}\mu_\beta^*$. A similar property hold for μ_β.

It remains to check that $\lambda_\beta = \lambda_\beta^*$. We prove that ψ_β is (up to a multiplicative constant) the only nonnegative L^1 eigenfunction of P_β^+. Let ψ be a nonnegative eigenfunction of P_β^+; obviously, ψ must be positive, continuous, and associated to a positive eigenvalue e^λ. We write

$$\begin{aligned} e^\lambda \int \psi d\mu_\beta^* &= \int P_\beta^+(\psi) d\mu_\beta^* \\ &= \int \psi d(P_\beta^{+*}\mu_\beta^*) = e^{\lambda_\beta^*} \int \psi d\mu_\beta^* \end{aligned}$$

so that we must have $\lambda = \lambda_\beta^*$, and ψ must be proportional to ψ_β. In particular, $\lambda_\beta = \lambda_\beta^*$.

Rephrasing what has just been done, we can say that the normalised operator N fixes a unique $\mathbb{Z}^d$-invariant Radon measure (up to a multiplicative factor), which is positive, and has density $\psi_\beta(x)\psi_\beta^*(x)$. We normalize it so that $\int_{[0,1)^d} \psi_\beta(x)\psi_\beta^*(x)dx = 1$.

Thus, the measure $\psi_\beta(x)\psi_\beta^*(x)dx$ appears as the unique $\mathbb{Z}^d$-periodic invariant measure for the transition semigroup generated by the following transition density:

$$P(x, dy) = \frac{\psi_\beta(y)}{e^{\lambda_\beta}\psi_\beta(x)} e^{-\beta L(x,y)} dy.$$

A stationary Markovian process on $(\mathbb{R}^d)^{\mathbb{N}}$, of initial distribution $\psi_\beta(x)\psi_\beta^*(x)dx$ and transition probabilities $P(x, dy) = \frac{\psi_\beta(y)}{e^{\lambda_\beta}\psi_\beta(x)} e^{-\beta L(x,y)} dy$, is realized by the measure μ_β on $(\mathbb{R}^d)^{\mathbb{N}}$ defined by :

$$\begin{aligned} (2.0.6) \quad & \mu_\beta(\{\gamma, \gamma_0 \in A_0, \gamma_1 \in A_1, ..., \gamma_n \in A_n\}) \\ & \quad = e^{-n\lambda_\beta} \int_{A_0 \times A_1 \times ... \times A_n} \psi_\beta^*(\gamma_0)\psi_\beta(\gamma_n) e^{-\beta \sum_{k=0}^{n-1} L(\gamma_k, \gamma_{k+1})} d\gamma_0 d\gamma_1 ... d\gamma_n \end{aligned}$$

for all $n \geq 0$, for all Borel sets $A_0, A_1, ..., A_n$. This defines a positive measure μ_β on $(\mathbb{R}^d)^{\mathbb{N}}$, as a consequence from Kolmogorov's extension theorem. By the invariance of the initial distribution, the measure μ_β is invariant under the 1-sided shift on $(\mathbb{R}^d)^{\mathbb{N}}$, so that it can be extended to a σ-invariant measure on $(\mathbb{R}^d)^{\mathbb{Z}}$, that we still denote μ_β. Actually, the periodicity properties of ψ_β, ψ_β^* and L imply that this measure is invariant under the action of $\mathbb{Z}^d$ on $(\mathbb{R}^d)^{\mathbb{Z}}$; finally, the measure of the fundamental domain $(\mathbb{R}^d)^{\mathbb{Z}_-^*} \times [0,1)^d \times (\mathbb{R}^d)^{\mathbb{Z}_+^*}$ is 1, due to our normalization of ψ_β, ψ_β^*.

Identifying the quotient space W to the fundamental domain $(\mathbb{R}^d)^{\mathbb{Z}_-^*} \times [0,1)^d \times (\mathbb{R}^d)^{\mathbb{Z}_+^*}$, we obtain a probability measure (that we still denote μ_β) on W, which is σ-invariant, and which we call *the Gibbs measure* for the potential L, at temperature $1/\beta$.

Note that, if we replace the potential $L(x,y)$ by $L(x,y) - u(y) + u(x) + c$, where u is a continuous $\mathbb{Z}^d$-periodic function and c a constant, then the eigenfunctions $\psi_\beta(x)$ and $\psi_\beta^*(x)$ are replaced respectively by $\psi_\beta(x)e^{-\beta u(x)}$ and $\psi_\beta^*(x)e^{\beta u(x)}$, and λ_β is replaced by $\lambda_\beta - \beta c$; the Gibbs measure μ_β is unchanged.

We now prove a property of "quasi-invariance" by spatial translations of the measure μ_β on W. We denote W_c the subset of W formed of elements $\bar\gamma$ such that: there exists $N \in \mathbb{N}$, there exist $r, s \in \mathbb{Z}^d$, such that $\gamma_k = s$ for $k \geq N$ and $\gamma_k = r$ for $k \leq -N$. Note that, if $\bar\gamma, \bar{\gamma'} \in W$, their sum $\bar\gamma + \bar{\gamma'} = \overline{\gamma + \gamma'} \in W$ is well defined.

Proposition 2.0.6. *For all $\bar z \in W_c$, for all measurable nonnegative continuous function f on W, we have*

$$\int_W f(\bar\gamma + \bar z) d\mu_\beta = \int_W f(\bar\gamma) e^{-\beta \sum_{k\in\mathbb{Z}} (L(\gamma_k - z_k, \gamma_{k+1} - z_{k+1}) - L(\gamma_k, \gamma_{k+1}))} d\mu_\beta.$$

Note that the $\sum_{k\in\mathbb{Z}}$ in the right-hand side is actually a sum on a finite number of terms, since $\bar z \in W_c$.

The reader is invited to compare this property with the definition of Gibbs measures given in [Ha90]; it is proved there (however, in a different situation) that this property actually characterizes the measure. We do not examine this problem here.

Proof. It is sufficient to check it when f depends only on a finite number of coordinates, by which we mean that f is of the form $g \circ \pi_I$ for some finite interval I and for some bounded measurable function g on W^I. Besides, by the σ-invariance of μ_β, it is enough to consider the case $I = [-n, 0]$. Without loss of generality, we may assume that $z_k = 0$ for $k \geq 0$ and $z_k = r \in \mathbb{Z}$ for $k \leq -n$.

To perform the calculation, it is simpler to identify W with the fundamental domain $(\mathbb{R}^d)^{\mathbb{Z}_-^*} \times [0,1)^d \times (\mathbb{R}^d)^{\mathbb{Z}_+^*}$. Now g is a nonnegative measurable function on $(\mathbb{R}^d)^{[-n,-1]} \times [0,1)^d$.

$$\int_W f(\bar\gamma + \bar z) d\mu_\beta = \int_{(\mathbb{R}^d)^{[-n,-1]} \times [0,1)^d} g(\gamma + z) \psi_\beta^*(\gamma_{-n}) \psi_\beta(\gamma_0) e^{-\beta(\sum_{k=-n}^{-1} L(\gamma_k, \gamma_{k+1}))} d\gamma$$

$$= \int_{(\mathbb{R}^d)^{[-n,-1]} \times [0,1)^d} g(\gamma) \psi_\beta^*(\gamma_{-n} - z_{-n}) \psi_\beta(\gamma_0 - z_0) e^{-\beta(\sum_{k=-n}^{-1} L(\gamma_k - z_k, \gamma_{k+1} - z_{k+1}))} d\gamma$$

$$= \int_{(\mathbb{R}^d)^{[-n,-1]} \times [0,1)^d} g(\gamma) \psi_\beta^*(\gamma_{-n}) \psi_\beta(\gamma_0) e^{-\beta(\sum_{k=-n}^{-1} L(\gamma_k - z_k, \gamma_{k+1} - z_{k+1}) - L(\gamma_k, \gamma_{k+1}))} \cdot$$

$$e^{-\beta(\sum_{k=-n}^{-1} L(\gamma_k, \gamma_{k+1}))} d\gamma$$

$$= \int_W f(\bar\gamma) e^{-\beta \sum_{k\in\mathbb{Z}} (L(\gamma_k - z_k, \gamma_{k+1} - z_{k+1}) - L(\gamma_k, \gamma_{k+1}))} d\mu_\beta$$

which proves the proposition. We have used the periodicity of ψ_β and ψ_β^*. $\square$

To end this section, we prove that the Gibbs measure μ_β satisfies a variational principle which looks like a thermodynamical variational principle. Once again we identify W with the fundamental domain $(\mathbb{R}^d)^{\mathbb{Z}_-^*} \times [0,1)^d \times (\mathbb{R}^d)^{\mathbb{Z}_+^*}$. For a probability measure μ on $(\mathbb{R}^d)^{\mathbb{Z}^*_-} \times [0,1)^d \times (\mathbb{R}^d)^{\mathbb{Z}^*_+}$, let us denote $d\mu(\gamma_{-1}|\gamma_{[0,+\infty)})$ the conditional law of γ_{-1} knowing $\gamma_{[0,+\infty)}$.

Proposition 2.0.7. *The measure μ_β maximizes the functional*

$$\mu \mapsto \begin{cases} -\int_{(\mathbb{R}^d)^{\mathbb{Z}^*_-} \times [0,1)^d \times (\mathbb{R}^d)^{\mathbb{Z}^*_+}} \log \left(\frac{d\mu(\gamma_{-1}|\gamma_{[0,+\infty)})}{e^{-\beta L(\gamma_{-1},\gamma_0)} d\gamma_{-1}} \right) d\mu(\gamma) \\ \textit{if } d\mu(\gamma_{-1}|\gamma_{[0,+\infty)}) \textit{ absolutely continuous w.r.t. } d\gamma_{-1}, \\ -\infty \textit{ otherwise} \end{cases}$$

over the set of σ-invariant probability measures on $W \simeq (\mathbb{R}^d)^{\mathbb{Z}^*_-} \times [0,1)^d \times (\mathbb{R}^d)^{\mathbb{Z}^*_+}$.

Proof. We use the following convexity inequality : if μ, ν are probability measures on some space X, and if μ is absolutely continuous with respect to ν, then

$$-\int_X \log(\frac{d\mu}{d\nu}) d\mu \leq 0,$$

with equality for $\mu = \nu$.

Note that

$$d\mu_\beta(\gamma_{-1}|\gamma_{[0,+\infty)}) = \frac{\psi^*_\beta(\gamma_{-1})}{\psi^*_\beta(\gamma_0)} e^{-\lambda_\beta - \beta L(\gamma_{-1},\gamma_0)} d\gamma_{-1}.$$

Thus, for any σ-invariant probability measure μ on W such that $d\mu(\gamma_{-1}|\gamma_{[0,+\infty)})$ is absolutely continuous with respect to the Lebesgue measure $d\gamma_{-1}$, we have

$$-\int d\mu(\gamma_{-1}|\gamma_{[0,+\infty)}) \log \left(\frac{d\mu(\gamma_{-1}|\gamma_{[0,+\infty)})}{\psi^*_\beta(\gamma_{-1}) \psi^{*-1}_\beta(\gamma_0) e^{-\lambda_\beta - \beta L(\gamma_{-1},\gamma_0)} d\gamma_{-1}} \right) \leq 0$$

for μ-almost every $\gamma_{[0,+\infty)}$, with equality for $\mu = \mu_\beta$. Integrating with respect to $\gamma_{[0,+\infty)}$, we obtain

$$-\int d\mu(\gamma) \log \left(\frac{d\mu(\gamma_{-1}|\gamma_{[0,+\infty)})}{\psi^*_\beta(\gamma_{-1}) \psi^{*-1}_\beta(\gamma_0) e^{-\lambda_\beta - \beta L(\gamma_{-1},\gamma_0)} d\gamma_{-1}} \right)$$

$$= -\lambda_\beta + \int (\log \psi^*_\beta(\gamma_{-1}) - \log \psi^*_\beta(\gamma_0)) d\mu - \int d\mu(\gamma) \log \left(\frac{d\mu(\gamma_{-1}|\gamma_{[0,+\infty)})}{\psi^*_\beta(\gamma_{-1}) \psi^{*-1}_\beta(\gamma_0) d\gamma_{-1}} \right) \leq 0$$

with equality for $\mu = \mu_\beta$.

If μ is σ-invariant, $\int (\log \psi^*_\beta(\gamma_{-1}) - \log \psi^*_\beta(\gamma_0)) d\mu = 0$. Thus, we get

$$-\int \log \left(\frac{d\mu(\gamma_{-1}|\gamma_{[0,+\infty)})}{e^{-\beta L(\gamma_{-1},\gamma_0)} d\gamma_{-1}} \right) d\mu(\gamma) \leq \lambda_\beta$$

for all σ-invariant measures μ, with equality for $\mu = \mu_\beta$. $\square$

After multiplication by $-1/\beta$, it would be tempting to decompose the functional of Proposition 2.0.7 in the form :

$$\int \bar{L} d\mu - \frac{1}{\beta} H(\mu)$$

where H would be the functional defined by

$$H(\mu) = \int \log \left(\frac{d\mu(\gamma_{-1}|\gamma_{[0,+\infty)})}{d\gamma_{-1}} \right) d\mu$$

Then, we would call $\int \bar{L} d\mu$ the mean energy, and look as $H(\mu)$ as a kind of entropy, so that $\int L d\mu - \frac{1}{\beta} H(\mu)$ would be a free energy.

However, this decomposition does not always make sense, since both terms may be infinite.

It would be interesting to see if Theorem 0.0.1 can be derived directly from Proposition 2.0.7 by letting $\beta \to +\infty$, and expanding the functional of Proposition 2.0.7 in powers of β.

Remark 2.0.8. The situation is considerably simpler when the configuration space is of the form $B^{\mathbb{Z}}$, where B is a *finite alphabet*. In that situation, the Gibbs measure μ_β for a potential L, at temperature $1/\beta$, minimizes the free energy $\int L d\mu - \frac{1}{\beta} h_\sigma(\mu)$ ([Ru78]); from the fact that h_σ is a bounded, lower semi-continuous functional, one can deduce directly that any limit of $(\mu_\beta)_{\beta\to\infty}$ is an energy-minimizing measure, and maximizes the entropy amongst energy-minimizing measures. See, for instance, Theorem 29 in [CLT01]; in that reference, the action of the shift on $B^{\mathbb{Z}}$ arises as the coding of an expansive map of the circle, and the potential L (depending on infinitely many coordinates) is the logarithm of the jacobian of the map; so that "energy-minimizing measures" are measures of minimal Lyapunov exponent.

When $B = \mathbb{R}^d$, difficulties arise from the fact that $H(\mu)$ is not a bounded functional and is not the metric entropy $h_\sigma(\mu)$; also, Lyapunov exponents appear when analyzing the gaussian fluctuations of the energy.

3. Proof of Theorem 0.0.1.

To start with, we give the general idea of the proof, and explain the role of assumptions (A1), (A2), (A3). These ideas are quite classical, their technical implementation is performed in Section 3.2.

On a finite-dimensional configuration space $(\mathbb{R}^d)^n$, let A be an energy functional, and

$$d\mu_\beta(x) = \frac{e^{-\beta A(x)}dx}{\int e^{-\beta A(y)}dy}$$

the associated "Gibbs measure". Assume that A has only non-degenerate minima $(x_i)_{i\in\mathbb{N}}$. Then, as $\beta \longrightarrow +\infty$, μ_β concentrates on the minima of A; more precisely, it converges to

$$\mu_\infty = (\sum_i [A''(x_i)]^{-1/2})^{-1} \sum_i [A''(x_i)]^{-1/2}\, \delta_{x_i}$$

where δ_{x_i} is the Dirac mass at x_i and $A''(x_i)$ the Hessian of A at x_i.

In other words, μ_∞ is the measure maximizing

$$-\sum_i \mu(x_i)\log\mu(x_i) - \frac{1}{2}\int \log[A''(x)]d\mu(x),$$

amongst measures carried by the minima of A, x_i.

We want to apply exactly this idea, when the configuration space, $(\mathbb{R}^d)^{\mathbb{Z}}/\mathbb{Z}^d$, is now infinite-dimensional. The difficulty is that both notions of Gibbs measures and metric entropy are defined, from the finite-dimensional model described above, by taking the thermodynamical limit $n \longrightarrow \infty$. We are in a situation where n goes first to ∞ (the thermodynamical limit), and then β (the low temperature limit). If we could first let β tend to ∞, then n, we would be done.

Assumptions (A1), (A2), (A3) contain what we need to apply the heuristics described above:

– non-degeneracy of the minimizers of the energy.

– technical possibility to reverse the orders of the two limits $n \longrightarrow +\infty$ and $\beta \longrightarrow +\infty$.

In the following, the space $\mathbb{R}^d$ is endowed with its canonical Euclidean structure, we denote $||.||$ the associated norm, $||.||_\infty$ the norm $||x||_\infty = \max_{i=1,..,d} |x^i|$, and $||.||_1$ the norm $||x||_1 = \sum_{i=1,..,d} |x^i|$.

3.1. **Preliminary results.**

Lemma 3.1.1. *Let $(\beta_k)_{k\geq 0}$ be a sequence such that $\beta_k \longrightarrow +\infty$. Then the families of functions $(-\frac{1}{\beta_k}\log\psi_{\beta_k})_k$ and $(-\frac{1}{\beta_k}\log\psi^*_{\beta_k})_k$ are equicontinuous.*

Proof. The function ψ_β satisfies

$$\begin{aligned}\psi_\beta(x) &= e^{-\lambda_\beta}\int_{\mathbb{R}^d} e^{-\beta L(x,y)}\psi_\beta(y)dy \\ &= e^{-\lambda_\beta}\int_{[0,1)^d} (\sum_{s\in\mathbb{Z}^d} e^{-\beta L(s+x,y)})\psi_\beta(y)dy\end{aligned}$$

The potential L being superlinear, there exists $M > 0$ such that

$$\sum_{||s||>M} e^{-\beta L(s+x,y)} \leq e^{-\beta \inf_{s\in\mathbb{Z}^d} L(s+x,y)}$$

for all $x, y \in [0,1)^d$. It follows that

$$\sum_{s\in\mathbb{Z}^d} e^{-\beta L(s+x,y)} \leq \sum_{||s||\leq M} e^{-\beta L(s+x,y)} + e^{-\beta\inf_s L(s+x,y)} \leq (M^d+1)e^{-\beta\inf_s L(s+x,y)}$$

so that

$$\limsup_{\beta\longrightarrow\infty} \frac{1}{\beta}\log(\sum_{s\in\mathbb{Z}^d} e^{-\beta L(s+x,y)}) \leq -\inf_s L(s+x,y).$$

On the other hand, since

$$\sum_{s\in\mathbb{Z}^d} e^{-\beta L(s+x,y)} \geq e^{-\beta\inf_s L(s+x,y)},$$

one has the lower bound:

$$\liminf_{\beta\longrightarrow\infty} \frac{1}{\beta}\log(\sum_{s\in\mathbb{Z}^d} e^{-\beta L(s+x,y)}) \geq -\inf_s L(s+x,y),$$

so that

$$\lim_{\beta\longrightarrow\infty} \frac{1}{\beta}\log(\sum_{s\in\mathbb{Z}^d} e^{-\beta L(s+x,y)}) = -\inf_s L(s+x,y). \tag{3.1.1}$$

Besides, the argument proves that the convergence is uniform in x, y.

For $x, y \in \mathbb{R}^d$, we denote $I(x,y) = \inf_{s\in\mathbb{Z}^d} L(s+x,y)$. This function is $\mathbb{Z}^d$-periodic in both variables; we show that it is a lipschitz function on $\mathbb{T}^d\times\mathbb{T}^d$. Because $L(x,y)$ goes to infinity as $||x-y|| \longrightarrow +\infty$, there exists $M > 0$ such that

$$I(x,y) = \inf_{||s||\leq M} L(s+x,y)$$

for all $x, y \in (0,1)^d$. Let us consider $x, y, x', y' \in (0,1/2)^d$. Assume that $I(x,y) = L(s_0+x,y)$ with $||s_0|| \leq M$. Then

$$I(x',y') \leq L(s_0+x',y') \leq L(s_0+x,y)+C(||x-x'||+||y-y'||) = I(x,y)+C(||x-x'||+||y-y'||)$$

where C is a bound on the norm of the derivative of L on $(-M-1, M+1)^d \times (0, 1/2)^d$.

Since (x, y) and (x', y') play symmetric roles, we have proved that I is lipschitz on $(0, 1/2)^d \times (0, 1/2)^d$, for the euclidean distance. Besides, in $(0, 1/2)^d \times (0, 1/2)^d$ the distance $||x - x'|| + ||y - y'||$ coincides with the distance of their images in the torus, $d_{\mathbb{T}^d}(\bar{x}, \bar{x}') + d_{\mathbb{T}^d}(\bar{y}, \bar{y}')$. This way, we can cover $\mathbb{T}^d \times \mathbb{T}^d$ by a finite number of charts in which I is lipschitz.

We now write

$$\begin{aligned}\frac{1}{\beta}\log\psi_\beta(x) - \frac{1}{\beta}\log\psi_\beta(y) = \quad & \frac{1}{\beta}\log\left(\frac{\int_{[0,1)^d}(\sum_{s\in\mathbb{Z}^d} e^{-\beta L(s+x,z)})\psi_\beta(z)dz}{\int_{[0,1)^d}(\sum_{s\in\mathbb{Z}^d} e^{-\beta L(s+y,z)})\psi_\beta(z)dz}\right) \\ & \leq \frac{1}{\beta}\log\left(\sup_z \frac{\sum_{s\in\mathbb{Z}^d} e^{-\beta L(s+x,z)}}{\sum_{s\in\mathbb{Z}^d} e^{-\beta L(s+y,z))}}\right) \\ & \underset{\beta\longrightarrow+\infty}{\longrightarrow} \sup_{z\in[0,1)^d} I(x,z) - I(y,z)\end{aligned}$$

the last line being a consequence of (3.1.1). Besides, the convergence is uniform in x, y.

Since I is lipschitz, there exists C such that $\sup_z I(x, z) - I(y, z) \leq C d_{\mathbb{T}^d}(\bar{x}, \bar{y})$.

Thus, for all $\varepsilon > 0$, there exists K such that, for all $k > K$,

$$\frac{1}{\beta_k}\log\psi_{\beta_k}(x) - \frac{1}{\beta_k}\log\psi_{\beta_k}(y) \leq \varepsilon + C d_{\mathbb{T}^d}(\bar{x}, \bar{y})$$

Since x and y play symmetric roles, it follows that $(\frac{1}{\beta_k}\log\psi_{\beta_k})$ is a uniformly equicontinuous family of $\mathbb{Z}^d$-periodic functions.

A similar argument yields the result for $(\frac{1}{\beta_k}\log\psi^*_{\beta_k})$. □

We introduce the value

$$c = -\inf\{\int \bar{L}d\mu, \mu \text{ a } \sigma\text{-invariant probability measure on } W\}.$$

Definition 3.1.2. (1) We say that a σ-invariant probability measure μ on W is energy-minimizing if $\int_W \bar{L}d\mu = -c$.

(2) We say that a configuration $\bar{\gamma}$ is strongly minimizing if, for all $m < n$, for all $m' < n'$, for all $\gamma'_{m'}, ..., \gamma'_{n'}$ such that $\gamma'_{m'} = \gamma_m$ and $\gamma'_{n'} = \gamma_n + s$ for some $s \in \mathbb{Z}^d$,

$$L(\gamma_m, \gamma_{m+1}) + ... + L(\gamma_{n-1}, \gamma_n) + c(m-n) \leq L(\gamma'_{m'}, \gamma'_{m'+1}) + ... + L(\gamma_{n'-1}, \gamma_{n'}) + c(m'-n').$$

Obviously, a strongly minimizing configuration is minimizing (but the converse is not necessarily true, see [Ber02] for a discussion of this issue in the context of Lagrangian dynamical systems).

We say that a configuration $\bar{\gamma}$ is recurrent if, for all k, for every $\varepsilon > 0$, there exists an infinity of positive indices j and of negative indices j such that

$$d_{W_1}(\overline{(\gamma_k, \gamma_{k+1})}, \overline{(\gamma_j, \gamma_{j+1})}) \leq \varepsilon$$

The Poincaré recurrence theorem implies that a configuration lying in the support of a σ-invariant probability measure on W is recurrent.

Proposition 3.1.3. *There exists a lipschitz $\mathbb{Z}^d$-periodic function u such that*

$$u(x) + L(x, y) - u(y) + c \geq 0, \text{ for all } x, y \in \mathbb{R}^d,$$

and $u(x) + L(x,y) - u(y) + c = 0$ *if there exists a configuration* $\bar\gamma \in W$ *which is recurrent strongly minimizing, or which lies in the support of an energy-minimizing measure, such that* $\gamma_0 = x, \gamma_1 = y$.

Proof. By Lemma 3.1.1, we can find a sequence $\beta_k \longrightarrow +\infty$ and a continuous $\mathbb{Z}^d$-periodic function u such that

$$\frac{1}{\beta_k}\log\psi^*_{\beta_k} \underset{k\longrightarrow+\infty}{\longrightarrow} -u$$

uniformly.

We may also assume that $\frac{\lambda_{\beta_k}}{\beta_k}$ converges in $\mathbb{R}\cup\{-\infty,+\infty\}$, say to a limit λ.

We use the following

Lemma 3.1.4. *Assume that* $(u_\beta)_{\beta>0}$ *is a family of functions on* $\mathbb{T}^d$ *which converges uniformly to a continuous function* u *as* $\beta \longrightarrow +\infty$. *Then*

$$\lim_{\beta\longrightarrow\infty}\frac{1}{\beta}\log\int_{\mathbb{T}^d} e^{\beta u_\beta(x)}dx = \sup_{x\in\mathbb{T}^d} u(x)$$

The proof goes as follows : the inequality $u_\beta \le u+\varepsilon \le \sup u + \varepsilon$, which holds for every $\varepsilon > 0$ and for β large enough, yields the upper bound on the limsup. The fact that $u_\beta \ge u - \varepsilon$, for every $\varepsilon > 0$ and for β large enough, and that $u \ge \sup u - \varepsilon$ on a set of positive Lebesgue measure, yields the lower bound on the liminf.

Now, taking the log of both sides of the equality

$$\psi^*_\beta(x) = e^{-\lambda_\beta}\int_{[0,1)^d}(\sum_{s\in\mathbb{Z}^d} e^{-\beta L(y,x+s)})\psi^*_\beta(y)dy,$$

dividing by β, and passing to the limit for the subsequence (β_k), we get

$$-u(x) = -\lambda - \inf_y\{I(y,x)+u(y)\} = -\lambda - \inf_y\{L(y,x)+u(y)\}$$

or $u(x) = \lambda + \inf_y\{L(y,x)+u(y)\}$. Since u is continuous, this implies already that λ is finite.

Imitating the notation of Fathi in [Fa97-1], we introduce the transformation $T^- : C(\mathbb{T}^d,\mathbb{R}) \longrightarrow C(\mathbb{T}^d,\mathbb{R})$:

$$T^-v(x) = \inf_y\{L(y,x)+v(y)\}.$$

If v is continuous, then T^-v is lipschitz. We admit this fact, whose proof is similar to the proof that I is lipschitz.

Thus we have $u = T^-u + \lambda$. This implies that u is lipschitz; besides, λ is necessarily equal to the critical value, c. This result is due to Fathi for a continuous time Lagrangian system ([Fa97-1]). Since the full proof is still unpublished, we give a general idea of it:

– The equality $u = T^-u+\lambda$ implies that $u(\gamma_0)+L(\gamma_0,\gamma_1)-u(\gamma_1)+\lambda \ge 0$, for all $\gamma_0,\gamma_1\in\mathbb{R}^d$. Taking the integral over an arbitrary σ-invariant probability measure μ on W yields $\lambda + \int \bar L d\mu \ge 0$, so that $\lambda \ge c$.

– Next, Fathi proves that $u = T^-u + \lambda$ is equivalent to the following: for all $\gamma_0 \in \mathbb{R}^d$, there exists a sequence $(\gamma_k)_{k\le 0}$ such that, for all $k \le 0$,

$$u(\gamma_k) + \sum_{j=k}^{-1} L(\gamma_j,\gamma_{j+1}) - u(\gamma_0) + |k|\lambda = 0.$$

Let us extend this sequence to a configuration $(\gamma_k)_{k\in\mathbb{Z}}$. Fathi shows that the sequence of Birkhoff sums

$$\frac{1}{|k|+1}\sum_{j=k}^{0}\delta_{\sigma^k\bar{\gamma}}$$

admits a convergent subsequence as $k \longrightarrow -\infty$, and that the limit μ is a σ-invariant probability measure satisfying $\int \bar{L}d\mu = -\lambda$. Thus, $\lambda = c$.

Now, let $\bar{\gamma}$ be a strongly minimizing recurrent configuration, and assume that there exists j such that $u(\gamma_j) - u(\gamma_{j+1}) + c + L(\gamma_j, \gamma_{j+1}) > 0$; for instance that $u(\gamma_1) - u(\gamma_2) + c + L(\gamma_1, \gamma_2) > 2\varepsilon$ for some $\varepsilon > 0$. Since $\bar{\gamma}$ is recurrent, we can find k arbitrarily large, such that $\overline{(\gamma_k, \gamma_{k+1})}$ comes arbitrarily close to $\overline{(\gamma_0, \gamma_1)}$; this implies that $(\gamma_0, \bar{\gamma}_{k+1})$ comes arbitrarily close to $(\gamma_0, \bar{\gamma}_1)$. Thus, for some k,

$$\begin{aligned} L(\gamma_0, \gamma_{k+1}) - u(\gamma_{k+1}) + u(\gamma_0) + c \leq \quad & L(\gamma_0, \gamma_1) - u(\gamma_1) + u(\gamma_0) + c + \varepsilon \\ < \quad & \sum_{j=0}^{k}(u(\gamma_j) - u(\gamma_{j+1}) + c + L(\gamma_j, \gamma_{j+1})) \\ & = \sum_{j=0}^{k} L(\gamma_j, \gamma_{j+1}) - u(\gamma_{k+1}) + u(\gamma_0) + (k+1)c \end{aligned}$$

To account for the second inequality, note that in the sum, all the terms are non-negative, the first one is $L(\gamma_0, \gamma_1) - u(\gamma_1) + u(\gamma_0) + c$, and the second one is greater than 2ε.

But this is in contradiction with the fact that $\bar{\gamma}$ is strongly minimizing. So, we must have $u(\gamma_j) - u(\gamma_{j+1}) + c + L(\gamma_j, \gamma_{j+1}) = 0$ for all j if $\bar{\gamma}$ is strongly minimizing and recurrent.

To prove the last assertion of the lemma, we know that, for every $\bar{\gamma} \in W$,

$$u(\gamma_1) - u(\gamma_0) \leq c + \bar{L}(\bar{\gamma})$$

and that

$$0 = \int_W (u(\gamma_1) - u(\gamma_0))d\mu(\bar{\gamma})) = c + \int_W \bar{L}(\bar{\gamma})d\mu(\bar{\gamma})$$

if μ is an energy-minimizing measure (in particular, σ-invariant). So, we must have equality $u(\gamma_0) - u(\gamma_1) = c + \bar{L}(\bar{\gamma})$ if $\bar{\gamma}$ lies in the support of an energy-minimizing measure.

□

We say that a sequence $(\mu_{\beta_k})_{k\in\mathbb{N}}$ $(\beta_k \to +\infty)$ converges to a measure μ_∞ if, for every k, for every bounded continuous function f on W_k,

$$\int f(\pi_k\bar{\gamma})d\mu_{\beta_k}(\bar{\gamma}) \underset{k\longrightarrow+\infty}{\longrightarrow} \int f(\pi_k\bar{\gamma})d\mu_\infty(\bar{\gamma})$$

Lemma 3.1.5. *Let (β_k) be a sequence such that $\beta_k \longrightarrow +\infty$. Then it is possible to extract from the sequence $(\mu_{\beta_k})_{k\in\mathbb{Z}}$ a subsequence which converges to a σ-invariant probability measure on W.*

Proof. We need to show that, for all I finite interval of $\mathbb{Z}$, for all $\varepsilon > 0$, there exists a compact subset $K \subset W_I$ such that $\mu_\beta \pi_I^{-1}(W_I \setminus K) \leq \varepsilon$ for β large enough. Once this is proved, we can apply Prohorod's theorem and a diagonal extraction procedure to find a subsequence of $\mu_{\beta_{k_n}}$ such that $\mu_{\beta_{k_n}}\pi_I^{-1}$ converges in the weak* topology, for all I :

$$\mu_{\beta_{k_n}}\pi_I^{-1} \longrightarrow \mu_\infty^I$$

Besides, if $J \subset I$, then $\mu_\infty^I.\pi_J^{-1} = \mu_\infty^J$ since $\mu_{\beta_{k_n}}\pi_I^{-1}\pi_J^{-1} = \mu_{\beta_{k_n}}\pi_J^{-1}$. Kolmogorov's extension theorem ensures that there exists a probability measure μ_∞ on W such that $\mu_\infty.\pi_I^{-1} = \mu_\infty^I$, for all I. Finally, μ_∞ is the limit of $\mu_{\beta_{k_n}}$.

We now prove the first claim. Let I be finite interval of $\mathbb{Z}$; since μ_β is σ-invariant, we can assume that $I = [0, n]$. There exists K such that, if $|x^i - y^i| > K$ for some $i = 1, ..., d$, then $L(x, y) \geq ||x - y||_1$.

By Lemma 3.1.1, if we normalize ψ_β and ψ_β^* such that $\psi_\beta(0) = \psi_\beta^*(0) = 1$, there exists M such that $\psi_\beta(x) \leq e^{\beta M}$, $\psi_\beta^*(x) \leq e^{\beta M}$ (for all x), and $\int_{\mathbb{T}^d} \psi_\beta(y)\psi_\beta^*(y)dy \geq e^{-\beta M}$.

Thus, from the expression (2.0.6) defining μ_β, for all $i = 1, ..., d, k = 0, ..., n$, and by the σ-invariance of μ_β,

$$\mu_\beta\pi_I^{-1}(\{|\gamma_{k+1}^i - \gamma_k^i| > K\}) \leq e^{3\beta M} \int_{\gamma_0 \in [0,1)^d, |\gamma_1^i - \gamma_0^i| > K} e^{-\beta||\gamma_1 - \gamma_0||_1} d\gamma_0 d\gamma_1 \leq e^{3\beta M} \frac{Cst.e^{-\beta K}}{\beta}$$

If we take K large enough, this term tends to zero as $\beta \longrightarrow \infty$, thus showing that $\mu_\beta\pi_I^{-1}$ concentrates on the compact set $\{||\gamma_{k+1} - \gamma_k||_\infty \leq K, \text{ for all } k\}$.

□

Lemma 3.1.6. *Let (β_k) be a sequence such that $\beta_k \longrightarrow +\infty$ and such that the sequence $(\mu_{\beta_k})_{k\in\mathbb{Z}}$ converges to a probability measure μ_∞ on W. Then each configuration $\bar{\gamma}$ in the support of μ_∞ is a minimizing configuration.*

Proof. Assume, on the contrary, that there exists a configuration $\bar{\xi}$ in the support of μ_∞, which is not minimizing. There exists $\bar{z} \in W_c$ such that

$$\sum_{k\in\mathbb{Z}} (L(\xi_k - z_k, \xi_{k+1} - z_{k+1}) - L(\xi_k, \xi_{k+1})) < 0.$$

By continuity of L, there exists a neighbourhood B of $\bar{\xi}$, such that, for all $\bar{\gamma} \in B$,

$$\sum_{k\in\mathbb{Z}} (L(\gamma_k - z_k, \gamma_{k+1} - z_{k+1}) - L(\gamma_k, \gamma_{k+1})) < 0.$$

Proposition 2.0.6 implies that

$$\int_W \chi_B(\bar{\gamma})d\mu_\beta(\bar{\gamma}) = \int_W \chi_B(\bar{\gamma} - \bar{z})e^{-\beta\sum_{k\in\mathbb{Z}}(L(\gamma_k - z_k, \gamma_{k+1} - z_{k+1}) - L(\gamma_k, \gamma_{k+1}))} d\mu_\beta(\bar{\gamma}).$$

The right-hand side term tends to 0, and so $\mu_\beta(B) \longrightarrow 0$, which contradicts the fact that $\bar{\xi}$ is in the support of μ_∞. □

Lemma 3.1.7. *The set of energy-minimizing configurations is relatively compact in W.*

Proof. A subset $K \subset W$ is relatively compact if and only if there exists M such that, for all $\bar{\gamma}_k \in K$, $\| \gamma_{k+1} - \gamma_k \| \leq M$ (remember that the topology is defined by the distance d_W introduced in Section 1).

Let $A = \sup\{|L(x,y)|, ||x-y||_\infty \leq 1\}$. Because L grows superlinearly, there exists M such that $||x-y|| > M \Longrightarrow L(x,y) > 2A$. For all $x, y \in \mathbb{R}^d$, there exists $s \in \mathbb{Z}^d$ such that $||x-y-s||_\infty \leq 1$; thus,

$$||x-y|| > M \Longrightarrow (\exists s \in \mathbb{Z}^d, L(x-s,y) < L(x,y)),$$

Let $\gamma \in (\mathbb{R}^d)^{\mathbb{Z}}$ be such that $||\gamma_{k+1} - \gamma_k|| > M$ for some k, say $k = 0$ for instance. Let $s \in \mathbb{Z}^d$ be such that $L(\gamma_0 - s, \gamma_1) < L(\gamma_0, \gamma_1)$, then

$$L(\gamma_0 - s, \gamma_1) + L(\gamma_1, \gamma_2) < L(\gamma_0, \gamma_1) + L(\gamma_1, \gamma_2)$$

so γ is not energy-minimizing. Thus, we have found M such that

$$\bar{\gamma} \text{ is energy-minimizing } \Rightarrow ||\gamma_{k+1} - \gamma_k|| \leq M \text{ for all } k \in \mathbb{Z}.$$

□

We can now prove a result, due to Mather ([Ma91]) in the context of lagrangian dynamical systems:

Theorem 3.1.1. *(a) Let μ be a σ-invariant probability measure on W. The three following assertions are equivalent:*
(i) μ is energy-minimizing.
(ii) the support of μ contains only strongly minimizing configurations.
(iii) the support of μ contains only minimizing configurations.

(b)Energy-minimizing measures do exist.

Proof. To prove the theorem, we note that the definition of an energy-minimizing measure and of a (strongly) minimizing configuration is unchanged if we replace L by $\tilde{L} \sim L+c$. Thus, using Proposition 3.1.3, we may assume that $\bar{L}$ is a nonnegative function, that

$$\inf\{\int \bar{L} d\mu, \mu \text{ a } \sigma - \text{invariant p. m.}\} = 0,$$

and that $\bar{L}$ vanishes on strongly minimizing configurations and on the support of any energy-minimizing measure. In this situation, it is clear that a σ-invariant measure μ is energy-minimizing if and only if $\bar{L}$ vanishes on its support, and a configuration $\bar{\gamma}$ is strongly minimizing if and only if $\bar{L}(\sigma^k \bar{\gamma}) = 0$ for all k.

This proves that $(i) \Leftrightarrow (ii)$.

It remains to prove that (iii)⇒(i). We note that Lemma 3.1.7 implies that μ is compactly supported. We show that the ergodic components of μ are energy-minimizing. Let $\bar{\gamma}$ be a point in the support of μ, such that the sequence of probability measures

$$\frac{1}{n} \sum_{k=0}^{n-1} \delta_{\sigma^k \bar{\gamma}}$$

converges weakly to a σ-invariant probability measure $\mu_{\bar{\gamma}}$ (this happens for μ-almost every point, by Birkhoff's ergodic theorem).

Let ν be an arbitrary σ-invariant probability measure, that we may assume ergodic, without loss of generality. Let ξ be a point such that the sequence of probability measures $\frac{1}{n} \sum_{k=0}^{n-1} \delta_{\sigma^k \bar{\xi}}$ converges weakly to ν. We choose representatives γ, ξ such that $||\gamma_0 - \xi_0|| \leq 1$. For all n, there exists $s_n \in \mathbb{Z}^d$ such that $||\gamma_n + s_n - \xi_n||_\infty \leq 1$. Assertion (ii) tells us that

$$L(\gamma_0, \gamma_1) + L(\gamma_1, \gamma_2) + ... + L(\gamma_{n-1}, \gamma_n) \leq L(\gamma_0, \xi_1) + L(\xi_1, \xi_2) + ... + L(\xi_{n-1}, \gamma_n + s_n)$$

We now choose a $C > 0$ such that $\nu(\{\bar{\gamma}, ||\gamma_1 - \gamma_0||_\infty \leq C\}) > 0$. We have $||\xi_n - \xi_{n-1}||_\infty \leq C$ for an infinity of $n \in \mathbb{N}$; we may also assume, without loss of generality,that $||\xi_1 - \xi_0|| \leq C$.

Now,

$$L(\gamma_0, \xi_1) + L(\xi_1, \xi_2) + ... + L(\xi_{n-1}, \gamma_n + s_n) \quad \leq \quad L(\xi_0, \xi_1) + L(\xi_1, \xi_2) + ... + L(\xi_{n-1}, \xi_n) + 2M$$

where M is an upper bound of $||\partial_1 L||$ and $||\partial_2 L||$ on the set $\{(x,y), ||x-y||_\infty \leq C+1\}$.

Dividing both sides by n, and letting $n \longrightarrow +\infty$, we get

$$\int_W \bar{L} d\mu_{\bar{\gamma}} \leq \int_W \bar{L} d\nu.$$

Thus, we have proved that all ergodic components of μ are energy-minimizing, implying that μ itself is energy-minimizing.

Assertion (b) follows from Lemmas 3.1.5 and 3.1.6. $\square$

We denote $\mathcal{M}$ the closure of the union of supports of energy minimizing measures:

$$\mathcal{M} = \overline{\cup_{\mu \text{ en. min.}} \, supp\mu} \subset W,$$

and call it "Mather set", for reasons explained in Section 4. It is a compact, σ-invariant subset of W.

Thanks to Proposition 3.1.3, we can operate a change of gage on the potential so that it becomes nonnegative, and vanishes on the Mather set. Although the change of gage is only lipschitz, the functions $\sum_{i=0}^{n-1} L(\gamma_i, \gamma_{i+1})$ remain of class C^3 with respect to the variables $\gamma_1, ..., \gamma_{n-1}$.

3.2. **The proof of Theorem 0.0.1.** We begin by proving a subaddivity property for the determinants $[{}_nA''(\bar{\gamma})]$, when $\bar{\gamma}$ is a minimizing configuration.

Recall that $A''(\bar{\gamma})$ is the hessian matrix at $\bar{\gamma} \in W$ of the (formal) sum $A(\bar{\gamma}) = \sum_{k \in \mathbb{Z}} L(\gamma_k, \gamma_{k+1})$. We see $A''(\bar{\gamma})$ as an infinite tridiagonal symmetric matrix, which can be decomposed into $d \times d$ blocks $(A''(i,j))_{i,j \in \mathbb{Z}}$:

$$A''(i,i) = \partial^2_{22} L(\gamma_{i-1}, \gamma_i) + \partial^2_{11} L(\gamma_i, \gamma_{i+1}),$$

$$A''(i, i+1) = \partial_{21} L(\gamma_i, \gamma_{i+1}),$$

and $A''(i,j) = 0$ if $|j - i| > 1$. The $nd \times nd$ submatrix ${}_nA''(\gamma)$, corresponding to indices $1 \leq i, j \leq n$, is the hessian matrix of the action $A(\gamma_{|[0,n+1]})$ with respect to the variables $\gamma_1, \cdots, \gamma_n$.

Notations: – In what follows, we denote $[M]$ the determinant of a square matrix M of any dimension.

– unless stated otherwise, we shall always represent matrices in d-block form; for instance, if M is an $nd \times nd$ matrix, M_{ij} or $M(i,j)$ $(1 \leq i, j \leq n)$ will be the $d \times d$ block in position (i,j).

– if $\gamma_0, \gamma_n \in \mathbb{R}^d$, we will denote ${}_nA''(\gamma_0, \gamma_n)$ the hessian of the energy $A : (\mathbb{R}^d)^{[0,n],(\gamma_0,\gamma_n)} \longrightarrow \mathbb{R}$ at its minimizer (which, for simplicity, has been assumed unique in Assumption (A1)). If $\bar{\gamma}$ is energy-minimizing, then ${}_nA''(\gamma_0, \gamma_n) = {}_nA''(\bar{\gamma})$.

– We recall that $h_n(\gamma_0, \gamma_n)$ denotes the value of the minimum of the action on $(\mathbb{R}^d)^{[0,n],(\gamma_0,\gamma_n)}$ (we have performed a change of gage so that $h_n \geq 0$). If $\bar{\gamma} \in W$, we will denote $h_n(\bar{\gamma}) = h_n(\gamma_0, \gamma_n)$.

Lemma 3.2.1. *Let M be a symmetric matrix, decomposed in the form*

$$M = \begin{pmatrix} A & {}^tC \\ C & B \end{pmatrix}$$

(where A and B are square symmetric matrices, and C is a rectangular matrix of appropriate dimension).

If M is definite semi-positive, then

$$[M] \leq [A].[B]$$

Proof. Assume first that A is invertible. Since the determinant of a matrix is unchanged when adding to one line a linear combination of the others, we see that the determinant of M is equal to that of the matrix

$$\begin{pmatrix} A & {}^tC \\ 0 & B - CA^{-1}\,{}^tC \end{pmatrix}$$

Thus, $[M] = [A].[B - CA^{-1}\,{}^tC]$.

We now use the fact that an $m \times m$ matrix M is definite semi-positive if and only if, for all $J \subset \{1, ..., m\}$, the determinant of the square submatrix $M_J := (M_{i,j})_{i,j \in J}$ is nonnegative.

In particular, if M is definite semi-positive, so are A and B.

Denote by k the dimension of A, and l the dimension of B.

Let $J \subset \{k+1, ..., k+l\}$ and $I = J \cup \{1, ..., k\}$. Like previously,

$$[M_I] = [A].[(B - CA^{-1}\,{}^tC)_J].$$

It follows that all the determinants of $(B - CA^{-1}\,{}^tC)_J$ are nonnegative for all J; thus, $B - CA^{-1}\,{}^tC$ is definite semi-positive.

To conclude, note that A^{-1} is a definite positive symmetric matrix, so that

$$B - CA^{-1}\,{}^tC \leq B,$$

meaning that

$${}^tX.(B - CA^{-1}\,{}^tC).X \leq {}^tX.B.X$$

for all X. But, if B and $B - CA^{-1}\,{}^tC$ are positive semi-definite matrices such that $B - CA^{-1}\,{}^tC \leq B$, we must have $[B - CA^{-1}\,{}^tC] \leq [B]$ (this can be checked by using the fact that there exists a matrix P such that both ${}^tPBP$ and ${}^tP(B - CA^{-1}\,{}^tC)P$ are diagonal).

This ends the proof of the lemma when A is invertible. If A is not invertible, we know by the previous result that $[M + \varepsilon I] \leq [A + \varepsilon I].[B + \varepsilon I]$ for all $\varepsilon > 0$, and we conclude by letting ε tend to 0. □

Lemma 3.2.1 implies a property of subadditivity of $\log[{}_nA''(\bar{\gamma})]$:

Lemma 3.2.2. *If $\bar{\gamma} \in W$ is an energy-minimizing configuration, then, for all $m \leq n$,*

$$[{}_nA''(\bar{\gamma})] \leq [{}_mA''(\bar{\gamma})].[{}_{n-m}A''(\sigma^m\bar{\gamma})]$$

According to the subbaditive ergodic theorem ([Ki73]), this implies the existence of $\lim \frac{1}{n} \log[{}_nA''(\gamma)]$ in $\mathbb{R} \cup \{-\infty\}$, for μ-almost every γ, if μ is energy-minimizing. We shall say more about this limit in Section 4; in particular, prove that it is in $\mathbb{R}$.

We now turn to the proof of Theorem 0.0.1.

Proof. For simplicity, we write the proof in the case $d = 1$.

Let μ_∞ be a limit point of μ_β, $(\beta \to +\infty)$, and let μ be an arbitrary energy-minimizing measure on W; without loss of generality, we assume that μ is ergodic.

For $\varepsilon > 0$ and $M > 0$, consider the following (countable) partition of $\mathbb{R}^{2d} = \mathbb{R}^2$:

$$\mathbb{R}^2 = \sqcup_{i,j} \tilde{P}_{ij}$$

where the union runs over $\{(i,j) \in \mathbb{Z}^2, \mid j - i \mid < \frac{M}{\varepsilon}\} \cup \{(i, \infty), i \in \mathbb{Z}\}$, and the $\tilde{P}_{ij}$s are defined as follows:

$$\tilde{P}_{ij} = \{(\gamma_0, \gamma_1), \gamma_0 \in [i\varepsilon, (i+1)\varepsilon), \gamma_1 \in [j\varepsilon, (j+1)\varepsilon)\}$$

for $\mid j - i \mid < \frac{M}{\varepsilon}$,

$$\tilde{P}_{i\infty} = \{(\gamma_0, \gamma_1), \gamma_0 \in [i\varepsilon, (i+1)\varepsilon), \exists j, \mid j - i \mid \geq \frac{M}{\varepsilon}, \gamma_1 \in [j\varepsilon, (j+1)\varepsilon)\}$$

If ε is the inverse of an integer, this gives a finite partition of the quotient $W_1 \simeq \mathbb{R}^2/\mathbb{Z}$, and hence a finite partition of $W = \sqcup P_{ij}$:

$$P_{ij} = \{\bar{\gamma} \in W, (\gamma_0, \gamma_1) \in \tilde{P}_{ij}\}$$

The number M will be fixed later – sufficiently large, whereas ε is doomed to tend to 0.

We assume that μ and μ_∞ do not charge the boundary of the elements of the partition P – if not so, we may translate the initial partition to a new partition $(\tilde{P}_{ij} + x)_{ij}$, $x \in \mathbb{R}^d$, so that this assumption is satisfied. For $\delta > 0$, we will denote $\mu(\delta)$ the μ-measure of a δ-neighbourhood of the boundary of the partition P. The function $\mu(\delta)$ tends to zero as δ goes to zero.

The choice of the partition P induces a symbolic dynamics over a subshift in the finite alphabet $\{P_{ij}\}$:

$$W^P = \{(\alpha_k)_{k \in \mathbb{Z}} \subset \{(ij)\}^{\mathbb{Z}}, P_{\alpha_k} \cap \sigma^{-1} P_{\alpha_{k+1}} \neq \emptyset\}$$

If μ is a σ-invariant measure on W we will denote μ^P its image on W^P.

Recall the following convexity inequality:

$$-\sum p_i \log p_i + \sum p_i \log q_i \leq 0 \tag{3.2.1}$$

whenever (p_i) and (q_i) are probability weights.

Hence, for all n,

$$-\sum_\alpha \mu(P_{\alpha_0} \cap ... \cap \sigma^{-n+1} P_{\alpha_{n-1}}) \log \mu(P_{\alpha_0} \cap ... \cap \sigma^{-n+1} P_{\alpha_{n-1}})$$
$$+ \sum_\alpha \mu(P_{\alpha_0} \cap ... \cap \sigma^{-n+1} P_{\alpha_{n-1}}) \log \mu_\beta(P_{\alpha_0} \cap ... \cap \sigma^{-n+1} P_{\alpha_{n-1}}) \leq 0$$

the sums running over all word of length n in W^P.

From now on, we will replace the $\cap$ by dots . in expressions of the type $P_{\alpha_0} \cap ... \cap \sigma^{-n+1} P_{\alpha_{n-1}}$.

We can rewrite this:

$$(3.2.2)\qquad -\sum \mu(P_{\alpha_0}...\sigma^{-n-1}P_{\alpha_{n-1}})\log\mu(P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}})$$

$$+\sum \mu(P_{\alpha_0}..\sigma^{-n}P_{\alpha_n})\log\left(\frac{\beta}{2\pi}\right)^{\frac{n}{2}}\int_{P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}}\psi^*_\beta(\gamma_0)e^{-\beta\sum_{i=0}^{n-1}L(\gamma_i,\gamma_{i+1})}\psi_\beta(\gamma_n)d\gamma_0..d\gamma_n$$

$$\leq -\sum \mu_\beta(P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}})\log\mu_\beta(P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}})+$$

$$\sum \mu_\beta(P_{\alpha}..\sigma^{-n+1}P_{\alpha_{n-1}})\log\left(\frac{\beta}{2\pi}\right)^{\frac{n}{2}}\int_{P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}}\psi^*_\beta(\gamma_0)e^{-\beta\sum_{i=0}^{n-1}L(\gamma_i,\gamma_{i+1})}\psi_\beta(\gamma_n)d\gamma_0..d\gamma_n$$

We have denoted $d\gamma_0...d\gamma_n$ the Lebesgue measure on W_n; we may see it as the Lebesgue measure on the fundamental domain $[0,1)\times\mathbb{R}^{n-1}$ for the action of $\mathbb{Z}$ on $\mathbb{R}^n$, which we identify to W_n. In calculations, it will be convenient to keep this identification in mind.

The rest of the proof is organized as follows: by the Laplace method, we first find an upper bound for the right hand side of inequality (3.2.2), then state a couple of results about tridiagonal matrices, and finally, find a lower bound for the left hand side of inequality (3.2.2).

The conclusion of Theorem 0.0.1 is obtained by dividing the resulting inequality by n, and first letting n tend to ∞; then, β to ∞, and then ε to 0.

Upper bound. We begin with finding an upper bound for the right hand side of inequality (3.2.2), in terms of the determinants $[A'']$. An integer N is fixed and we take $n=kN$ in the inequality above.

Lemma 3.2.3. *(a)* **(Laplace method).** *Let $\gamma_0,\gamma_N\in\mathbb{R}^2$. Then, assuming as in (A1) that the minimizer of the energy in $(\mathbb{R}^d)^{[0,N],(\gamma_0,\gamma_N)}$ is unique and non-degenerate, we have*

$$\left(\frac{\beta}{2\pi}\right)^{\frac{N-1}{2}}\int_{\mathbb{R}^{N-1}}e^{-\beta\sum_{i=0}^{N-1}L(\gamma_i,\gamma_{i+1})}d\gamma_1...d\gamma_{N-1}=\frac{e^{-\beta h_N(\gamma_0,\gamma_N)}}{[\ {}_{N-1}A''(\gamma_0,\gamma_N)]^{1/2}}(1+\underset{\beta\to\infty}{o(1)})$$

$$\leq\frac{1}{[\ {}_{N-1}A''(\gamma_0,\gamma_N)]^{1/2}}(1+\underset{\beta\to\infty}{o(1)})$$

where, for fixed N, $\underset{\beta\to\infty}{o(1)}$ is uniform on each set $\{|\gamma_N-\gamma_0|\leq K\}$.

(b) If the constant M, involved in the construction of the partition P, is chosen large enough, then, for all $\gamma_0\in\mathbb{R}$,

$$\left(\frac{\beta}{2\pi}\right)^{\frac{N}{2}}\int_{\bar\gamma\in P_{\alpha_0}..\sigma^{-N+1}P_{\alpha_{N-1}}}e^{-\beta\sum_{i=0}^{N-1}L(\gamma_i,\gamma_{i+1})}d\gamma_1..d\gamma_N\leq\left(\frac{\beta}{2\pi}\right)^{\frac{N}{2}}e^{-\beta M}\leq 1$$

for β large enough, as soon as one the α_k's is of the form $i\infty$.

Assertion (a) comes from the Laplace method for estimating integrals decaying exponentially ([Di68], IV.2, or [Co65]). Since the method is very classical, we do not provide a proof; we shall provide one later, when we will need an estimate uniform in N. Assertion (a) requires the non-degeneracy of minima of the action, contained in Assumption (A1). The remainder term $o(1)$ is bounded in terms of the second and third derivatives with respect to $\gamma_1,...,\gamma_{n-1}$ of the energy $\sum_{i=0}^{N-1}L(\gamma_i,\gamma_{i+1})$, so that it is uniform on compact sets.

For (b), take M such that $\mid \gamma_1 - \gamma_0 \mid > M \Rightarrow L(\gamma_0, \gamma_1) \geq \mid \gamma_1 - \gamma_0 \mid$; and use the fact that $L \geq 0$ elsewhere.

We define functions F_N and G_N^β on W^P, depending on N coordinates, as follows:

$$F_N(\alpha_0, .., \alpha_{N-1}) = 1$$

if one of the α_j's is of the form $i\infty$, and

$$F_N(\alpha_0, .., \alpha_{N-1}) = \sup\{\frac{1}{[{}_{N-1}A''(\gamma_0, \gamma_N)]^{1/2}}, \bar{\gamma} \in P_{\alpha_0}..\sigma^{-N+1}P_{\alpha_{N-1}}\}$$

otherwise;

$$G_N^\beta(\alpha_0, .., \alpha_{N-1}) = 1$$

if one of the α_j's is of the form $i\infty$, and

$$G_N^\beta(\alpha_0, .., \alpha_{N-1}) = \left(\frac{\beta}{2\pi}\right)^{1/2} \sup_{\gamma_0} \int_{\mathbb{R}} e^{-\beta h_N(\gamma_0, \gamma_N)} d\gamma_N$$

otherwise.

Assumption (A3) ensures that G_N^β is bounded, independently of β, by B_N growing subexponentially with N.

Lemma 3.2.4. *If the constant M, involved in the construction of the partition P, is chosen large enough, then there exists $C(\beta) \geq 0$ and, for all $N \in \mathbb{N}^*$, a real $\beta(N) > 0$, such that: for all k, and for all $\alpha_0, ..., \alpha_{kN-1}$,*

$$\left(\frac{\beta}{2\pi}\right)^{\frac{kN}{2}} \int_{P_{\alpha_0}..\sigma^{-kN+1}P_{\alpha_{kN-1}}} \psi_\beta^*(\gamma_0) e^{-\beta \sum_{i=0}^{kN-1} L(\gamma_i, \gamma_{i+1})} \psi_\beta(\gamma_{kN}) d\gamma_0 ... d\gamma_{kN}$$
$$\leq C(\beta) \prod_{j=0}^{k-1} F_N(\alpha_{jN}, .., \alpha_{(j+1)N-1}) \prod_{l=0}^{k-1} G_N(\alpha_{jN}, .., \alpha_{(j+1)N-1})(1 + \underset{\beta\to\infty}{o(1)})^k$$

for all $\beta > \beta(N)$, and with a uniform $\underset{\beta\to\infty}{o(1)}$.

Proof. We first note that there exists $C(\beta) > 0$ such that $C(\beta)^{-1/2} \leq \psi_\beta \leq C(\beta)^{1/2}$, and $C(\beta)^{-1/2} \leq \psi_\beta^* \leq C(\beta)^{1/2}$, because they are continuous positive $\mathbb{Z}^d$-periodic functions.

Applying Fubini's theorem, we first estimate the integral with respect to $\gamma_{(k-1)N+1}, ..., \gamma_{kN}$, while $\gamma_0, ..., \gamma_{(k-1)N}$ are fixed.

If one of the P_{α_j}'s ($j = (k-1)N, ..., kN-1$) is of the form $P_{i\infty}$, we use Lemma 3.2.3 (b), and get

$$\left(\frac{\beta}{2\pi}\right)^{\frac{N}{2}} \int_{(\gamma_{(k-1)N}, ..., \gamma_{kN}) \in P_{\alpha_{(k-1)N}}..\sigma^{-N+1}P_{\alpha_{kN-1}}} e^{-\beta \sum_{i=(k-1)N}^{kN-1} L(\gamma_i, \gamma_{i+1})} d\gamma_{(k-1)N+1}..d\gamma_{kN}$$
$$\leq 1 = F_N(\alpha_{(k-1)N}, .., \alpha_{kN-1}) G_N^\beta(\alpha_{(k-1)N}, .., \alpha_{kN-1})$$

Otherwise, we use Lemma 3.2.3 (a), and write

(3.2.3)
$$\left(\frac{\beta}{2\pi}\right)^{\frac{N}{2}} \int_{(\gamma_{(k-1)N},..,\gamma_{kN})\in P_{\alpha_{(k-1)N}}..\sigma^{-N+1}P_{\alpha_{kN-1}}} e^{-\beta\sum_{i=(k-1)N}^{kN-1}L(\gamma_i,\gamma_{i+1})}d\gamma_{(k-1)N+1}..d\gamma_{kN}$$
$$\leq F_N(\alpha_{(k-1)N},..,\alpha_{kN-1})(1+o(1))\left(\frac{\beta}{2\pi}\right)^{\frac{1}{2}}\int e^{-\beta h_N(\gamma_{(k-1)N},\gamma_{kN})}d\gamma_{kN}$$
$$\leq F_N(\alpha_{(k-1)N},..,\alpha_{kN-1})(1+o(1))G_N^{\beta}(\alpha_{(k-1)N},..,\alpha_{kN-1})$$

The latter bound does not depend on $\gamma_{(k-1)N}$; hence,

$$\left(\frac{\beta}{2\pi}\right)^{\frac{kN}{2}} \int_{P_{\alpha_0}..\sigma^{-kN+1}P_{\alpha_{kN-1}}} e^{-\beta\sum_{i=0}^{kN-1}L(\gamma_i,\gamma_{i+1})}d\gamma_0...d\gamma_{kN}$$
$$\leq \left(\frac{\beta}{2\pi}\right)^{\frac{(k-1)N}{2}} \int_{P_{\alpha_0}..\sigma^{-(k-1)N+1}P_{\alpha_{(k-1)N-1}}} e^{-\beta\sum_{i=0}^{(k-1)N-1}L(\gamma_i,\gamma_{i+1})}d\gamma_0...d\gamma_{(k-1)N}$$
$$\times F_N(\alpha_{(k-1)N},..,\alpha_{kN-1})G_N^{\beta}(\alpha_{(k-1)N},..,\alpha_{kN-1})(1+\underset{\beta\to\infty}{o(1)})$$

Lemma 3.2.4 can now be proved by induction on k.

□

About tridiagonal matrices. Before going on estimating integrals, we need a few facts about tridiagonal matrices. We call a matrix $(A_{ij})_{1\leq i,j\leq n}$ *tridiagonal* if
$$A_{ij}\neq 0 \Rightarrow |\, i-j \,|\leq 1$$
The hessian of the energy is a tridiagonal matrix.

The following lemma is essentially proved in [AMB92] (p.128):

Lemma 3.2.5. *For all $\alpha>0$, there exists $r(\alpha)>0$ such that: if A is an invertible symmetric tridiagonal matrix with $|\, A_{i,i+1} \,|\leq 1$, then*
$$\|\, A^{-1} \,\|_2 \leq \alpha \textit{ implies } \|\, A^{-1} \,\|_\infty \leq r(\alpha)$$
independently of the dimension.

Proof. For $1\leq j\leq n$, let $f^j=A^{-1}e^j$, where (e^j) is the canonical basis of $\mathbb{R}^n$. Note that
$$\|\, A^{-1} \,\|_\infty \leq \sup_k \sum_j |\, f_k^j \,| = \sup_j \sum_k |\, f_k^j \,|$$
since A^{-1} is symmetric.

Let us fix j, and denote $f=f^j$. For $m>j$, we define a vector ${}^mf$ with coordinates
$${}^mf_k=0$$
for $k<m$, and
$${}^mf_k=f_k$$
for $k\geq m$. Then $\eta=A.{}^mf$ has coordinates
$$\eta_{m-1}=A_{m-1,m}f_m$$
$$\eta_m=-A_{m,m-1}f_{m-1}$$

and

$$\eta_k = 0$$

otherwise.

Since, by assumption,

$$\| \ {}^m f \|_2 \leq \alpha \| \eta \|_2,$$

we get, for all $m > j$,

$$P_m := \sum_{k \geq m} | f_k |^2 \leq \alpha^2 (| f_m |^2 + | f_{m-1} |^2)$$

As proved in [AMB92], p. 128, this inequality implies

$$| f_k | \leq \left(\frac{2\alpha^2}{1 + (1 + 4\alpha^4)^{1/2}} \right)^{\frac{k-j-2}{2}} ||f||_2 \leq \left(\frac{2\alpha^2}{1 + (1 + 4\alpha^4)^{1/2}} \right)^{\frac{k-j-2}{2}} \sqrt{\alpha}$$

for $k \geq j$. Remembering that f stands for f^j:

$$\sum_{k \geq j} | f_k^j | \leq \sqrt{\alpha} \sum_{k=0}^{+\infty} \left(\frac{2\alpha^2}{1 + (1 + 4\alpha^4)^{1/2}} \right)^{(k-2)/2} =: r(\alpha)/2$$

We can use a similar trick for $k < j$, and get that

$$\sum_{1 \leq k \leq n} | f_k^j | \leq r(\alpha),$$

independently of j and of the dimension n. □

We shall also need the following result, which is a part of the main result of [AMB92] (Theorem 2):

Theorem 3.2.6. *([AMB92]) Let M be a symmetric tridiagonal $\mathbb{Z} \times \mathbb{Z}$ matrix, such that there exists $K > 0$ such that, for all i, $K^{-1} \leq |M_{i,i+1}| \leq K$ and $|M_{i,i}| \leq K|$. Assume that M defines a continuous, invertible endomorphism of $l^2(\mathbb{Z}, \mathbb{R})$.*

Then, the kernel of M in $\mathbb{R}^{\mathbb{Z}}$ is 2-dimensional and admits a basis $s, u \in \mathbb{R}^{\mathbb{Z}}$ such that

$$|s_{n+m}| \leq C\theta^m ||(s_n, s_{n+1})||$$
$$|u_{n-m}| \leq C\theta^m ||(u_n, u_{n+1})||$$

for all $n \in \mathbb{Z}$, $m \geq 0$, for some constants $C > 0$, $0 < \theta < 1$.

Theorem 3.2.6 implies the existence of a real number $L > 0$ such that, for all $\delta > 0$, for all $0 < n$, for all $\gamma \in \mathbb{R}^{\mathbb{Z}}$ such that $M\gamma = 0$,

$$|\gamma_0| \leq \delta \text{ and } |\gamma_n| \leq \delta \Rightarrow |\gamma_j| \leq L\delta, \text{ for all } j = 1, ..., n-1.$$

Indeed, fix $j = 0, ..., n-1$; there exist $\lambda_1, \lambda_2 \in R$ such that

$$(\gamma_k, \gamma_{k+1}) = \lambda_1 \frac{(s_k, s_{k+1})}{||(s_j, s_{j+1})||} + \lambda_2 \frac{(u_k, u_{k+1})}{||(u_j, u_{j+1})||},$$

for all k. Writing that $|\gamma_0| \leq \delta$ and $|\gamma_n| \leq \delta$, we obtain

$$C|\lambda_1|\theta^{-j} - C|\lambda_2|\theta^j \leq \delta$$

and

$$C|\lambda_2|\theta^{-(n-j)} - C|\lambda_1|\theta^{n-j} \leq \delta,$$

which implies

$$C|\lambda_1|(1 - \theta^{2n}) \leq 2\delta\theta^j$$

and

$$C|\lambda_2|(1-\theta^{2n}) \leq 2\delta\theta^{n-j}.$$

Hence,

$$||(\gamma_j, \gamma_{j+1})|| \leq |\lambda_1| + |\lambda_2| \leq 4\delta/C.$$

Lower bound. Let us turn to the left hand side of inequality (3.2.2), which we will try to bound below before letting $n = kN$ tend to ∞. Since μ is a minimizing measure, we note that the term $\mu(P_{\alpha_0}....\sigma^{-n+1}P_{\alpha_{n-1}})$ is non zero only if all the P_{α_i}'s are included in $\{| \gamma_1 - \gamma_0 | \leq M\}$ (if M is large enough); besides, the cylinder $P_{\alpha_0}....\sigma^{-n+1}P_{\alpha_{n-1}}$ must contain a trajectory, say $\bar{\xi}$, in the Mather set.

In the coming calculations, the cylinder $P_{\alpha_0}....\sigma^{-n+1}P_{\alpha_{n-1}}$ is fixed, as well as $\bar{\xi} \in P_{\alpha_0}....\sigma^{-n+1}P_{\alpha_{n-1}} \cap \mathcal{M}$. We need to estimate from below the integral

$$\int_{P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}} \psi^*_\beta(\gamma_0) e^{-\beta\sum_{i=0}^{n-1} L(\gamma_i,\gamma_{i+1})} \psi_\beta(\gamma_n) d\gamma_0...d\gamma_n.$$

As previously, we shall use the Laplace method. However, since we need a uniform estimate with respect to the length n of the path, we shall now give the details.

Before starting, recall Assumptions (A2):

There exists $\varepsilon_0 > 0$ such that, for all $0 \leq \delta \leq \varepsilon_0$, there exists a sequence $(c_n) \in [0,1]^{\mathbb{N}}$ satisfying:

– $\lim_n \frac{\log c_n}{n} = 0$,

and :

– for all $n > 0$, for all $\gamma_0, \gamma_n \in \mathbb{R}^d$ such that

$$\| \gamma_0 - \xi_0 \| \leq c_n \delta$$

$$\| \gamma_n - \xi_n \| \leq c_n \delta$$

for some energy-minimizing configuration ξ, there exists a minimizer γ of the energy:

$$\begin{aligned} A : (\mathbb{R}^d)^{[0,n],\gamma_0,\gamma_n} &\longrightarrow \mathbb{R} \\ (\gamma_0, ..., \gamma_n) &\mapsto \sum_{i=0}^{n-1} L(\gamma_i, \gamma_{i+1}), \end{aligned}$$

such that $\| \gamma_k - \xi_k \| \leq \delta$ for all $0 \leq k \leq n$.

We denote $(\gamma_0, \hat{\gamma}_1, ..., \hat{\gamma}_{n-1}, \gamma_n)$ the minimizer. Its energy is, by definition of the function h_n,

$$h_n(\gamma_0, \gamma_n) = L(\gamma_0, \hat{\gamma}_1) + \sum_{i=1}^{n-2} L(\hat{\gamma}_i, \hat{\gamma}_{i+1}) + L(\hat{\gamma}_{n-1}, \gamma_n).$$

Applying a Taylor formula to the function $L(\gamma_0,\gamma_1)+\sum_{i=1}^{n-2} L(\gamma_i,\gamma_{i+1})+L(\gamma_{n-1},\gamma_n)$ at the minimizer $(\hat{\gamma}_1, \hat{\gamma}_2, \cdots, \hat{\gamma}_{n-1})$, we can write

$$\int_{P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}} \psi^*_\beta(\gamma_0) e^{-\beta\sum_{i=0}^{n-1} L(\gamma_i,\gamma_{i+1})} \psi_\beta(\gamma_n) d\gamma_0...d\gamma_n$$

$$= \int_{P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}} \psi^*_\beta(\gamma_0) e^{-\beta h_n(\gamma_0,\gamma_n) - \frac{\beta}{2}\ {}_{n-1}A''(\gamma_0,\gamma_n).(\gamma-\hat{\gamma})^2 - \beta R_n(\gamma-\hat{\gamma})} \psi_\beta(\gamma_n) d\gamma_0..d\gamma_n$$

where the remainder R_n is given by the integral formula:

$$R_n(\gamma-\hat{\gamma}) = \int_0^1 \frac{(1-t)^2}{2} A^{(3)}(\hat{\gamma}+t(\gamma-\hat{\gamma})).((\gamma-\hat{\gamma}))^3 dt$$

so that

$$|R_n(\gamma-\hat{\gamma})| \leq C \parallel \gamma-\hat{\gamma} \parallel_3^3 \leq C \parallel \gamma-\hat{\gamma} \parallel_\infty \parallel \gamma-\hat{\gamma} \parallel_2^2 \leq C\varepsilon \parallel \gamma-\hat{\gamma} \parallel_2^2$$

where C is a bound onthe third derivative of L on the set $\{(x,y)\in\mathbb{R}^2, |x-y|\leq M\}$, and ε is the diameter of the elements of the partition P.

Moreover, if the cylinder $P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}$ contains a configuration $\bar{\xi}$ in the Mather set, then, for all $\bar{\gamma}\in P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}$,

$$\begin{aligned} 0 \leq h_n(\gamma_0,\gamma_n) &\leq L(\gamma_0,\xi_1)+L(\xi_1,\xi_2)+...+L(\xi_{n-1},\gamma_n) \\ &\leq L(\xi_0,\xi_1)+L(\xi_1,\xi_2)+...+L(\xi_{n-1},\xi_n)+C\varepsilon = C\varepsilon, \end{aligned}$$

where C is a Lipschitz constant for L on $\{(x,y)\in\mathbb{R}^2, |x-y|\leq M\}$.

Thus,

$$\begin{aligned} &\int_{P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}} \psi^*_\beta(\gamma_0)e^{-\beta\sum_{i=0}^{n-1}L(\gamma_i,\gamma_{i+1})}\psi_\beta(\gamma_n)d\gamma_0...d\gamma_n \\ &\geq C(\beta)^{-1}e^{-\beta C\varepsilon}\int_{P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}} e^{-\beta(\frac{1}{2}\ _{n-1}A''(\gamma_0,\gamma_n)+C\varepsilon I_{n-1}).(\gamma_1-\hat{\gamma}_1,..,\gamma_{n-1}-\hat{\gamma}_{n-1})^2}d\gamma_0...d\gamma_n \end{aligned}$$

We have denote I_{n-1} the identity matrix of dimension $n-1$ and, as previously, $C(\beta)^{-1/2}$ is a lower bound for both ψ_β and ψ^*_β.

If we were sure that $P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}$ contained a neighbourhood of $(\gamma_0,\hat{\gamma}_1,...,\hat{\gamma}_{n-1},\gamma_n)$, for every γ_0,γ_n, our job would be quite easier and we could go directly to the estimate (3.2.10) a couple of pages ahead. However, this is not necessarily the case: the problem occurs when (ξ_j,ξ_{j+1}) comes too close to the boundary of the partition. The technical complications of the next few pages arise from the necessity of dealing with this problem.

To begin with, we can write a very rough estimate:

$$\begin{aligned} (3.2.4)\quad & C(\beta)^{-1}e^{-\beta C\varepsilon}\left(\frac{\beta}{2\pi}\right)^{\frac{n}{2}}\int_{P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}} d\gamma_0 d\gamma_n \times \\ &\int e^{-\beta(\frac{1}{2}\ _{n-1}A''(\gamma_0,\gamma_n)+C\varepsilon I_{n-1}).(\gamma_1-\hat{\gamma}_1,\cdots,\gamma_{n-1}-\hat{\gamma}_{n-1})^2}d\gamma_1..d\gamma_{n-1} \\ &\geq C(\beta)^{-1}e^{-\beta C\varepsilon}\left(\frac{\beta}{2\pi}\right)^{\frac{n}{2}}\varepsilon^n e^{-nK\beta\varepsilon^2} \end{aligned}$$

where K is an upper bound on the norm of $\frac{1}{2}A''+C\varepsilon I$ in $l^2(\mathbb{Z},\mathbb{R})$.

Let us now try and give a more subtle estimate: we write

$$C(\beta)^{-1}e^{-\beta C\varepsilon}\int_{P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}} d\gamma_0 d\gamma_n \times$$
$$\int e^{-\beta(\frac{1}{2}\ _{n-1}A''(\gamma_0,\gamma_n)+C\varepsilon I_{n-1}).(\gamma_1-\hat{\gamma}_1,\cdots,\gamma_{n-1}-\hat{\gamma}_{n-1})^2} d\gamma_1..d\gamma_{n-1}$$
$$\geq C(\beta)^{-1}e^{-\beta C\varepsilon}\int_{(1)} d\gamma_0 d\gamma_n \times$$
$$\int e^{-\beta(\frac{1}{2}\ _{n-1}A''(\gamma_0,\gamma_n)+C\varepsilon I_{n-1}).(\gamma_1-\hat{\gamma}_1,\cdots,\gamma_{n-1}-\hat{\gamma}_{n-1})^2} d\gamma_1..d\gamma_{n-1}$$

where the integral runs over the set

$$(1) = P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}} \cap \{|\gamma_0-\xi_0| \leq c_n\delta, |\gamma_n-\xi_n| \leq c_n\delta\},$$

for some $\delta > 0$. We choose

$$\delta^2 = B\beta^{-1},$$

where $B > 0$ is arbitrary.

At this stage, it is useful to remember that (thanks to our definition of the partition P) the cylinder $P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}} \subset [0,1)\times\mathbb{R}^n$ is the product of its projections $B_0, ..., B_n$ on the successive coordinates:

$$P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}} = B_0 \times ... \times B_n.$$

Moreover, the B_i's are segments of length ε: $B_i = [a_i, b_i)$.

We denote

$$J(\xi) = \{i \in [0,n] \text{ such that } \xi_i \notin (a_i + 2\delta, b_i - 2\delta)\}.$$

One has

$$C(\beta)^{-1}e^{-\beta C\varepsilon}\int_{(1)} d\gamma_0 d\gamma_n \times \tag{3.2.5}$$
$$\int e^{-\beta(\frac{1}{2}\ _{n-1}A''(\gamma_0,\gamma_n)+C\varepsilon I_{n-1}).(\gamma_1-\hat{\gamma}_1,\cdots,\gamma_{n-1}-\hat{\gamma}_{n-1})^2} d\gamma_1..d\gamma_{n-1}$$
$$\geq C(\beta)^{-1}e^{-\beta C\varepsilon}\int_{(1)\cap(2)} d\gamma_0 d\gamma_n \times$$
$$\int e^{-\beta(\frac{1}{2}\ _{n-1}A''(\gamma_0,\gamma_n)+C\varepsilon I_{n-1}).(\gamma_1-\hat{\gamma}_1,\cdots,\gamma_{n-1}-\hat{\gamma}_{n-1})^2} d\gamma_1..d\gamma_{n-1}$$

where

$$(2) = \{\gamma_i \in B_i \text{ for } i \in J(\xi), \gamma_j \in B_j \text{ and } |\gamma_j - \hat{\gamma}_j| \leq \delta \text{ for } j \notin J(\xi)\}.$$

By Assumption (A2), if $\gamma \in (1)$, then, for $j \notin J(\xi)$,

$$|\gamma_j - \hat{\gamma}_j| \leq \delta \Longrightarrow \gamma_j \in B_j,$$

so that actually

$$(2) = \{\gamma_i \in B_i \text{ for } i \in J(\xi), |\gamma_j - \hat{\gamma}_j| \leq \delta \text{ for } j \notin J(\xi)\}.$$

We claim that

$$(3.2.6)\quad C(\beta)^{-1}e^{-\beta C\varepsilon}\int_{(1)\cap(2)} d\gamma_0 d\gamma_n \times$$
$$\int e^{-\beta(\frac{1}{2}\ _{n-1}A''(\gamma_0,\gamma_n)+C\varepsilon I_{n-1}).(\gamma_1-\hat{\gamma}_1,\cdots,\gamma_{n-1}-\hat{\gamma}_{n-1})^2} d\gamma_1..d\gamma_{n-1}$$
$$\geq C(\beta)^{-1}e^{-\beta C\varepsilon}\left((2\pi)^{-1/2}\sqrt{2K^{-1}B}(L+1)e^{-4K^{-1}B(L+1)^2}\right)^{|J(\xi)|}\int_{(3)} d\gamma_0 d\gamma_n \times$$
$$\int e^{-\beta(\frac{1}{2}\ _{n-1}A''(\gamma_0,\gamma_n)+C\varepsilon I_{n-1}).(\gamma_1-\hat{\gamma}_1,\cdots,\gamma_{n-1}-\hat{\gamma}_{n-1})^2} d\gamma_1..d\gamma_{n-1}$$

where:
– $B = \beta\delta^2$
– K is an upper bound on the norms of $(\frac{1}{2}A'' + C\varepsilon I)$ and $(\frac{1}{2}A'' + C\varepsilon I)^{-1}$ in $l^2(\mathbb{Z},\mathbb{R})$
– L is a real number such that, for all $\delta > 0$, for all $\gamma \in \mathbb{R}^{\mathbb{Z}}$, for all $n > 0$,

$$(\frac{1}{2}A'' + C\varepsilon I).\gamma = 0, |\gamma_0| \leq \delta, |\gamma_n| \leq \delta \Rightarrow |\gamma_j| \leq L\delta \text{ for all } j = 1, ..., n-1.$$

The existence of L is ensured by the remarks following Theorem 3.2.6 (note that L depends on ε).

Finally, (3) is the set

$$(3) = \{|\gamma_0 - \xi_0| \leq c_n\delta, |\gamma_n - \xi_n| \leq c_n\delta, |\gamma_j - \hat{\gamma}_j| \leq \delta \text{ for } j \notin J(\xi)\}.$$

To prove (3.2.6), write $J(\xi)$ as a disjoint union of intervals:

$$J(\xi) = [k_1, l_1] \cup [k_2, l_2] \cup ... \cup [k_r, l_r].$$

Integrate $\int_{(1)\cap(2)}$ with respect to the variables $\gamma_{k_1}, ..., \gamma_{l_1}$, the other variables fixed. Since $|\gamma_{k_1-1} - \hat{\gamma}_{k_1-1}| \leq \delta$ and $|\gamma_{l_1+1} - \hat{\gamma}_{l_1+1}| \leq \delta$, we know that the critical point of the function

$$(\gamma_{k_1}, ..., \gamma_{l_1}) \mapsto (\frac{1}{2}\ _{n-1}A''(\gamma_0,\gamma_n) + C\varepsilon I_{n-1}).(\gamma_1 - \hat{\gamma}_1, \cdots, \gamma_{n-1} - \hat{\gamma}_{n-1})^2$$

is at uniform distance $L\delta$ from $(\hat{\gamma}_{k_1}, ..., \hat{\gamma}_{l_1})$. From there, we shall prove that

$$(3.2.7)\quad \int_{\gamma\in(1)\cap(2)} e^{-\beta(\frac{1}{2}\ _{n-1}A''(\gamma_0,\gamma_n)+C\varepsilon I_{n-1}).(\gamma_1-\hat{\gamma}_1,\cdots,\gamma_{n-1}-\hat{\gamma}_{n-1})^2} d\gamma_{k_1}..d\gamma_{l_1}$$
$$\geq \left((2\pi)^{-1/2}\sqrt{2K^{-1}B}(L+1)e^{-4K^{-1}B(L+1)^2}\right)^{l_1-k_1+1}$$
$$\int_{\gamma\in(3)} e^{-\beta(\frac{1}{2}\ _{n-1}A''(\gamma_0,\gamma_n)+C\varepsilon I_{n-1}).(\gamma_1-\hat{\gamma}_1,\cdots,\gamma_{n-1}-\hat{\gamma}_{n-1})^2} d\gamma_{k_1}..d\gamma_{l_1}$$

All we can say about the domain of integration of (3.2.7) is that: if $\gamma \in (1)$, then for $j \in J(\xi)$ the domain of integration with respect to γ_j contains either $\{\gamma_j - \hat{\gamma}_j \in [\delta, \varepsilon]\}$, or $\{\gamma_j - \hat{\gamma}_j \in [-\varepsilon, -\delta]\}$. Consider, for instance, the first situation.

Integrate (3.2.7) with respect to k_1. Say, for instance, $k_1 = 2$, and denote $M = \frac{1}{2}\ {}_3A''(\gamma_0, \gamma_n) + C\varepsilon I_3$. Then, estimate (3.2.7) goes as follows:

$$\int_{\gamma_2=\delta}^{\varepsilon} e^{-\beta(2M_{12}\gamma_1\gamma_2+M_{22}\gamma_2^2+2M_{23}\gamma_2\gamma_3)}d\gamma_2$$
$$= e^{-\beta(2M_{12}\gamma_1\gamma_2(\min)+M_{22}\gamma_2(\min)^2+2M_{23}\gamma_2(\min)\gamma_3)} \int_{\gamma_2=\delta}^{\varepsilon} e^{-\beta M_{22}(\gamma_2-\gamma_2(\min))^2}d\gamma_2$$

where $\gamma_2(\min)$ is the point achieving the minimum of $2M_{12}\gamma_1\gamma_2+M_{22}\gamma_2^2+2M_{23}\gamma_2\gamma_3$. We know that $|\gamma_2(\min)| \leq L\delta$, so that, if we perform the change of variable $\gamma_2 \mapsto \sqrt{2\beta M_{22}}(\gamma_2 - \gamma_2(\min))$, and if $K^{-1} \leq M_{22} \leq K$, we have

$$e^{-\beta(2M_{12}\gamma_1\gamma_2(\min)+M_{22}\gamma_2(\min)^2+2M_{23}\gamma_2(\min)\gamma_3)} \int_{\gamma_2=\delta}^{\varepsilon} e^{-\beta M_{22}(\gamma_2-\gamma_2(\min))^2}d\gamma_2 \tag{3.2.8}$$
$$\geq e^{-\beta(2M_{12}\gamma_1\gamma_2(\min)+M_{22}\gamma_2(\max)^2+2M_{23}\gamma_2(\max)\gamma_3)}(2\beta M_{22})^{-1/2} \int_{\sqrt{2\beta K^{-1}}\delta(L+1)}^{\sqrt{2K\beta}(\varepsilon-L\delta)} e^{-\gamma_2^2/2}d\gamma_2$$
$$= e^{-\beta(2M_{12}\gamma_1\gamma_2(\min)+M_{22}\gamma_2(\min)^2+2M_{23}\gamma_2(\min)\gamma_3)}(2\beta M_{22})^{-1/2} \int_{\sqrt{2K^{-1}B}(L+1)}^{\sqrt{2K\beta}(\varepsilon-L\delta)} e^{-\gamma_2^2/2}d\gamma_2$$

if we remember that β and δ are linked by $\beta\delta^2 = B$. Now, for β large enough, we can bound the last integral from below by the integral on the interval $[\sqrt{2K^{-1}B}(L+1), 2\sqrt{2K^{-1}B}(L+1)]$, which is itself larger than

$$e^{-\beta(2M_{12}\gamma_1\gamma_2(\min)+M_{22}\gamma_2(\min)^2+2M_{23}\gamma_2(\min)\gamma_3)}(2\beta M_{22})^{-1/2}\sqrt{2K^{-1}B}(L+1)e^{-4K^{-1}B(L+1)^2} \tag{3.2.9}$$
$$= (2\pi)^{-1/2}\int_{\mathbb{R}} e^{-\beta(2M_{12}\gamma_1\gamma_2+M_{22}\gamma_2^2+2M_{23}\gamma_2\gamma_3)}d\gamma_2.\sqrt{2K^{-1}B}(L+1)e^{-4K^{-1}B(L+1)^2}$$

We have integrated (3.2.7) with respect to γ_{k_1}. We can iterate the procedure and integrate successively with respect to $\gamma_{k_1+1}, ..., \gamma_{l_1}$, to prove (3.2.7). From (3.2.7), (3.2.6) is obtained by combining the similar estimates for all the intervals $[k_j, l_j]$.

It remains to estimate the integral $\int_{(3)}$ (cf (3.2.6)). The integral $\int_{(3)}$ runs over $\gamma_j \in \mathbb{R}$, for all $j \in J(\xi)$. For an index $i \notin J(\xi)$, it still runs over the set $\{|\gamma_i - \hat{\gamma}_i| \leq \delta\}$.

For a break, we prove the following corollary of Lemma 3.2.5:

Corollary 3.2.7. *There exists $\rho(\varepsilon)$ such that, for all n, for all $\gamma \in W$,*

$$\| \left({}_{n-1}A''(\gamma) + 2C\varepsilon I_{n-1}\right)^{-1/2} \|_\infty \leq \frac{1}{\rho(\varepsilon)}$$

Proof. Obviously, the spectrum of ${}_{n-1}A''(\gamma) + 2C\varepsilon I_{n-1}$ is included in an interval $[2C\varepsilon, \lambda]$ independent of the dimension n. Let $\mathcal{C}$ be a closed contour in $\mathbb{C} \setminus \mathbb{R}^-$, going once around $[2C\varepsilon, \lambda]$. The matrix $\left({}_{n-1}A''(\gamma) + 2C\varepsilon I_{n-1}\right)^{-1/2}$ is given by holomorphic functional calculus:

$$({}_{n-1}A''(\gamma) + 2C\varepsilon I_{n-1})^{-1/2} = \frac{1}{2i\pi}\int_{\mathcal{C}} z^{-1/2}(zI_{n-1} - ({}_{n-1}A''(\gamma) + 2C\varepsilon I_{n-1}))^{-1}dz$$

Now, for all $z \in \mathcal{C}$,

$$\| \left(zI_{n-1} - ({}_{n-1}A''(\gamma) + 2C\varepsilon I_{n-1})\right)^{-1} \|_2$$

is bounded, independently of n, by

$$\alpha(z) = \sup_{x \in [2C\varepsilon, \lambda]} \frac{1}{\mid z - x \mid}$$

By Lemma 3.2.5,

$$\| \left(zI_{n-1} - ({}_{n-1}A''(\gamma) + 2C\varepsilon I_{n-1}) \right)^{-1} \|_\infty \leq r(\alpha(z))$$

independently of n, and

$$\| \left({}_{n-1}A''(\gamma) + 2C\varepsilon I_{n-1} \right)^{-1/2} \|_\infty \leq \frac{1}{2\pi} \int_{\mathcal{C}} \mid z \mid^{-1/2} r(\alpha(z)) dz := \frac{1}{\rho(\varepsilon)}$$

$\square$

We resume the calculations from (3.2.6):

(3.2.10)

$$\int_{(3)} d\gamma_0 d\gamma_n . \int e^{-\beta(\frac{1}{2}\ {}_{n-1}A''(\gamma_0,\gamma_n) + C\varepsilon I_{n-1}).(\gamma_1 - \hat{\gamma}_1, \cdots, \gamma_{n-1} - \hat{\gamma}_{n-1})^2} d\gamma_1 .. d\gamma_{n-1}$$

$$\geq \int_{|\gamma_0 - \xi_0| \leq c_n\delta, |\gamma_n - \xi_n| \leq c_n \delta} d\gamma_0 d\gamma_n \times$$

$$\int_{\|(\gamma - \hat{\gamma})\|_\infty \leq \delta} e^{-\beta(\frac{1}{2}\ {}_{n-1}A''(\gamma_0,\gamma_n) + C\varepsilon I_{n-1}).(\gamma_1 - \hat{\gamma}_1, \cdots, \gamma_{n-1} - \hat{\gamma}_{n-1})^2} d\gamma_1 .. d\gamma_{n-1}$$

$$\geq \int_{|\gamma_0 - \xi_0| \leq c_n\delta, |\gamma_n - \xi_n| \leq c_n \delta} d\gamma_0 d\gamma_n \times$$

$$\int_{\|(\frac{ {}_{n-1}A''(\gamma_0,\gamma_n)}{2} + C\varepsilon I_n)^{1/2}.(\gamma - \hat{\gamma})\|_\infty \leq \rho(\varepsilon)\delta} e^{-\beta(\frac{1}{2}\ {}_{n-1}A''(\gamma_0,\gamma_n) + C\varepsilon I_{n-1}).(\gamma_1 - \hat{\gamma}_1, \cdots, \gamma_{n-1} - \hat{\gamma}_{n-1})^2} d\gamma_1 .. d\gamma_{n-1}$$

$$= \left(\frac{\beta}{2\pi} \right)^{-(n-1)/2} \int_{|\gamma_0 - \xi_0| \leq c_n\delta, |\gamma_n - \xi_n| \leq c_n \delta} d\gamma_0 d\gamma_n \times$$

$$\frac{1}{[\ {}_{n-1}A''(\gamma_0, \gamma_n) + 2C\varepsilon I_{n-1}]^{1/2}} \times \frac{1}{(2\pi)^{(n-1)/2}} \int_{\|y\|_\infty \leq \sqrt{\beta}\rho(\varepsilon)\delta} e^{-\frac{(y,y)}{2}} dy_1 .. dy_{n-1}$$

$$\geq (c_n\delta)^2 \left(\frac{\beta}{2\pi} \right)^{-(n-1)/2} \frac{1}{\max_{\gamma \in \alpha} [\ {}_{n-1}A''(\gamma) + 2C\varepsilon I_{n-1}]^{1/2}} (1 - e^{-\beta\rho(\varepsilon)^2 \frac{\delta^2}{2}})^{n-1}$$

The max in the last line is to be read as the max over all the $\gamma \in P_{\alpha_0} .. \sigma^{-n+1} P_{\alpha_{n-1}}$.

To get the last inequality, we have used the following estimate on tails of the Gaussian distribution on $\mathbb{R}$:

$$\frac{1}{(2\pi)^{1/2}} \int_{|y| \geq Y} e^{-\frac{|y|^2}{2}} dy \leq 2 \frac{e^{-\frac{|Y|^2}{2}}}{Y}$$

which yields, in dimension $n - 1$,

$$\text{(3.2.11)} \qquad \frac{1}{(2\pi)^{(n-1)/2}} \int_{\|y\|_\infty \leq Y} e^{-\frac{(y,y)^2}{2}} dy_1 .. dy_{n-1} \geq (1 - e^{-\frac{Y^2}{2}})^{n-1}$$

for $Y > 2$. We apply it to $Y = \sqrt{\beta}\rho(\varepsilon)\delta = \rho(\varepsilon)\sqrt{B}$; we take B large enough to ensure that $Y > 2$.

The main point in estimate (3.2.10) is summarized in the following lemma:

Lemma 3.2.8. *There exists $\rho = \rho(\varepsilon)$ such that, for all γ_0, γ_n,*

$$\int_{\|(\gamma-\hat{\gamma})\|_\infty \leq \delta} e^{-\beta(\frac{1}{2}\ {}_{n-1}A''(\gamma_0,\gamma_n)+C\varepsilon I_{n-1}).(\gamma_1-\bar{\gamma}_1,\cdots,\gamma_{n-1}-\bar{\gamma}_{n-1})^2} d\gamma_1..d\gamma_{n-1}$$
$$\geq \left(\frac{\beta}{2\pi}\right)^{-(n-1)/2} \frac{1}{[\ {}_{n-1}A''(\gamma)+2C\varepsilon I_{n-1}]^{1/2}}(1-e^{-\beta\rho^2\frac{\delta^2}{2}})^{n-1}$$

And more generally,

Lemma 3.2.9. *For all $K \geq 0$, and for all $\varepsilon > 0$, there exists a $\rho > 0$ such that, for all n, for all $nd \times nd$ block-tridiagonal positive symmetric matrix Q satisfying*
- $\| Q_{i,i+1} \| \leq K$ *for all* i.
- $Q \geq \varepsilon I_n$,

then

$$\left(\frac{\beta}{2\pi}\right)^{n/2} \int_{\|x\|_\infty \leq \delta} e^{-\beta\frac{(Qx,x)}{2}} dx_1..dx_n \geq (1-e^{-\beta\rho\frac{\delta^2}{2}})^n \left(\frac{\beta}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} e^{-\beta\frac{(Qx,x)}{2}} dx_1..dx_n$$
$$= \frac{(1-e^{-\beta\rho\frac{\delta^2}{2}})^n}{[Q]^{1/2}}$$

as soon as $\beta\delta^2$ is large enough.

To sum up, the calculations of the last pages lead to the following lower bound:

Lemma 3.2.10. *Assume that $P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}$ contains an element, denoted ξ, in the Mather set. Then, if $|\gamma_0 - \xi_0| \leq c_n\delta$ and $|\gamma_n - \xi_n| \leq c_n\delta$*

$$\left(\frac{\beta}{2\pi}\right)^{\frac{n-1}{2}} \int_{P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}} \exp\left(-\beta\sum_{i=0}^{n-1} L(\gamma_i,\gamma_{i+1})\right) d\gamma_1...d\gamma_{n-1}$$
$$\geq e^{-\beta C\varepsilon}\left((2\pi)^{-1/2}\sqrt{2K^{-1}B}(L+1)e^{-4K^{-1}B(L+1)^2}\right)^{|J(\xi)|} \frac{1}{[\ {}_{n-1}A''(\gamma_0,\gamma_n)+2C\varepsilon I_{n-1}]^{1/2}}(1-e^{-\rho^2\frac{B}{2}})^{n-1}$$

for some $\rho = \rho(\varepsilon) > 0$ depending only on ε.

Let us define

$$Bad(n) = \{(\alpha_0, ..., \alpha_{n-1}),\ \text{for all } \xi \in \mathcal{M} \cap P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}, |J(\xi)| > 2\mu(2\delta)n\}.$$

(The definition of $\mu(\delta)$ was given at the beginning of the proof; it is the μ-measure of a δ-neighbourhood of the boundary of the partition P.)

By Birkhoff's ergodic theorem,

$$\mu(Bad(n)) := \mu\left(\cup_{(\alpha_0,...,\alpha_{n-1})\in Bad(n)} P_{\alpha_0}..\sigma^{-n+1}P_{\alpha_{n-1}}\right) \underset{n\longrightarrow+\infty}{\longrightarrow} 0.$$

To end the proof of Theorem 0.0.1: take $n = kN$ in inequality (3.2.2). Bound the right-hand side from above using Lemma 3.2.4. Bound the left-hand side from below, using the rough estimate (3.2.4) for the cylinders $(\alpha_0, ..., \alpha_{n-1}) \in Bad(n)$, and Lemma 3.2.10 for the other cylinders, for which we know that $|J(\xi)| \leq 2\mu(2\delta)n$.

$$(3.2.12) \quad -\sum \mu(P_{\alpha_0}....\sigma^{-kN+1}P_{\alpha_{kN-1}})\log\mu(P_{\alpha_0}...\sigma^{-kN+1}P_{\alpha_{kN-1}})$$
$$-\frac{1}{2}\sum_{\alpha}\mu(P_{\alpha_0}..\sigma^{-kN+1}P_{\alpha_{kN-1}})\log(\max_{\gamma\in\alpha}[\ _{kN-1}A''(\gamma)+2C\varepsilon I_{kN-1}])+(kN-1)\log(1-e^{-\rho^2\frac{B}{2}})$$
$$-\log(C(\beta)(c_{kN}\delta)^2e^{-\beta C\varepsilon}\frac{\beta}{2\pi})$$
$$-(kN)(\beta K\varepsilon^2-\log\varepsilon)\mu(Bad(kN))-2kN\mu(2\delta)\log\left((2\pi)^{-1/2}\sqrt{2K^{-1}B}(L+1)e^{-4K^{-1}B(L+1)^2}\right)$$
$$\leq -\sum \mu_\beta(P_{\alpha_0}....\sigma^{-kN+1}P_{\alpha_{kN-1}})\log\mu_\beta(P_{\alpha_0}....\sigma^{-kN+1}P_{\alpha_{kN-1}})+\log C(\beta)+k\log(1+\underset{\beta\to\infty}{o(1)})$$
$$+k\sum \mu_\beta(P_{\alpha_0}....\sigma^{-N+1}P_{\alpha_{N-1}})\log F_N(\alpha_0,..,\alpha_{N-1})+k\log B_N$$

for β large enough. Remember that $B=\beta\delta^2$ is fixed (and arbitrary).

We notice that $\log\max_{\gamma\in\alpha}[\ _{n-1}A''(\gamma)+2C\varepsilon I_{n-1}]$, as a function of the sequence $(\alpha_0,..,\alpha_{n-1})$, has the following subadditivity property: if $(\alpha_0,..,\alpha_{n-1})$ intersects the Mather set, then

$$\log\max_{\gamma\in\alpha}[\ _{n-1}A''(\gamma)+2C\varepsilon I_{n-1}]$$
$$\leq \log\max_{\gamma\in\alpha}[\ _mA''(\gamma)+2C\varepsilon I_m]+\log\max_{\gamma\in\alpha}[\ _{n-1-m}A''(\sigma^m\gamma)+2C\varepsilon I_{n-1-m}]$$

This follows directly from Lemma 3.2.1.

As a consequence, if μ is an (invariant) minimizing measure, then

$$\frac{1}{kN}\sum \mu(P_{\alpha_0}..\sigma^{-kN+1}P_{\alpha_{kN-1}})\log\max_{\gamma\in\alpha}[\ _{kN-1}A''(\gamma)+2C\varepsilon I_{kN-1}]$$

converges to its infimum, as $k\longrightarrow+\infty$. And in particular, the limit is less than

$$\frac{1}{N}\sum \mu(P_{\alpha_0}..\sigma^{-N+1}P_{\alpha_{N-1}})\log\max_{\gamma\in\alpha}[\ _{N-1}A''(\gamma)+2C\varepsilon I_{N-1}]$$

Thus, if we divide both side of (3.2.12) by kN and let k tend to ∞ (β being kept fixed), we get the inequality:

$$h_\sigma(\mu,P)-\frac{1}{2}\int_{W^P}\frac{1}{N}\log\max_{\gamma\in\alpha}[\ _{N-1}A''(\gamma)+2C\varepsilon I_{N-1}]d\mu^P(\alpha)$$
$$-\log(1-e^{-\rho^2\frac{B}{2}})-2\mu(2\delta)\log\left((2\pi)^{-1/2}\sqrt{2K^{-1}B}(L+1)e^{-4K^{-1}B(L+1)^2}\right)^{|J(\xi)|}$$
$$\leq h_\sigma(\mu_\beta,P)+\int_{W^P}\frac{1}{N}\log F_N(\alpha)d\mu_\beta^P(\alpha)+\frac{1}{N}(\underset{\beta\to\infty}{o(1)})+\frac{\log B_N}{N}$$

We used the fact (assumed in (A2)) that $\log c_n/n\longrightarrow 0$. The first term $h_\sigma(\mu,P)$ is the metric entropy of the invariant measure μ, with respect to the partition P and the transformation σ on W; in other words, it is the metric entropy of the measure μ^P on the subshift of finite type W^P.

Now, let $\beta\longrightarrow+\infty$; or more precisely, take a sequence β_k such that μ_{β_k} converges weakly to μ_∞. Since we have assumed that μ_∞ does not charge the boundary of the elements of the partition, we get

$$h_\sigma(\mu,P)-\frac{1}{2}\int_{W^P}\frac{1}{N}\log\max_\alpha[\,{}_{N-1}A''(\gamma)+2C\varepsilon I_{N-1}])d\mu^P(\alpha)$$
$$-\log(1-e^{-\rho^2\frac{B}{2}})\leq h_\sigma(\mu_\infty,P)+\int_{W^P}\frac{1}{N}\log F_N(\alpha)d\mu_\infty^P(\alpha)+\frac{\log B_N}{N}$$

The point in fixing N was to integrate only functions depending on a finite number of coordinates, so as to be able to pass to the weak limit.

At this stage, we can let $B\longrightarrow+\infty$, so that $\log(1-e^{-\rho^2\frac{B}{2}})\longrightarrow 0$.

Now, letting ε (the diameter of the partition P) tend to 0, and recalling the definition of F_N,

$$h_\sigma(\mu)-\frac{1}{2}\int\frac{1}{N}\log[\,{}_{N-1}A''(\gamma)]d\mu(\gamma)$$
$$\leq h_\sigma(\mu_\infty)-\frac{1}{2}\int\frac{1}{N}\log[\,{}_{N-1}A''(\gamma)]d\mu_\infty(\gamma)+\frac{\log B_N}{N}$$

and, finally, letting $N\to+\infty$ (and using Assumption (A3)), we get the result. □

This ends the proof of Theorem 0.0.1. It remains to prove that the functional

$$\mu\mapsto h_\sigma(\mu)-\frac{1}{2}\lim_{n\to\infty}\int_W\frac{1}{n}\log[{}_nA''(\bar\gamma)]d\mu(\bar\gamma)$$

is finite on energy-minimizing measures.

4. Frenkel-Kontorova models and twist-maps

4.1. We now give (without proofs) a few links between Frenkel-Kontorova models and symplectic twist diffeomorphisms of $\mathbb{R}^d\times\mathbb{R}^d$. We refer to [AMB92] for a detailed discussion. This section will allow us to prove that the term $h_\sigma(\mu)-\frac{1}{2}\lim_{n\to\infty}\int_W\frac{1}{n}\log[{}_nA''(\bar\gamma)]d\mu(\bar\gamma)$ is finite in Theorem 0.0.1. It also provides a link with Part II, which is more focused on the lagrangian aspects of the problem.

If L satisfies the "twist property" (cf Section 1), it is shown in [AMB92] how to associate to the Frenkel-Kontorova model, discussed above, a symplectic "twist diffeomorphism" of $\mathbb{R}^d\times\mathbb{R}^d$ to itself : this map ϕ^* is defined by

$$(x',p')=\phi^*(x,p)\Longleftrightarrow\partial_2L(x,x')=p',-\partial L_1(x,x')=p.$$

Recall the definition of a stationary configuration for the Frenkel-Kontorova model: it is a sequence $(\gamma_k)_{k\in\mathbb{Z}}$ such that

$$\partial_2L(\gamma_{k-1},\gamma_k)+\partial_1L(\gamma_k,\gamma_{k+1})=0, \tag{4.1.1}$$

for all k. There is a homeomorphism between $\mathbb{R}^d\times\mathbb{R}^d$ and the set of stationary configurations of the Frenkel-Kontorova model, given by

$$(x,p)\mapsto(\gamma_k)_{k\in\mathbb{Z}}$$

where, for all k, γ_k is the projection of $\phi^{*k}(x,p)$ on the first factor $\mathbb{R}^d$. Besides, this homeomorphism is a conjugacy between ϕ^* and the shift σ restricted to the set of stationary configurations.

This way, one can associate to every Frenkel-Kontorova satisfying the "twist property", a symplectic twist diffeomorphism; and conversely, to every symplectic twist map of $\mathbb{R}^d\times\mathbb{R}^d$, a Frenkel-Kontorova model with configuration space $(\mathbb{R}^d)^{\mathbb{Z}}$.

We can also introduce another diffeomorphism ϕ of $\mathbb{R}^d \times \mathbb{R}^d$ to itself, defined by

$$\phi(\gamma_{-1}, \gamma_0) = (\gamma_0, \gamma_1)$$

where $\gamma_{-1}, \gamma_0, \gamma_1$ are related by (4.1.1) with $k = 0$; equivalently,

$$\phi^*(\gamma_{-1}, p_{-1}) = (\gamma_0, p_0), \phi^*(\gamma_0, p_0) = (\gamma_1, p_1)$$

for some p_{-1}, p_0, p_1. The bijection

$$\theta : (\gamma_k)_{k\in\mathbb{Z}} \mapsto (\gamma_0, \gamma_1) \tag{4.1.2}$$

between the set of stationary configurations and $\mathbb{R}^d \times \mathbb{R}^d$ is a conjugacy between the shift and ϕ. For this reason, we will sometimes call stationary configurations "trajectories of ϕ". The conjugacy θ also allows to identify ϕ-invariant probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ and σ-invariant probability measures carried by the set of stationary configurations.

If L is $\mathbb{Z}^d$-periodic, it is natural to take the quotient space $W = (\mathbb{R}^d)^{\mathbb{Z}}/\mathbb{Z}^d$ as configuration space for the Frenkel-Kontorova model. A stationary configuration in W is defined as the image of a stationary configuration in $(\mathbb{R}^d)^{\mathbb{Z}}$ in the quotient. The diffeomorphism ϕ^* can be then be defined on the quotient $\mathbb{R}^d/\mathbb{Z}^d \times \mathbb{R}^d = \mathbb{T}^d \times \mathbb{R}^d$, and the diffeomorphism ϕ, on the quotient space $(\mathbb{R}^d \times \mathbb{R}^d)/\mathbb{Z}^d$, the action of $\mathbb{Z}^d$ in this last case is defined by

$$s.(x, y) = (x + s, y + s)$$

for all $s \in \mathbb{Z}^d$, for all $x, y \in \mathbb{R}^d$.

The conjugacies defined previously between the action of the shift σ on the set of stationary configurations, and ϕ or ϕ^*, are compatible with the actions of $\mathbb{Z}^d$, and thus go to the quotient spaces.

The description of energy-minimizing configurations (rather called action-minimizing in this context) is precisely the heart of what is called "Mather theory" in the study of symplectic exact diffeomorphisms of $\mathbb{T}^d \times \mathbb{R}^d$. In this context, what is classically called "Mather set" is the subset of $\theta(\mathcal{M}) \subset (\mathbb{R}^d \times \mathbb{R}^d)/\mathbb{Z}^d$, where $\mathcal{M} \subset W$ is the set defined at the end of 3.1. The Mather set, as a subset of $(\mathbb{R}^d \times \mathbb{R}^d)/\mathbb{Z}^d$, is compact (Lemma 3.1.7), and ϕ-invariant.

We will say more about Mather theory in Part II, in the case of a continuous time dynamical system. The function L will be called a Lagrangian, ϕ will be the associated Euler-Lagrange flow, and ϕ^* the corresponding Hamiltonian flow.

4.2. **Determinants and Lyapunov exponents.** Motivated by a paper by Thouless ([Th72]) in dimension $d = 1$, we now give a relation between the Hessian of the energy, and Lyapunov exponents. This relation is not new; in the case of a continuous time Lagrangian systems, it is known as the Levit-Smilansky formula (Paragraph 6.3).

Lyapunov exponents are defined by Oseledets' theorem (see for instance [KH95], Supplement), which we use in the following form:

Oseledets' theorem: Let $\phi : (\mathbb{R}^d \times \mathbb{R}^d)/\mathbb{Z}^d \longrightarrow (\mathbb{R}^d \times \mathbb{R}^d)/\mathbb{Z}^d$ be a C^1 diffeomorphism, and let ν be a ϕ-invariant probability measure, carried by a compact subset of $(\mathbb{R}^d \times \mathbb{R}^d)/\mathbb{Z}^d$. Then, there exists a set $Y \subset (\mathbb{R}^d \times \mathbb{R}^d)/\mathbb{Z}^d$ such that $\nu(Y) = 1$, $\phi(Y) = Y$, and such that:

For all $y \in Y$, there exists $r(y) \in \mathbb{N}$, and real numbers $\chi_1(y) < \chi_2(y) < ... < \chi_{r(y)}(y)$, such that the tangent space $T_y((\mathbb{R}^d \times \mathbb{R}^d)/\mathbb{R}^d) \simeq \mathbb{R}^d \times \mathbb{R}^d$ admits a decomposition

$$T_y(\mathbb{T}^d \times \mathbb{R}^d) = E_1(y) \oplus ... \oplus E_{r(y)}(y)$$

satisfying:

$$\forall v \in E_i(y) \setminus \{0\}, \frac{1}{n} \log ||d(\phi)^n_y.v|| \underset{n \longrightarrow \pm\infty}{\longrightarrow} \chi_i(y).$$

The decomposition is ϕ-invariant, in the sense that $r(\phi(y)) = r(y)$ and $E_i(\phi(y)) = d\phi_y.E_i(y)$.

The subspace $E^s(y) = \oplus_{\chi_i(y)<0} E_i(y)$ is called the stable subspace at y, $E^u(y) = \oplus_{\chi_i(y)>0} E_i(y)$ is the unstable subspace, and $E^0(y) = E_i(y)$ fir $\chi_i(y) = 0$ is called the neutral subspace.

In our situation, we adopt a slightly different convention for the Lyapunov exponents. Since ϕ is conjugate to the symplectic diffeomorphism ϕ^*, its Lyapunov exponents come into pairs $(\lambda, -\lambda)$. We denote

$$-\lambda_1^+(y) \leq -\lambda_2^+(y) + \leq ... \leq -\lambda_d^+(y) \leq 0 \leq \lambda_d^+(y) \leq ... \leq \lambda_1^+(y)$$

the Lyapunov exponents; they are now repeated with multiplicity, according to the dimensions of the corresponding subspaces E_j.

Lemma 4.2.1. *Let γ be a trajectory of ϕ, and let $n \in \mathbb{N}$. Let us consider the equation (4.1.1) linearized at $(\gamma_i)_{i\in\mathbb{Z}}$:*

$$\partial_{12}L(\gamma_{i-1}, \gamma_i).Y_{i-1} + (\partial_{22}L(\gamma_{i-1}, \gamma_i) + \partial_{11}L(\gamma_i, \gamma_{i+1})).Y_i + \partial_{21}L(\gamma_i, \gamma_{i+1}).Y_{i+1} = 0,$$

for all $i \in \mathbb{Z}$.

Fix the initial condition $Y_0 = 0$.

Then, for all n, the determinant of the linear map $Y_1 \longmapsto Y_n$ (from $\mathbb{R}^d$ to $\mathbb{R}^d$) is equal to the determinant

$$(-1)^{nd} (\prod_{k=1}^{n} [A''_{k+1,k}])^{-1} \times [\ {}_{n-1}A''].$$

Proof. A vector $Y = (Y_1, \cdots, Y_n)$ $(Y_i \in \mathbb{R}^d)$ satisfies ${}_nA''.Y = (0, 0, 0, \cdots, 0, *)$, if and only if Y is the solution of the linearized equation with $Y_0 = 0$.

Equivalently,

$$(Y_{n-1}, Y_n) = d(\phi^{n-1})_{(\gamma_0, \gamma_1)}.(0, Y_1)$$

The matrix ${}_{n-1}A''$ is non-invertible if and only if there exists Y with $Y_1 \neq 0$ and $Y_n = 0$ such that ${}_nA''.Y = (0, 0, 0, \cdots, 0, *)$; that is, the map $Y_1 \mapsto Y_n$ is not invertible. In this case, Lemma 4.1.1 is obvious. Thus, let us assume that the map $Y_1 \mapsto Y_n$ is invertible.

Also assume for the momen that ${}_nA''(\gamma)$ is invertible. Me may then decompose the matrix $G = \ {}_nG = {}_nA''(\gamma)^{-1}$ into $d \times d$ blocks $(G_{ij})_{1\leq i,j\leq n}$. The components Y_1 and Y_n are related by:

$$Y_n = G_{nn}.G_{n1}^{-1} Y_1$$

Let us evaluate the determinant of $G_{nn}.G_{n1}^{-1}$ in terms of the determinant of $A''(\gamma)$. We first define a sequence of $d \times d$ matrices $(a_0, a_1, \cdots, a_{n-1})$ by $a_0 = Id$ and

$$a_k = -A''_{k+1,k}(A''_{kk} + a_{k-1}A''_{k-1,k})^{-1},$$

agreeing here that $A''_{01} = 0$ (the sequence is well defined if ${}_nA''$ has been assumed invertible).

We also define an $nd \times nd$ matrix T decomposed into $d \times d$ blocks $(T_{ij})_{1 \leq i,j \leq n}$ with

$$T_{ii} = Id$$
$$T_{ij} = \prod_{k=j}^{i} a_{i-k}$$

(this way, T is lower block-triangular). In fact, the matrix T is constructed in such a way that $D = T.\ {}_nA''$ is an upper block triangular matrix, with blocks on the diagonal

$$D_{kk} = D_k = A''_{kk} + a_{k-1}A''_{k-1,k}$$

We have $G = D^{-1}T$ which yields $G_{nn}G_{n1}^{-1} = D_n T_{n1}^{-1} D_n^{-1}$ so that

$$\begin{aligned} [G_{nn}G_{n1}^{-1}] &= [T_{n1}]^{-1} \\ &= (\prod_{k=1}^{n-1}[a_{n-k}])^{-1} \\ &= (-1)^{nd}(\prod_{k=1}^{n}[A''_{k+1,k}])^{-1} \times \prod_{k=1}^{n-1}[D_k] \\ &= (-1)^{nd}(\prod_{k=1}^{n}[A''_{k+1,k}])^{-1} \times [\ {}_{n-1}A''] \end{aligned}$$

where the last equality comes from the observation that $[\ {}_{n-1}A''] = [\ {}_{n-1}D]$.

Thus, the determinant of $Y_1 \mapsto Y_n$ is equal to $(-1)^{nd}(\prod_{k=1}^{n}[A''_{k+1,k}])^{-1} \times [\ {}_{n-1}A'']$. By continuity of both expressions with respect to the energy functional, Lemma 4.1.1 remains valid even when ${}_nA''$ is not invertible. □

Proposition 4.2.2. *If μ is an energy-minimizing measure on W, then the limit $\lim \frac{1}{n}\log[{}_nA''(\bar{\gamma})]$ exists for μ-almost every $\bar{\gamma}$, and is equal to*

$$\sum_{1}^{d} \lambda_i^+(\gamma_0, \gamma_1) + \lim \frac{1}{n}\sum_{i=0}^{n-1} \log \mid [\partial^2_{12}L(\gamma_i, \gamma_{i+1})] \mid,$$

where the $\lambda_i^+(\gamma_0, \gamma_1)$ are the d-first (nonnegative) Lyapunov exponents of $\overline{(\gamma_0, \gamma_1)}$ under the diffeomorphism ϕ.

Proof. The existence of the limit $\lim \frac{1}{n}\sum_{i=0}^{n-1} \log \mid [\partial^2_{12}L(\gamma_i, \gamma_{i+1})] \mid$ for μ-almost every $\bar{\gamma}$ is guaranteed by Birkhoff's theorem, applied to the function

$$\begin{aligned} W &\longrightarrow \mathbb{R} \\ \bar{\gamma} &\mapsto \log \mid [\partial^2_{12}L(\gamma_0, \gamma_1)] \mid \end{aligned}$$

We denote $\theta * \mu$ the image of μ under the conjugacy θ (4.1.2). It is a ϕ-invariant probability measure. Let us now denote $\Lambda^d(\mathbb{R}^d \times \mathbb{R}^d)$ the d-fold exterior product of $\mathbb{R}^d \times \mathbb{R}^d$. It is endowed with the euclidean structure coming from the euclidean structure on $\mathbb{R}^d \times \mathbb{R}^d$. The Oseledets theorem implies that, for $\theta * \mu$-almost every $y \in (\mathbb{R}^d \times \mathbb{R}^d)/\mathbb{Z}^d$, for every $P \in \Lambda^d(\mathbb{R}^d \times \mathbb{R}^d)$, the limit

$$\lim_{n \longrightarrow \pm\infty} \frac{1}{n} \log ||d(\phi^n)_y.P||$$

exists (besides, it is of the form $\sum_{i=1}^{d} \varepsilon(i)\lambda^+_{\alpha(i)}$, where $\varepsilon(i) = \pm 1$, $\alpha(i) \in \{1, ..., d\}$, the pairs $(\varepsilon(i), \alpha(i))$ all distinct). We denote this limit $\lambda_P(y)$.

Denote $(e_1, e_2, ..., e_{2d})$ an orthonormal basis of $\mathbb{R}^d \times \mathbb{R}^d$ such that $(e_1, e_2, ..., e_d)$ is an orthonormal basis of $\mathbb{R}^d \times \{0\}$ and $(e_{d+1}, ..., e_{2d})$ is an orthonormal basis of $\{0\} \times \mathbb{R}^d$. The Birkhoff and Oseledets theorems, combined with Lemma 4.2.1, imply that, for μ-almost every $\bar{\gamma}$,

$$\lim \frac{1}{n} \log[\,{}_nA''(\bar{\gamma})] = \lambda_P(\gamma_0, \gamma_1) + \lim \frac{1}{n} \sum_{i=0}^{n-1} \log \mid [\partial^2_{12} L(\gamma_i, \gamma_{i+1})] \mid$$

where $P = e_{d+1} \wedge e_{d+2} ... \wedge e_{2d}$. Indeed, λ_P is precisely the exponential growth rate of the determinant of $Y_1 \mapsto Y_n$, for the fixed initial condition $Y_0 = 0$. This relation also shows that $\lambda_P(\gamma_0, \gamma_1) = \lambda_P(\gamma_1, \gamma_2)$ for μ-almost every $\bar{\gamma}$.

Let $\overline{(\gamma_0, \gamma_1)}$ be in the Mather set, and $\bar{\gamma} = \theta^{-1}\overline{(\gamma_0, \gamma_1)} \in \mathcal{M}$. We show that $\{0\} \times \mathbb{R}^d \subset T_{\overline{(\gamma_0,\gamma_1)}}(\mathbb{R}^d \times \mathbb{R}^d)/\mathbb{Z}^d$ is transverse to the stable subspace at $\overline{(\gamma_0, \gamma_1)}$. Otherwise, there would exist an element $(\xi_i)_{i\in\mathbb{Z}}$ in the kernel of $A''(\bar{\gamma})$, such that $\xi_0 = 0$ and $\xi_i \underset{i\to+\infty}{\longrightarrow} 0$ exponentially fast. Define an element $\zeta \in (\mathbb{R}^d)^{\mathbb{Z}}$ by

$$\zeta_j = 0 \text{ for } j \leq 0,$$

$$\zeta_j = \xi_j \text{ for } j \geq 0.$$

Then $\zeta \in l^2(\mathbb{Z}, \mathbb{R}^d)$, and, since $A''(\gamma).\xi = 0$ and the first coordinate of ζ vanishes,

$$\langle A''(\bar{\gamma})\zeta, \zeta\rangle = 0.$$

But, since $\bar{\gamma}$ is in the Mather set, $A''(\bar{\gamma})$ is a positive semi-definite operator in $l^2(\mathbb{Z}, \mathbb{R}^d)$. Thus, the function $\langle A''(\bar{\gamma}).,.\rangle$ achieves a minimum at ζ. Its derivative at ζ must vanish: $A''(\gamma).\zeta = 0$.

An element of $KerA''(\bar{\gamma})$ is entirely determined by two successive coordinates. Since $\zeta_{-1} = \zeta_0 = 0$, we have $\zeta = 0$. The same argument now shows that $\xi = 0$.

We have thus shown that $\{0\} \times \mathbb{R}^d \subset T_{(\gamma_0,\gamma_1)}(\mathbb{R}^d \times \mathbb{R}^d)$ is transverse to the stable subspace for ϕ.

Now, let $\overline{(\gamma_0, \gamma_1)}$ be a point satisfying the conclusions of Oseledets' theorem with respect to μ, and $\bar{\gamma} = \theta^{-1}\overline{(\gamma_0, \gamma_1)}$. Let $E^0(\gamma_0, \gamma_1)$ be the neutral subspace, and $d_0 = dimE^0(\gamma_0, \gamma_1)$. Note that d_0 is even. We have

$$\lambda_P(\gamma_0, \gamma_1) = \sum_{i=1}^{d} \lambda_i^+(\gamma_0, \gamma_1)$$

unless $dim(E^0(\gamma_0, \gamma_1) \cap (\{0\} \times \mathbb{R}^d)) > d_0/2$.

Remember however that $\lambda_P(\gamma_0, \gamma_1) = \lambda_P(\gamma_1, \gamma_2)$. We cannot have simultaneously $dim(E^0(\gamma_0, \gamma_1) \cap (\{0\} \times \mathbb{R}^d)) > d_0/2$ and $dim(E^0(\gamma_1, \gamma_2) \cap (\{0\} \times \mathbb{R}^d)) > d_0/2$. Otherwise,

$$\theta^{-1}\left(E^0(\gamma_0, \gamma_1) \cap (\{0\} \times \mathbb{R}^d)\right)\right) \oplus \theta^{-1}\sigma^{-1}\left(E^0(\gamma_1, \gamma_2) \cap (\{0\} \times \mathbb{R}^d)\right)\right)$$

would be a subspace of dimension $> d_0$ of $KerA''(\bar{\gamma})$, composed of sequences $(\xi_k)_{k\in\mathbb{Z}}$ such that

$$\lim_{k\to\pm\infty} \frac{1}{k} \log ||(\xi_k, \xi_{k+1})|| = 0,$$

a contradiction with the definition of d_0. □

An example. Consider the example

$$L(\gamma_0,\gamma_1) = \frac{||\gamma_1-\gamma_0||^2}{2} - V(\gamma_0) - \langle \omega, \gamma_1-\gamma_0\rangle,$$

ω a vector in $\mathbb{R}^d$.

Then, $\partial^2_{12}L$ is $-Id$, so that $\log|[\partial^2_{12}L]| = 0$. In this situation, we simply get that any limit point μ_∞ of the (μ_β)s maximizes the functional

$$\mu \mapsto h_\sigma(\mu) - \frac{1}{2}\int_W \sum_{i=1}^{d} \lambda_i^+(\gamma_0,\gamma_1)d\mu(\bar\gamma).$$

The reason for the additional term $\lim \frac{1}{n}\sum_{i=0}^{n-1} \log \mid [\partial^2_{12}L(\gamma_i,\gamma_{i+1})] \mid$ in the general situation is that, in the definition of Gibbs measures, we have chosen the flat Lebesgue measure on $\mathbb{R}^d$ as the reference measure, although there was no reason to favour this choice amongst other smooth measures. In the special case $L(\gamma_0,\gamma_1) = \frac{||\gamma_1-\gamma_0||^2}{2} - V(\gamma_0) - \langle\omega,\gamma_1-\gamma_0\rangle$, the function L is defined in reference to a certain euclidean structure on $\mathbb{R}^d$, so that it is natural to take the associated Lebesgue measure as reference measure.

In the special case $\omega = 0$, we may assume that $\max V = 0$, then the Mather set is

$$\mathcal{M} = \{\bar\gamma, \exists x \in \mathbb{R}^d, V(x) = 0, \gamma_i = x \text{ for all } i\}.$$

The entropy of any energy-minimizing measure vanishes, since $\mathcal{M}$ consists in fixed points of σ. So, in Theorem 0.0.1, the functional reduces to the sum of nonnegative Lyapunov exponents. Finally, in this situation, we will prove in paragraph 6.4 that assumptions (A2) and (A3) are always satisfied.

Part 2. Lagrangian dynamics

5. Hamilton-Jacobi, Aubry-Mather and Schrödinger

Let $\mathbb{R}^d$ be endowed with its usual euclidean structure denoted $\langle .,.\rangle$, and let us consider the lagrangian

$$\mathcal{L}(x,v) = \frac{\| v \|^2}{2} - V(x)$$

on $\mathbb{R}^d \times \mathbb{R}^d$, V being a $\mathbb{Z}^d$-periodic potential of class C^3, and $||.||$ being the norm associated to the scalar product $\langle .,.\rangle$.

For $\omega \in \mathbb{R}^d$, perform the change of gage

$$\mathcal{L}_\omega(x,v) = \frac{\| v \|^2}{2} - V(x) - \langle \omega, x\rangle,$$

in the definition of the Lagrangian.

The corresponding Hamiltonian (energy) is then

$$H_\omega(x,p) = \frac{||p+\omega||^2}{2} + V(x)$$

on $\mathbb{R}^d \times \mathbb{R}^d$.

The Euler-Lagrange flow is the flow (ϕ_t) on $\mathbb{R}^d \times \mathbb{R}^d$, defined by $\phi_t(x_0,v_0) = (\gamma(t),\dot\gamma(t))$ where γ is the solution to the second order equation

$$\ddot\gamma_t = -V'(\gamma_t),$$

with the initial condition $\gamma(0) = x_0$, $\dot\gamma(0) = v_0$.

Trajectories of the Euler-Lagrange flow are characterized by a variational principle: let $\gamma : [a,b] \longrightarrow \mathbb{R}^d$ be a C^1 curve. Then $t \mapsto (\gamma_t, \dot{\gamma}_t)$ is trajectory of the Euler-Lagrange flow if and only if γ is a critical point of the action functional

$$\mathcal{A}(\xi) = \int_a^b \mathcal{L}(\xi_t, \dot{\xi}_t)dt,$$

restricted to the set of C^1 curves $\xi : [a,b] \longrightarrow \mathbb{R}^d$ such that $\xi(a) = \gamma(a), \xi(b) = \gamma(b)$.

The dynamics is described in an equivalent way by the Hamiltonian flow, whose trajectories are solution to the system of Hamilton equations

$$\begin{cases} \dot{x} = \partial_p H_\omega(x,p) \\ \dot{p} = -\partial_x H_\omega(x,p) \end{cases}$$

on $\mathbb{R}^d \times \mathbb{R}^d$. Moreover, the energy is constant along trajectories of the flow.

Since V is periodic, both the Euler-Lagrange and the Hamiltonian flow can be defined on the quotient space $\mathbb{T}^d \times \mathbb{R}^d$.

When one tries to understand the action of the Hamiltonian flow on the phase space $\mathbb{T}^d \times \mathbb{R}^d$, it is natural to try and find invariant regions. Of particular interest are invariant lagrangian graphs, that is, invariant subsets of the form

$$\{(x, \omega + du(x))\} \subset \mathbb{T}^d \times \mathbb{R}^d,$$

($\omega \in \mathbb{R}^d$, $u : \mathbb{T}^d \longrightarrow \mathbb{R}$ as smooth as possible). Such a subset, if it exists, projects diffeomorphically to the base $\mathbb{T}^d$.

For this subset to be invariant, it is necessary and sufficient that there exist a constant C such that u satisfies the stationary Hamilton-Jacobi equation (HJ):

$$H_\omega(x, du(x)) = C$$

for all x.

However, generally speaking, the Hamilton-Jacobi equation (HJ) may have no smooth solution. There are two ways out: the theory of viscosity solutions, and Mather theory. The connection between the two approaches has been made very clear by the recent, and still mostly unpublished, work of Fathi ([Fa]).

Viscosity solutions.

Let us consider the equation

$$-\varepsilon \Delta u + H_\omega(x, d_x u) = C,$$

for $\varepsilon \geq 0$. When $\varepsilon = 0$, it is equation (HJ); otherwise, it is called the viscous, stationary Hamilton-Jacobi equation, (HJV).

A continuous $\mathbb{Z}^d$-periodic function u is called a viscosity solution of the equation if, for every C^1-function (C^2 in the case of (HJV)) ϕ,

– if $u - \phi$ attains a local maximum at y_0, then $-\varepsilon\Delta\phi(y_0) + H(x, d_x\phi(y_0)) \leq C$, and

– if $u - \phi$ attains a local minimum at y_0, then $-\varepsilon\Delta\phi(y_0) + H(x, d_x\phi(y_0)) \geq C$.

See, for instance, [CEL84], [CL83], [Ba94]. A justification of this definition is that it coincides with the classical notion of solutions if u is smooth.

It may be checked that, if a viscosity solution exists, then

$$\inf H(x,0) \leq C \leq \sup H(x,0),$$

by applying the definition to $\phi = 0$ and y_0, successively, a local maximum or minimum of u.

Besides, since $H_\omega(x,p)$ tends to infinity with $||p||$, viscosity solutions are lipschitz and share a common lipschitz constant. Indeed, consider $\phi_x(y) = K||x-y||$, and apply the definition. Assume that $u - \phi_x$ has a local maximum at $y \neq x$, then

$$H_\omega(x, K\frac{y-x}{||y-x||}) \leq C,$$

which is not possible if K has been chosen large enough. Thus, $u - \phi_x$ attains its maximum at x, which means that

$$u(y) - u(x) \leq K||x-y||.$$

The definition of viscosity solution also holds, with obvious modifications, for the evolutive Hamilton-Jacobi equation:

$$\partial_t u + H_\omega(x, d_x u) = 0.$$

We are in the situation when $H_\omega(x,p)$ is C^2, superlinear in p, strictly convex. In this case, given a continuous initial condition u_0, the solution u_t is unique ([Ba94], Theorem 2.8), and given by the expression:

$$u_t(x) = \inf_\gamma\{u(\gamma(0)) + \int_{-t}^0 \mathcal{L}_\omega(\gamma(s), \dot\gamma(s))ds\},$$

where the inf is taken over all curves $\gamma : [-t, 0] \longrightarrow V$ with square integrable derivatives, and such that $\gamma(-t) = x$. Moreover, u_t is lipschitz in x for all $t > 0$ ([Fa97-1]).

The semi-group $(T_t^-)_{t\geq 0}$ on $C(\mathbb{T}^d, \mathbb{R})$, defined by

$$T_t^- u(x) = \inf_\gamma\{u(\gamma(0)) + \int_{-t}^0 \mathcal{L}_\omega(\gamma(s), \dot\gamma(s))ds\},$$

is called the Hopf-lax or Lax-Oleinik semi-group, according to sources.

There is also a second semi-group $(T_t^+)_{t\geq 0}$

$$T_t^+ v(x) = \sup_{\gamma:[0,t]\to\mathbb{T}^d, \gamma(0)=x}\{v(\gamma(t)) - \int_0^t \mathcal{L}_\omega(\gamma(s), \dot\gamma(s))ds\},$$

which gives the opposite of the solutions of the equation

$$\partial_t u + H_\omega(x, -d_x u) = 0,$$

which corresponds to reversing the orientation of time.

Notice that u is a viscosity solution of the stationary (HJ) equation if and only if $u - Ct$ is a solution of the evolutive equation. Thus, looking for solutions of the stationary equation is the same as solving the fixed point problem:

$$T_t^- u = u + Ct,$$

for all t.

The existence of such fixed points is given by a theorem of Fathi, called the "weak KAM theorem":

Theorem 5.0.3. *(*[Fa], [Fa97-1]*) For a unique constant $c = c(\omega) \in \mathbb{R}$, there exist continuous functions u_- and u_+ on $\mathbb{T}^d$, solutions to the following fixed points problem:*

$$T_t^- u_- = u_- + ct$$

and

$$T_t^+ u_+ = u_+ - ct.$$

They are lipschitz, and satisfy the two following properties:
– for all C^1 curve $\gamma : [0,t] \longrightarrow V$,

$$u_\pm(\gamma(t)) - u_\pm(\gamma(0)) \leq \int_0^t \mathcal{L}_\omega(\gamma(s),\dot\gamma(s))ds + c(\omega)t$$

– for all $x \in V$, there exist two curves $\gamma_- : (-\infty, 0] \longrightarrow V$ and $\gamma_+ : [0,+\infty) \longrightarrow V$ with $\gamma_-(0) = \gamma_+(0) = x$, such that, for all $t \geq 0$,

$$u_-(x) - u_-(\gamma_-(-t)) = \int_{-t}^0 \mathcal{L}_\omega(\gamma_-(s),\dot\gamma_-(s))ds + c(\omega)t$$

and

$$u_+(\gamma_+(t)) - u_+(x) = \int_0^t \mathcal{L}_\omega(\gamma_+(s),\dot\gamma_+(s))ds + c(\omega)t$$

Note that the theorem does not assert the uniqueness of solutions. Obviously, if u is a solution, then $u + K$ also, for all $K \in \mathbb{R}$, but there may even be solutions which do not differ by a constant.

Of course, the existence of viscosity solutions of the stationary (HJ) solution was known before this theorem. It was usually proved by the “vanishing viscosity method”, which consists, first, in finding a solution u_ε for (HJV), then in letting $\varepsilon \longrightarrow 0$ and proving a “stability” result: any limit of u_ε in the uniform topology is a viscosity solution of the non-viscous (HJ) equation ([CEL84], Theorem 3.1).

If u_- is, as above, a (lipschitz) viscosity solution of $H_\omega(x, d_x u) = C$, then it is differentiable almost everywhere: the graph

$$Graph(du_-) \subset \mathbb{T}^d \times \mathbb{R}^d$$

is a graph lying over a set of full Lebesgue measure in $\mathbb{T}^d$, and invariant under the positive times of the Hamiltonian flow. Similarly, if u_+ is as above, then $Graph(-du_+) \subset \mathbb{T}^d \times \mathbb{R}^d$ is a graph over a set of full Lebesgue measure in $\mathbb{T}^d$, and invariant under the negative times of the Hamiltonian flow (see [Fa]).

Applying Theorem 5.0.3 to various ωs, one obtains weak solutions to the problem of finding invariant Lagrangian graphs.

Mather theory. The starting point of Mather theory is the remark that, if $\gamma : \mathbb{R} \longrightarrow \mathbb{T}^d$ is a trajectory of the Euler-Lagrange flow, lying in an invariant lagrangian torus of the form $Graph(\omega + du)$, then γ is a *global* minimizer of the action, meaning that

$$\int_a^b \mathcal{L}_\omega(\gamma_t, \dot\gamma_t)dt \leq \int_a^b \mathcal{L}_\omega(\xi_t, \dot\xi_t)dt$$

for all $a < b$, for all $\xi : [a,b] \longrightarrow \mathbb{T}^d$ of class C^1 such that $\xi(a) = \gamma(a)$, $\xi(b) = \gamma(b) \in \mathbb{T}^d$. Note that this notion of action-minimizing trajectory depends on ω, contrarily to the definition of the Euler-Lagrange flow. It would be more adequate to speak of “ω-action-minimizing” trajectories.

Thus, even if an invariant lagrangian graph associated to ω does not exist, one may still look for action-minimizing trajectories. Or, if we are only interested in invariant measures of the flow, for (ω-)action-minimizing measures : these are defined as probability measures on the phase space, invariant under the flow, and achieving the minimum of the integral

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} \mathcal{L}_\omega(x,v)d\mu(x,v)$$

over the set of all invariant probability measures.

Theorem 5.0.4. *(*[Ma91]*) (a) For each ω, action-minimizing measures do exist.*

(b) For each ω, let us define the Mather set $\tilde{\mathcal{M}}_\omega \subset \mathbb{T}^d \times \mathbb{R}^d$ as the (closure of) the union of supports of ω-action-minimizing measures:

$$\tilde{\mathcal{M}}_\omega = \overline{\cup_{\mu \ act.\ min.} supp \mu}.$$

Then $\tilde{\mathcal{M}}_\omega$ is a compact set, invariant under the Euler-Lagrange flow.

(c) A probability measure, invariant under the Euler-Lagrange flow, is ω-action-minimizing if and only if its support lies in the Mather set $\tilde{\mathcal{M}}_\omega$.

(e) A probability measure, invariant under the Euler-Lagrange flow, is ω-action-minimizing if and only if the trajectories in its support are ω-action-minimizing.

(f) (The Graph Theorem) The projection $\pi : \mathbb{T}^d \times \mathbb{R}^d \longrightarrow \mathbb{T}^d$, restricted to $\tilde{\mathcal{M}}_\omega$, is injective. Its inverse, defined on

$$\mathcal{M}_\omega := \pi(\tilde{\mathcal{M}}_\omega),$$

is Lipschitz.

The links with the theory of viscosity solutions have been made explicit in the (still mostly unpublished) work of Fathi ([Fa], [Fa97-1], [Fa97-2]). For instance, the ω-action-minimizing trajectories of the Euler-Lagrange flow, when carried by Legendre duality to the cotangent bundle, are the *complete* trajectories of the Hamiltonian flow lying in $Graph(du)$, for some viscosity solution u of $H_\omega(x, d_x u) = c(\omega)$ (recall that $Graph(du)$ is, a priori, only invariant by the positive times of the Hamiltonian flow). The Mather set $\tilde{\mathcal{M}}_\omega$ (transported by Legendre duality to the cotangent bundle) is contained in $Graph(du)$ for any such u. Finally, the value of $c(\omega)$ is

$$c(\omega) = -\inf\{\int_{\mathbb{T}^d \times \mathbb{R}^d} \mathcal{L}_\omega(x,v) d\nu(x,v), \nu \text{ a } \phi\text{-invariant probability measure}\}.$$

The constant $c(\omega)$, called the effective hamiltonian in PDE, is called the Mather function (seen as a function of ω) by others; or sometimes, Mañe's critical value for the Lagrangian $\mathcal{L}_\omega$.

Let us end this paragraph by a proposition, due to Fathi, which will be useful later.

Proposition 5.0.5. *(*[Fa97-2]*) For every fixed point of u_- of the semi-group $(T_t^- - c(\omega)t)_{t \geq 0}$, there exists a unique fixed point u_+ of the semi-group $(T_t^+ + c(\omega)t)_{t \geq 0}$ such that u_- and u_+ coincide on $\mathcal{M}_\omega$. They satisfy $u_- \geq u_+$. The fixed points u_- and u_+ are then called conjugate.*

The Schrödinger equation, the viscous Hamilton-Jacobi equation, and the associated stochastic process.

Given a scalar β, we consider the Schrödinger operator on $\mathbb{R}^d$:

$$\mathcal{H}_\beta = \frac{\Delta}{2\beta^2} + V,$$

and we define

$$\mathcal{H}_\beta^\omega = e^{-\beta\langle\omega,x\rangle} \circ \mathcal{H}_\beta \circ e^{\beta\langle\omega,x\rangle}.$$

Although we do not precise for the moment on which space it acts, we can note that, for any β, $\mathcal{H}_\beta^\omega$ preserves (formally) the set of $\mathbb{Z}^d$-periodic functions. For

$\beta = i/h$ (h being the Planck constant), the operator $\mathcal{H}_\beta^\omega$ is the quantization of the classical Hamiltonian H_ω defined earlier. In this paper, however, we consider the case when $\beta > 0$.

We will use the following properties of the operator $\mathcal{H}_\beta$, proved for instance in [AS82], [Si79], II.6 (for the moment, without the change of gage represented by ω):

The operator $\mathcal{H}_\beta$ is essentially self-adjoint, bounded from above. By using functional calculus, one can define the semi-group $(\exp(t\beta\mathcal{H}_\beta))_{t\geq 0}$ of bounded operators on $L^2(\mathbb{R}^d)$. For every $t > 0$, $\exp(t\beta\mathcal{H}_\beta)$ is an integral operator, with a positive kernel $K_\beta^t(x, y)$ depending continuously on $(t, x, y) \in (0, +\infty) \times \mathbb{R}^d \times \mathbb{R}^d$, given by the Feynman-Kac formula ([AS82], [Si79] II.6):

$$K_\beta^t(x, y) = \int_{C([0,t],\mathbb{R}^d)} e^{\beta \int_0^t V(\gamma_u)du} d\mathcal{W}_{[0,t]}^{\beta,(x,y)}(\gamma),$$

where $d\mathcal{W}_{[0,t]}^{\beta,(x,y)}$ denotes the brownian bridge between x and y. It is a positive measure on the set of continuous paths $C([0, t], \mathbb{R}^d)$; its definition is recalled a bit later.

From this formula, and the fact that $\int d\mathcal{W}_{[0,t]}^{\beta,(x,y)}(\gamma) = \left(\frac{\beta}{2\pi t}\right)^{d/2} e^{-\frac{\beta||x-y||^2}{2t}}$, one sees in particular that

$$K_\beta^t(x, y) \leq e^{\beta M t} \left(\frac{\beta}{2\pi t}\right)^{d/2} e^{-\frac{\beta||x-y||^2}{2t}},$$

where M is an upper bound on V.

We now *define* $\exp(t\beta\mathcal{H}_\beta^\omega)$ as

$$\exp(t\beta\mathcal{H}_\beta^\omega) = e^{-\beta\langle\omega,x\rangle} \circ \exp(t\beta\mathcal{H}_\beta) \circ e^{\beta\langle\omega,x\rangle}.$$

For each $t > 0$, it is a kernel operator, with a continuous kernel given by

$$K_{\beta,\omega}^t(x, y) = e^{-\beta\langle\omega,x\rangle} K_\beta^t(x, y) e^{\beta\langle\omega,y\rangle}.$$

It also acts as a kernel operator on the set of $\mathbb{Z}^d$-periodic functions.

Remark 5.0.6. When writing this paper for the first time, the author was not aware that the Feynman-Kac formula also holds (with the necessary modifications) in the presence of a magnetic field, i.e. when ω is replaced by a non-closed 1-form: see [Si79], V.15. This seems to indicate that all the results below also hold in the presence of a magnetic field.

We can apply the results of Section 2 to the operator $P_\beta^+ = \exp(t\beta\mathcal{H}_\beta^\omega)$. There exist positive, $\mathbb{Z}^d$-continuous functions ψ_β and ψ_β^*, and a real number λ_β, such that

$$\exp(t\beta\mathcal{H}_\beta^\omega)\psi_\beta = e^{t\beta\lambda_\beta}\psi_\beta \tag{5.0.1}$$

and

$$\exp(t\beta\mathcal{H}_\beta^\omega)^*\psi_\beta^* = e^{t\beta\lambda_\beta}\psi_\beta^*, \tag{5.0.2}$$

for all t. For each $t > 0$, $e^{t\beta\lambda_\beta}$ is the spectral radius of $\exp(t\beta\mathcal{H}_\beta^\omega)$ in $L^2(\mathbb{T}^d)$, and it is a simple eigenvalue.

(More precisely, the proof of Section 2 would allow us to find such ψ_β, ψ_β^* and λ_β for each given t. But since the operators $\exp(t\beta\mathcal{H}_\beta^\omega)$ commute and since ψ_β, ψ_β^* and λ_β are defined uniquely by equations (5.0.1), (5.0.2), they must be the same for all t.)

Besides, the differentiation of equation (5.0.1) with respect to time yields that

$$\mathcal{H}^{\omega}_{\beta}\psi_\beta = \lambda_\beta\psi_\beta \tag{5.0.3}$$

in the sense of distributions. Since the Laplace operator is elliptic, this implies that ψ_β is of class C^2 (at least) and that (5.0.3) holds in the strong sense. Similarly, ψ^*_β is of class C^2, and

$$\mathcal{H}^{\omega *}_{\beta}\psi^*_\beta = \lambda_\beta\psi^*_\beta. \tag{5.0.4}$$

If we let $u_\beta = -\frac{\log\psi^*_\beta}{\beta}$ and $v_\beta = -\frac{\log\psi_\beta}{\beta}$, a simple computation shows that

$$-\frac{\Delta u}{2\beta} + H_\omega(x, d_x u) = \lambda_\beta$$

and that

$$-\frac{\Delta v}{2\beta} + H_\omega(x, -d_x v) = \lambda_\beta,$$

in other words, u_β is a solution of (HJV) with viscosity coefficient $1/\beta$, and v_β is a solution of (HJV) for the reversed orientation of time.

Remark 5.0.7. We have seen that (u_β) is a uniformly lipschitz family, and that $(\lambda_\beta)_\beta$ is bounded. It follows from the stability result for viscosity solutions ([CEL84], Theorem 3.1), that any limit point u of u_β in the uniform topology is a viscosity solution of

$$H_\omega(x, d_x u) = C,$$

where C is a limit point of λ_β. Since we know (Theorem 5.0.3) that this equation has solutions for $C = c(\omega)$ only, this implies that

$$\lambda_\beta \underset{\beta\to+\infty}{\longrightarrow} c(\omega).$$

We are now interested in the behaviour of the measure

$$d\mu^0_\beta(x) = \psi_\beta(x)\psi^*_\beta(x)dx,$$

which we normalize to give a probability measure on the torus.

Theorem 5.0.8. *Let $\mathcal{H}^{\omega}_{\beta} = e^{-\beta\langle\omega,x\rangle} \circ \left(\frac{\Delta}{2\beta^2} + V(x)\right) \circ e^{\beta\langle\omega,x\rangle}$ and $\mathcal{H}^{\omega *}_{\beta} = \mathcal{H}^{-\omega}_{\beta}$. Let ψ_β, ψ^*_β be the C^2 positive eigenfunctions, defined above.*

Then, as $\beta \to +\infty$, the measure

$$\mu^0_\beta = \frac{\psi_\beta(x)\psi^*_\beta(x)dx}{\int_{\mathbb{T}^d}\psi_\beta(y)\psi^*_\beta(y)dy}$$

on $\mathbb{T}^d$ concentrates on the Mather set $\mathcal{M}_\omega$.

Assume the system satisfies the properties (A1), (A2), (A3) below. If μ^0_∞ is a limit point of (μ^0_β), and μ_∞ is the corresponding action-minimizing measure on $\mathbb{T}^d \times \mathbb{R}^d$, then μ_∞ maximizes the functional

$$h_\phi(\mu) - \frac{1}{2}\int_{\mathbb{T}^d\times\mathbb{R}^d}(\sum_{i=1}^{d}\lambda^+_i(x,v))d\mu(x,v)$$

amongst all action-minimizing measures.

The fact that μ_∞^0 can be lifted in a unique way to an action-minimizing measure comes from the Graph Theorem (Theorem 5.0.4 (f)).

In the theorem, $h_\phi(\mu)$ stands for the metric entropy of the invariant measure μ with respect to the Euler-Lagrange flow (ϕ_t), and the $\lambda_i^+(x,v)$ stand for the non-negative Lyapunov exponents of (x,v) under the action of the flow. See Paragraph 4.2 for a definition of Lyapunov exponents.

Remark 5.0.9. The first point of the theorem (i.e. the concentration on the Mather set) was already known to a number of people, see for instance [Go02] (Section 8) where the measures μ_β^0 appear under the name "stochastic Mather measures". The point in the theorem is the variational principle satisfied by μ_∞.

Before going on, we need to introduce notations for various path spaces and for the Wiener measure(s).

We denote $H_{[0,t]}$ the Hilbert space of paths $[0,t] \longrightarrow \mathbb{R}^d$, with L^2 derivative. For $x, y \in \mathbb{R}^d$, $H_{[0,t]}^x$ denotes the affine subspace of paths starting at x, and $H_{[0,t]}^{x,y}$ the space of paths with endpoints x, y. The space $H_{[0,t]}^{0,0}$ is endowed with the scalar product

$$\langle \gamma, \xi \rangle = \int_0^t \langle \dot{\gamma}_u, \dot{\xi}_u \rangle du.$$

We denote $C_{[0,t]}$ the space of continuous paths $[0,t] \longrightarrow \mathbb{R}^d$. The topology is that of uniform convergence; $C_{[0,t]}^x$ and $C_{[0,t]}^{x,y}$ are, respectively, the affine subspaces of paths starting at x, and with endpoints x, y.

$C_{\mathbb{R}} = C(\mathbb{R}, \mathbb{R}^d)$ is the space of continuous paths from $\mathbb{R}$ to $\mathbb{R}^d$, endowed with the topology of uniform convergence on compact intervals.

We let $\mathbb{Z}^d$ act on $C(\mathbb{R}, \mathbb{R}^d)$ or $C_{[0,t]}$ by

$$(s.\gamma)(u) = \gamma(u) + s$$

for all $u \in \mathbb{R}$, $s \in \mathbb{Z}^d$, $\gamma \in C(\mathbb{R}, \mathbb{R}^d)$. The quotient space $C(\mathbb{R}, \mathbb{R}^d)/\mathbb{Z}^d$ (respectively $C_{[0,t]}/\mathbb{Z}^d$) is naturally identified with the space of continuous paths $\mathbb{R} \longrightarrow \mathbb{T}^d$ (respectively $[0,t] \longrightarrow \mathbb{T}^d$), and denoted W (resp. $W_{[0,t]}$).

There is also a natural action of $\mathbb{R}$ by translation of time on $C(\mathbb{R}, \mathbb{R}^d)$ or on W:

$$\sigma^t(\gamma)(u) = \gamma(u+t)$$

for $\gamma \in C(\mathbb{R}, \mathbb{R}^d)$ or W.

Later, we will be interested in measures on $C(\mathbb{R}, \mathbb{R}^d)$ or W. When we speak of measures, let us precise that the Borel σ-field on path spaces is the smallest for which all the maps $\gamma \mapsto \gamma_u$ are measurable.

Measures on W will be naturally identified with measures on $C(\mathbb{R}, \mathbb{R}^d)$, invariant under the action of $\mathbb{Z}^d$.

The space $W_{[0,t]}$ can be endowed with the Wiener measure starting at x, a probability measure denoted $\mathcal{W}_{[0,t]}^{\beta,x}$ and carried on $W_{[0,t]}^x$. The brownian bridge $\mathcal{W}_{[0,t]}^{\beta,(x,y)}$ is a positive measure carried on $W_{[0,t]}^{x,y}$, and whose definition is recalled thereafter. The parameter $\beta > 0$ is the inverse of the diffusion coefficient.

We refer to [Si79], II.4,5 for the construction of Wiener processes. For $x, y \in \mathbb{R}^d$, the brownian bridge $\mathcal{W}_{[t_0,t_n]}^{\beta,(x,y)}$ with diffusion coefficient $1/\beta$, starting at x and ending at y, in the time interval $[t_0, t_n]$ is defined as the unique positive measure on $C([t_0, t_n], \mathbb{R}^d)$ such that

$$\mathcal{W}^{\beta,(x,y)}_{[t_0,t_n]}(B) \quad = \quad \int_{z_1 \in B_1, \ldots, z_{n-1} \in B_{n-1}} \prod_{i=1}^{n} \frac{e^{-\frac{\beta||z_{i+1}-z_i||^2}{2(t_{i+1}-t_i)}}}{(2\pi(t_{i+1}-t_i)/\beta)^{d/2}} dz_i$$

for all $t_0 < t_1 < \cdots < t_n$, for all $B \subset C([t_0, t_n], \mathbb{R}^d)$ of the form $B = \{\gamma, \gamma_{t_i} \in B_i, \forall i = 1, ..., n-1\}$, where the B_i's are measurable subsets of $\mathbb{R}^d$.

The Wiener measure started at x, $\mathcal{W}^{\beta,x}_{[0,t]}$, is related to the brownian bridges $\mathcal{W}^{\beta,(x,y)}_{[0,t]}$ by

$$\mathcal{W}^{\beta,x}_{[0,t]}(B) = \int \mathcal{W}^{\beta,(x,y)}_{[0,t]}(B) dy$$

for every measurable B.

We now state Assumptions (A1), (A2), (A3).

Assumptions : Recall that we have defined the action of a path $\gamma : [0, t] \longrightarrow \mathbb{R}^d$ as

$$\mathcal{A}(\gamma) = \int_0^t \mathcal{L}(\gamma_s, \dot{\gamma}_s) ds.$$

Assumption (A1) For all n, for all endpoints $\xi_0, \xi_n \in \mathbb{R}^d$, the minima of the action $\mathcal{A}$ in the space $H^{\xi_0,\xi_n}_{[0,n]}$ are non-degenerate (we mean thereby that the hessian of $\mathcal{A}$ at each minimum is invertible as an operator in the Hilbert space $H^{0,0}_{[0,n]}$). Besides, the number of minimizers is bounded, independently of n, ξ_0, ξ_n.

Assumption (A2) There exists $\varepsilon_0 > 0$ such that, for all $0 \leq \varepsilon \leq \varepsilon_0$, there exists a sequence $(c_n) \in [0, 1]^{\mathbb{N}}$ satisfying:

– $\lim_n \frac{\log c_n}{n} = 0$,

and :

– for all $n > 0$, for all $\gamma_0, \gamma_n \in \mathbb{R}^d$ such that

$$\parallel \gamma_0 - \xi_0 \parallel \leq c_n \varepsilon$$

$$\parallel \gamma_n - \xi_n \parallel \leq c_n \varepsilon$$

for some trajectory ξ in the Mather set (lifted to $\mathbb{R}^d$), then there exists a minimizer γ of $\mathcal{A} : H^{\xi_0,\xi_n}_{[0,n]} \longrightarrow \mathbb{R}$ such that $\parallel \gamma_k - \xi_k \parallel \leq \varepsilon$ for all $0 \leq k \leq n$.

Change of gage: It follows from the weak KAM theorem 5.0.3 that there exists a $\mathbb{Z}^d$-periodic, Lipschitz function u, and a constant c, such that the functional

$$\gamma \mapsto \mathcal{A}(\gamma) - u(\gamma_t) + u(\gamma_0) + ct$$

is nonnegative on $H_{[0,t]}$ for all t, and vanishes if γ is a portion of a trajectory in the Mather set. In all the definitions given above, we can replace the action $\mathcal{A}$ by this new functional, without changing the definition of Euler-Lagrange flow, Mather set, etc... The fact that u is not smooth is not really a problem, since we only need to differentiate the action functional with respect to variations of the path leaving endpoints fixed. As far as Schrödinger operators are concerned, this change in the choice of the action functional would amount to replacing the kernel $K^t_{\beta,\omega}(x, y)$ by $e^{-\beta ct} e^{-\beta u(x)} K^t_{\beta,\omega}(x, y) e^{\beta u(y)}$, which would lead to replacing $\psi_\beta(x)$ by $\psi_\beta(x)e^{-\beta u(x)}$ and $\psi^*_\beta(x)$ by $\psi^*_\beta(x)e^{-\beta u(x)}$. Eventually, the measure μ^0_β would remain unchanged. *In the rest of this part, we renormalize the action functional so that it is nonnegative, and vanishes on the Mather set.*

After this modification of the action, we introduce the function

$$h_n(x,y) = \inf_{H^{x,y}_{[0,n]}} \mathcal{A},$$

defined on $\mathbb{R}^d \times \mathbb{R}^d$.

Assumption (A3) There exists a sequence $B_n \geq 0$ satisfying $\lim_n \frac{\log B_n}{n} = 0$, such that, for all n,

$$sup_{\gamma_0} \beta^{d/2} \int_{\mathbb{R}^d} e^{-\beta h_n(\gamma_0,\gamma_n)} d\gamma_n \leq B_n.$$

The non-degeneracy of minimima of the action is necessary for the Laplace method (Section 6); the second part of (A1), about the number of minimizers, seems less crucial. Assumptions (A2), (A3) are not very easy to interprete, although we check in 6.4 that they are always satisfied when $\omega = 0$. (A3) says something about the behaviour of h_n near its minima, *uniformly in n*. We formulate a conjecture about a different assumption under which the theorem could hold:

Conjecture: These assumptions can be replaced by the assumption that the Mather set is a uniformly hyperbolic set for the action of the Euler-Lagrange flow.

Of course, this assumption is not easier to check on examples than the previous ones, but it is more conceptual. A proof seems close at hand for discrete time systems (i.e the situation of twist maps described in Section 4), however not for continuous time systems.

In paragraph 6.4, we show that (A2) and (A3) are always satisfied in the case $\omega = 0$. In this case, the theorem seems to be part of the folklore in the study of the tunnelling effect in semi-classical mechanics:

Corollary 5.0.10. *Let $\mathcal{H}_\hbar = \hbar^2 \frac{\Delta}{2} + V$, and let $\psi_\hbar$ be the unique $\mathbb{Z}^d$-periodic positive eigenfunction, corresponding to the largest eigenvalue of $\mathcal{H}_\hbar$ in $L^2(\mathbb{T}^d)$.*

Then, as $\hbar \longrightarrow 0$, the probability measure

$$\frac{\psi_\hbar^2(x)dx}{\int_{\mathbb{T}^d} \psi_\hbar^2(y)dy}$$

concentrates on the maxima of V.

Assume furthermore that the system satisfies (A1). If we consider the expansion of V in orthonormal coordinates near a maximum x_0, in the form

$$V(x_0+y) = V(x_0) - \frac{1}{2}\sum |a_i(x_0)|^2 y_i^2 + O(y^3),$$

then the measure $\psi_\hbar^2(x)dx$ concentrates on those x_0s for which the quantity

$$\sum |a_i(x_0)|$$

is the smallest.

Note that the maxima are not assumed to be non-degenerate nor isolated. The same result (assuming non-degenerate maxima) is contained, but hidden amongst deeper theorems, in paragraph 4.4 of [He88], where $\sum |a_i(x_0)|$ appears as the bottom of the spectrum of the quantum harmonic oscillator:

$$-\hbar^2 \frac{\Delta}{2} + \frac{1}{2} \sum |a_i(x_0)|^2 y_i^2,$$

and the result is obtained by BKW estimates. Corollary 5.0.10 is in agreement with the results therein, if we change the sign in front of the laplacian and replace the word "maximum" by "minimum".

As in Part I, Theorem 0.0.2 (=5.0.8) is proved by associating to the Schrödinger equation (5.0.3), (5.0.4) a stationary stochastic process of initial distribution μ^0_β. This process is the Markov process of initial distribution μ^0_β and with transition semi-group:

$$f \mapsto P^t_{\beta,\omega} f = \frac{1}{e^{t\beta\lambda_\beta}\psi_\beta} \exp t\beta\mathcal{H}^\omega_\beta.(\psi_\beta f).$$

Since ψ_β and ψ^*_β satisfy (5.0.1), (5.0.2), μ^0_β is indeed the invariant $\mathbb{Z}^d$-periodic distribution.

The process corresponds to a measure μ_β on the path space $C(\mathbb{R}, \mathbb{R}^d)$, defined by the formula

$$\begin{aligned} &\mu_\beta(\{\gamma, \gamma_{|[0,t]} \in K\}) \\ &= \frac{e^{-t\beta\lambda_\beta}}{\int_{\mathbb{T}^d} \psi_\beta \psi^*_\beta} \int_{\gamma_0 \in \mathbb{T}^d} \psi^*_\beta(\gamma_0) d\gamma_0 \left(\int_{\gamma_{|[0,t]} \in K} e^{\beta \int_0^t V(\gamma_s) ds + \beta\langle \omega, \gamma_t - \gamma_0 \rangle} \psi_\beta(\gamma_t) d\mathcal{W}^{\beta,\gamma_0}_{[0,t]}(\gamma) \right) \end{aligned}$$

when K is a measurable subset of $C_{[0,t]}$.

Since μ^0_β is invariant under the transition semi-group, the measure μ_β is invariant by the translations of time $(\sigma^t)_{t\in\mathbb{R}}$, as well as by the action of $\mathbb{Z}^d$ on $C(\mathbb{R}, \mathbb{R}^d)$. Thus, it defines a probability measure μ_β on the quotient

$$C(\mathbb{R}, \mathbb{R}^d)/\mathbb{Z}^d = C(\mathbb{R}, \mathbb{T}^d) =: W.$$

6. Proof of Theorem 0.0.2

6.1. **Preliminary results.**

Lemma 6.1.1. *For each sequence $\beta_n \longrightarrow +\infty$, there exists a subsequence $\beta_{n(k)}$ and a probability measure μ_∞ on W, such that, for all $l \in \mathbb{N}$, for all $t_1 < t_2 ... < t_l$,*

$$\mu_{\beta_{n(k)}}(g(\gamma_{t_1}, ..., \gamma_{t_l})) \underset{k \longrightarrow \infty}{\longrightarrow} \mu_\infty(g(\gamma_{t_1}, ..., \gamma_{t_l})),$$

for every bounded continuous function g on $(\mathbb{R}^d)^l/\mathbb{Z}^d$.

In this case, we shall say that the sequence $(\mu_{\beta_{n(k)}})$ converges to μ_∞.

Proof. Fix $T > 0$. To get rid of some constants, assume that $\| \omega \| \leq 1$ and $| V | \leq 1$. We denote $\mu_\beta(.|\gamma_0 = x)$ the measure μ_β conditioned with respect to γ_0.

For all $0 < t \leq T$, for all $x \in \mathbb{R}^d$,

$$\mu_\beta(\| \gamma_t - \gamma_0 \|_\infty \geq 4t\, |\gamma_0 = x) = \frac{\int_{C_{[0,t]}} \mathbf{1}_{\{\|\gamma_t - x\|_\infty \geq 4t\}} e^{\beta \int_0^t V(\gamma_s)ds + \beta\langle \omega, \gamma_t - x\rangle} d\mathcal{W}^{\beta,x}_{[0,t]}(\gamma)}{\int_{C_{[0,t]}} e^{\beta \int_0^t V(\gamma_s)ds + \beta\langle \omega, \gamma_t - x\rangle} d\mathcal{W}^{\beta,x}_{[0,t]}(\gamma)}$$

$$\leq \frac{\int_{W_{[0,t]}} \mathbf{1}_{\{\|\gamma_t - x\|_\infty \geq 4t\}} e^{\beta(t + \|\gamma_t - x\|)} d\mathcal{W}^{\beta,x}_{[0,t]}(\gamma)}{\int_{W_{[0,t]}} e^{-\beta(t + \|\gamma_t - x\|)} d\mathcal{W}^{\beta,x}_{[0,t]}(\gamma)}$$

$$= \frac{e^{2\beta t} \int_{\mathbb{R}^d} \mathbf{1}_{\{\|y\|_\infty \geq 4t\}} e^{\beta\|y\| - \beta \frac{\|y\|^2}{2t}} dy}{\int_{\mathbb{R}^d} e^{-\beta\|y\| - \beta\frac{\|y\|^2}{2t}} dy}$$

$$\leq \frac{e^{2\beta t} \int_{\mathbb{R}^d} \mathbf{1}_{\{\|y\|_\infty \geq 4t\}} e^{-\beta \frac{\|y\|^2}{4t}} dy}{\int_{\mathbb{R}^d} e^{-\beta(\|y\| + \frac{\|y\|^2}{2t})} dy} \lesssim C(t) e^{-6\beta t} \beta^{(d-1)/2}$$

for all $t > 0$, for β large enough; we have used the following estimate on the tail of a gaussian distribution on $\mathbb{R}^d$:

$$\frac{1}{(2\pi)^{d/2}} \int_{\|y\|_\infty \geq \delta} e^{-\frac{\|y\|^2}{2}} dy \leq 2d \frac{e^{-\frac{\delta^2}{2}}}{\delta}$$

as well as the fact that

$$\int_{\mathbb{R}^d} e^{-\beta(\|y\| + \frac{\|y\|^2}{2t})} dy = \frac{1}{\beta^d} \int_{\mathbb{R}^d} e^{-(\|y\| + \frac{\|y\|^2}{2\beta t})} dy \sim \frac{C}{\beta^d}$$

as $\beta \longrightarrow +\infty$. As a consequence, for all $0 \leq s < t \leq T$,

$$\mu_\beta(\| \gamma_t - \gamma_s \|_\infty \geq 4(t-s)\ |) \lesssim C(t-s) e^{-6\beta(t-s)} \beta^{(d-1)/2} \tag{6.1.1}$$

This implies in particular the tightness of the laws of γ_t under $(\mu_\beta)_{\beta > 0}$, for all t. Thus, we can find a subsequence $\beta_{n(k)} \longrightarrow +\infty$, and a probability measure μ_∞ on $W_{[0,T]\cap\mathbb{Q}}$, such that, for all $t_1 < ... < t_l \in [0,T] \cap \mathbb{Q}$,

$$\mu_{\beta_{n(k)}}(g(\gamma_{t_1}, ..., \gamma_{t_l})) \underset{k \longrightarrow +\infty}{\longrightarrow} \mu_\infty(g(\gamma_{t_1}, ..., \gamma_{t_l}))$$

for every bounded continuous function g on $(\mathbb{R}^d)^l/\mathbb{Z}^d$.

But actually, thanks to inequality (6.1.1), the convergence will take place for all $t_1 < ... < t_l \in [0,T]$, and every bounded continuous function g on $(\mathbb{R}^d)^l/\mathbb{Z}^d$. □

Proposition 6.1.2. *(a) Let ψ_β, ψ^*_β satisfiy (5.0.1), (5.0.2). Then the families of functions $(-\frac{1}{\beta}\log\psi_\beta)_{\beta>0}, (-\frac{1}{\beta}\log\psi^*_\beta)_{\beta>0}$ are equilipschitz.*

(b) If $\beta_k \longrightarrow +\infty$ is a sequence such that

$$-\frac{1}{\beta_k}\log\psi_{\beta_k} \longrightarrow -u_+$$

and

$$-\frac{1}{\beta_k}\log\psi^*_{\beta_k} \longrightarrow v_-$$

in the uniform topology, for some continuous functions u_+ and v_-, then $T^+_t u_+ = u_+ - c(\omega)t$ and $T^-_t v_- = v_- + c(\omega)t$, for all $t \geq 0$.

(c) Let $J = \inf(v_- - u_+)$, so that

$$-\frac{\log\psi_{\beta_k} + \log\psi^*_{\beta_k}}{\beta_k} + \frac{\log\int\psi_{\beta_k}(y)\psi^*_{\beta_k}(y)dy}{\beta_k} \longrightarrow v_- - u_+ - J,$$

and let u_- be the fixed point of $(T_t^- - c(\omega)t)$ which is conjugate to u_+. Then $u_- \le v_- - J$.

Proof. Assertion (a) was proved in Section 5 (remember that $-\frac{1}{\beta}\log\psi_\beta$ and $-\frac{1}{\beta}\log\psi^*_\beta$ are solutions of viscous Hamilton-Jacobi equations).

Assertion (b) follows from the stability result for viscosity solutions ([CEL84], Theorem 3.1).

The fact that

$$\frac{\log\int\psi_{\beta_k}(y)\psi^*_{\beta_k}(y)dy}{\beta_k} \longrightarrow -J$$

follows from Lemma 3.1.4. As to the last assertion, it is a consequence of the inequality $v_- - u_+ - J \ge 0$, and the characterization of u_- as the smallest fixed point of $(T_t^- - c(\omega_t)$ satisfying $u_- - u_+ \ge 0$ ([Fa97-2]). □

Proposition 6.1.3. *(Large deviation upper bound) Let $t \ge 0$. Then for any subset $K_t \subset W_{[0,t]}$, closed for the uniform topology,*

$$\limsup \frac{1}{\beta}\log\mu_\beta(\{\gamma, \gamma_{|[0,t]} \in K_t\}) \le -\inf_{\gamma\in K_t}\inf_{(u_-,u_+)} u_-(\gamma_0)+\mathcal{A}(\gamma_{|[0,t]})-u_+(\gamma_t)+tc(\omega)$$

where the second inf *is taken over the set of conjugate fixed points of the Hopf-Lax semi-groups.*

Corollary 6.1.4. *If μ_∞ is a limit point of μ_β, it is carried by action-minimizing trajectories of the Euler-Lagrange flow.*

Proof. (Proposition 6.1.3) Recall the expression of $\mu_\beta(\{\gamma,\gamma_{|[0,t]} \in K_t\})$, for a measurable $K_t \subset W_{[0,t]}$:

$$(6.1.2)\quad \mu_\beta(\{\gamma, \gamma_{|[0,t]} \in K_t\})$$
$$= \frac{e^{-t\beta\lambda_\beta}}{\int_{\mathbb{T}^d}\psi_\beta\psi^*_\beta}\int_{\gamma_0\in\mathbb{T}^d}\psi^*_\beta(\gamma_0)d\gamma_0\left(\int_{\gamma_{|[0,t]}\in K_t} e^{\beta\int_0^t V(\gamma_s)ds+\beta\langle\omega,\gamma_t-\gamma_0\rangle}\psi_\beta(\gamma_t)d\mathcal{W}^{\beta,\gamma_0}_{[0,t]}(\gamma)\right)$$

We have seen that $\lambda_\beta \underset{\beta\longrightarrow+\infty}{\longrightarrow} c(\omega)$. We also recall that, if K_t is closed, then, for all $x \in \mathbb{T}^d$,

$$(6.1.3)\quad \limsup\frac{1}{\beta}\log\int_{\gamma_{|[0,t]}\in K_t} e^{\beta\int_0^t V(\gamma_s)ds+\beta\langle\omega,\gamma_t-\gamma_0\rangle}e^{-\beta u(\gamma_t)}d\mathcal{W}^{\beta,x}_{[0,t]}(\gamma)$$
$$\le -\inf_{\gamma_{|[0,t]}\in K_t,\gamma_0=x}\mathcal{A}(\gamma_{|[0,t]})+u(\gamma_t)$$

for every continuous function u on $\mathbb{T}^d$. This follows from an application of the large deviation result of Schilder ([Sc66], [DZ92]: Theorem 5.2.3) combined with “Varadhan’s lemma” ([Va67], [DZ92]: Theorem 4.3.1).

Finally, let us consider a sequence $\beta_k \to +\infty$ such that $\frac{1}{\beta_k}\log\mu_{\beta_k}(\{\gamma,\gamma_{|[0,t]} \in K_t\})$ converges in $\mathbb{R}\cup\{-\infty\}$. Keeping the notations of Proposition 6.1.2, we may also assume (after further extractions) that

$$-\frac{1}{\beta_k}\log\psi_{\beta_k} \longrightarrow -u_+$$

$$-\frac{1}{\beta_k}\log\psi^*_{\beta_k} \longrightarrow v_-$$

and

$$\frac{1}{\beta_k} \log \int_{\mathbb{T}^d} \psi_\beta \psi_\beta^* \longrightarrow -J,$$

with $v_- - J$ larger than the function u_- conjugate to u_+.

Combining this with (6.1.2), we get

$$\begin{aligned}
\limsup \frac{1}{\beta_k} \log \mu_{\beta_k}(\{\gamma, \gamma_{|[0,t]} \in K_t\}) &\leq -\inf_{\gamma_{|[0,t]} \in K_t} v_-(\gamma_0) + \mathcal{A}(\gamma_{|[0,t]}) - u_+(\gamma_t) + tc(\omega) - J \\
&\leq -\inf_{\gamma_{|[0,t]} \in K_t} u_-(\gamma_0) + \mathcal{A}(\gamma_{|[0,t]}) - u_+(\gamma_t) + tc(\omega) \\
&\leq -\inf_{\gamma_{|[0,t]} \in K_t} \inf_{(u_-, u_+)} u_-(\gamma_0) + \mathcal{A}(\gamma_{|[0,t]}) - u_+(\gamma_t) + tc(\omega)
\end{aligned}$$

Since this is true for every subsequence β_k, we have proved Proposition 6.1.3. □

Proof. (Corollary 6.1.4) Let K be a closed subset of W, and K_t its projection on $W_{[0,t]}$. Then

$$\mu_\beta(K) \leq \mu_\beta(\{\gamma, \gamma_{|[0,t]} \in K_t\})$$

and, after Proposition 6.1.3, the measure of K will go to zero exponentially fast unless

$$\inf_{\gamma \in K_t} \inf_{(u_-, u_+)} u_-(\gamma_0) + \mathcal{A}(\gamma_{|[0,t]}) - u_+(\gamma_t) + tc(\omega) = 0,$$

for all t. But, for a path $\gamma_{|[0,t]}$,

$$u_-(\gamma_0) + \mathcal{A}(\gamma_{|[0,t]}) - u_+(\gamma_t) + tc(\omega) = 0$$

implies that

$$\mathcal{A}(\xi_{|[0,t]}) \geq \mathcal{A}(\gamma_{|[0,t]})$$

for every path ξ such that $\xi_0 = \gamma_0$ and $\xi_t = \gamma_t$.

We have thus proved that μ_∞ is carried by action-minimizing trajectories of the Euler-Lagrange flow. □

The measure μ_∞, as a measure on W carried by action-minimizing trajectories and invariant by translation of time $(\sigma^t)_{t\in\mathbb{R}}$, can be naturally identified to an action-minimizing measure on $\mathbb{T}^d \times \mathbb{R}^d$. The identification takes place via the map

$$\begin{array}{rcl}
\{\gamma : \mathbb{R} \longrightarrow \mathbb{T}^d,\ \text{trajectories of the E-L flow}\} & \longrightarrow & \mathbb{T}^d \times \mathbb{R}^d \\
\gamma & \mapsto & (\gamma_0, \dot\gamma_0)
\end{array}$$

which is a conjugacy between the translation of time (σ^t) and the Euler-Lagrange flow.

In particular, the measure μ_∞^0 on $\mathbb{T}^d$, defined by

$$\mu_\infty^0(A) = \mu_\infty\{\gamma \in W, \gamma_0 \in A\}$$

is carried by the set $\mathcal{M}_\omega$. We have thus proved the first assertion of Theorem 5.0.8: the measure μ_β^0 concentrates on $\mathcal{M}_\omega$.

6.2. **Sketch of proof. Definition of the hessian of the action, and of its determinant.**

Let $x, y \in \mathbb{R}^d$. The action $\mathcal{A} : H^{x,y}_{[0,t]} \longrightarrow \mathbb{R}$ is twice differentiable, and its second derivative at a point γ, $d^2\mathcal{A}(\gamma)$, is a symmetric bilinear form on $H^{0,0}_{[0,t]}$; one may write it as

$$d^2\mathcal{A}(\gamma).\xi.\xi = \langle \mathcal{A}''(\gamma)\xi, \xi\rangle$$

where $\mathcal{A}''(\gamma)$ is a self-adjoint operator on H: the hessian of $\mathcal{A}$ at γ.

Remembering the expression of $\mathcal{A}$, one can actually write

$$\mathcal{A}''(\gamma) = I + f,$$

f being defined by

$$\langle f\xi, \xi\rangle = \int_0^t V''(\gamma_s).\xi_s.\xi_s ds$$

This last bilinear form may be extended to a continuous symmetric bilinear form on W; and this implies that f is a trace operator ([Ku75], p.83): the sum of the eigenvalues of f, $(\lambda_i)_{i\in\mathbb{N}}$, is absolutely convergent.

Thus, one may define the determinant of $I+f$ as $[I+f] := \prod_{i\in\mathbb{N}}(1+\lambda_i)$, which is well defined (possibly zero). This determinant will be non zero if and only if -1 is not an eigenvalue of f, that is, if and only if the operator $\mathcal{A}''(\gamma)$ is invertible in $H^{0,0}_{[0,t]}$.

If γ is a critical point of $\mathcal{A} : H^{x,y}_{[0,t]} \longrightarrow \mathbb{R}$ such that $\mathcal{A}''(\gamma)$ is invertible, we will say that γ is a non-degenerate critical point of $\mathcal{A} : H^{x,y}_{[0,t]} \longrightarrow \mathbb{R}$.

Like in Part I, if $\gamma \in H_{[0,t']}$ for some $t' \geq t$, we will denote $[\ {}_tA''(\gamma)]$ the determinant of the hessian of $A(\gamma_{|[0,t]}) : H^{\gamma_0,\gamma_t}_{[0,t]} \longrightarrow \mathbb{R}$, at γ.

We prove the following proposition, which is the direct transposition of Theorem 0.0.1 in continuous time:

Proposition 6.2.1. *Let μ_∞ be a limit point of $(\mu_\beta)_{\beta\longrightarrow+\infty}$, and let μ be a σ-invariant probability measure on W, carried by action-minimizing trajectories of the Euler-Lagrange flow. Then, under assumption (A1), (A2) and (A3),*

$$h_\sigma(\mu) - \frac{1}{2}\int \lim_t \frac{1}{t}\log[{}_t\mathcal{A}''(\gamma)]d\mu(\gamma) \leq h_\sigma(\mu_\infty) - \frac{1}{2}\int \lim_t \frac{1}{t}\log[{}_t\mathcal{A}''(\gamma)]d\mu_\infty(\gamma)$$

The proof of Proposition 6.2.1 goes along the same lines as that of Theorem 0.0.1. The main difference is a higher degree of technicality in the writing of the Laplace method for estimating path integrals. We do not rewrite the proof in its entirety, but indicate how to adapt Paragraph 3.2 to the new situation.

To simplify notations we consider again the case $d = 1$.

The proof starts, as in Paragraph 3.2, with the construction of a partition $\tilde{P}$ of $C([0,1],\mathbb{R})$:

$$\tilde{P}_{ij} = \{\gamma \in C([0,1],\mathbb{R}), \gamma_0 \in [i\varepsilon,(i+1)\varepsilon), \gamma_1 \in [j\varepsilon,(j+1)\varepsilon)\}$$

for $|\ j-i\ | < \frac{M}{\varepsilon}$, and

$$\tilde{P}_{i\infty} = \{\gamma, \gamma_0 \in [i\varepsilon,(i+1)\varepsilon), \exists j, |\ j-i\ | \geq \frac{M}{\varepsilon}, \gamma_1 \in [j\varepsilon,(j+1)\varepsilon)\}$$

If ε is the inverse of an integer, the partition goes to the quotient $W_{[0,1]} = C([0,1],\mathbb{R})/\mathbb{Z}$, and gives a finite partition P of W.

The proof is then identical to that of Theorem 0.0.1 until the statement of Lemma 3.2.3 (a). The integrands

$$\left(\frac{\beta}{2\pi}\right)^{\frac{n-1}{2}} e^{-\beta \sum_{i=0}^{n-1} L(\gamma_i,\gamma_{i+1})} d\gamma_1..d\gamma_{n-1}$$

just need to be replaced by

$$e^{\beta \int_0^n V(\gamma_s)ds+\beta\langle\omega,\gamma_t-\gamma_0\rangle} d\mathcal{W}^{\beta,(\gamma_0,\gamma_n)}_{[0,n]}(\gamma).$$

We now indicate how to adapt Lemma 3.2.3.

Laplace method (fixed time interval). Lemma 3.2.3 (b), which is a consequence of the superlinear growth of the Lagrangian, can be obtained from the estimate

$$K^t_{\beta,\omega}(x,y) \leq e^{\beta M t}\left(\frac{\beta}{2\pi t}\right)^{d/2} e^{-\frac{\beta||x-y||^2}{2t}}.e^{\beta||\omega||.||x-y||},$$

mentioned in Section 5.

As to Lemma 3.2.3 (a), it comes from the following:

Theorem 6.2.2. *(*[Be88], [BDS93]*) Let $\gamma_0,\gamma_t \in \mathbb{R}^2$. Assume that the action $\mathcal{A} : H^{\gamma_0,\gamma_t}_{[0,t]} \longrightarrow \mathbb{R}$ has only one minimum $\hat\gamma$, which is non degenerate, and let Ω be a neighbourhood of $\hat\gamma$ in the uniform topology. Then*

$$\int_{W^{\gamma_0,y\gamma_t}_{[0,t]}\cap\Omega} e^{\beta(\int_0^t V(\gamma_s)ds+\langle\omega,\gamma_t-\gamma_0\rangle)} d\mathcal{W}^{\beta,(\gamma_0,\gamma_t)}_{[0,t]}(\gamma) = \frac{e^{-\beta\mathcal{A}(\hat\gamma)}}{[{}_t\mathcal{A}''(\hat\gamma)]^{1/2}}(1+\underset{\beta\to\infty}{o(1)}).$$

For fixed t and K, the $\underset{\beta\to\infty}{o(1)}$ is uniform on the set $\{|\gamma_t-\gamma_0|\leq K\}$.

Let us give a general idea of how this estimate may be obtained (the reader is referred to [Be88], [BDS93] for a complete proof). Exactly as in the case of an integral over a finite dimensional space, one begins by applying a Taylor expansion of order 2 of the function:

$$\gamma \mapsto \int_0^t V(\gamma_s)ds + \langle\omega,\gamma_t-\gamma_0\rangle$$

at the minimizer of the action $\hat\gamma$, and in the space $W^{\gamma_0,\gamma_t}_{[0,t]}$

$$(6.2.1)\quad \int_{W^{\gamma_0,\gamma_t}_{[0,t]}\cap\Omega} e^{\beta\int_0^t V(\gamma_s)ds} d\mathcal{W}^{\beta,\gamma_0,\gamma_t}_{[0,t]}(\gamma) =$$

$$e^{\beta\int_0^t V(\hat\gamma_s)ds}\int_{W^{\gamma_0,\gamma_t}_{[0,t]}\cap\Omega} e^{\beta\int_0^t V'_{\hat\gamma_s}.(\gamma_s-\hat\gamma_s)ds+\frac{\beta}{2}\int_0^t V''_{\hat\gamma_s}.(\gamma_s-\hat\gamma_s)^2ds+\beta R(\gamma-\hat\gamma)} d\mathcal{W}^{\beta,\gamma_0,\gamma_t}_{[0,t]}(\gamma) =$$

$$e^{\beta\int_0^t V(\hat\gamma_s)ds}\int_{W^{\gamma_0,\gamma_t}_{[0,t]}\cap\Omega} e^{\beta\langle\hat\gamma_s,\gamma_s-\hat\gamma_s\rangle+\frac{\beta}{2}\int_0^t V''_{\hat\gamma_s}.(\gamma_s-\hat\gamma_s)^2ds+\beta R(\gamma-\hat\gamma)} d\mathcal{W}^{\beta,\gamma_0,\gamma_t}_{[0,t]}(\gamma)$$

$$= e^{-\beta\frac{\int_0^t\dot{\hat\gamma}^2}{2}+\beta\int_0^t V(\hat\gamma_s)ds}\int_{W^{\gamma_0,\gamma_t}_{[0,t]}\cap\Omega-\hat\gamma} e^{\frac{\beta}{2}\int_0^t V''_{\hat\gamma_s}.\gamma_s^2ds+\beta R(\gamma)} d\mathcal{W}^{\beta,(0,0)}_{[0,t]}(\gamma)$$

where the last line is obtained by the Cameron-Martin formula ([Ku75], p.111), and the line before comes from the fact that $\hat\gamma$ is a critical point of the action.

The remainder $R(\gamma)$, given by Taylor's integral formula, is bounded (independently of n) by $C \parallel \gamma \parallel_3^3$, where C is a bound on the third derivative of V; and actually, if Ω is a uniform neighbourhood of radius ε around $\hat\gamma$, $R(\gamma)$ is bounded by by $C\varepsilon \parallel \gamma \parallel_2^2$. One shows that this remainder does not interfer in the estimate of Theorem 6.2.2.

The final ingredient is the formula

$$\int_{W^{0,0}_{[0,t]}} e^{-\beta\langle f\gamma,\gamma\rangle} d\mathcal{W}^{\beta,0,0}_{[0,t]}(\gamma) = [I+f]^{-1/2}$$

valid as soon as $\langle f.,.\rangle$ is a continuous symmetric bilinear form on $H^{0,0}_{[0,t]}$ which admits a continuous extension to $W^{0,0}_{[0,t]}$. It is obtained by diagonalization of f in an orthonormal basis for $\langle .,.\rangle$. It yields Theorem 6.2.2 when applied to

$$\langle f\gamma,\gamma\rangle = \int_0^t V''_{\hat{\gamma}_s}.\gamma_s^2 ds.$$

Lemma 3.2.4 is unchanged, with the necessary modifications in the expression of path integrals, as explained earlier. We now need to adapt Lemma 3.2.8.

Laplace method (lower bound, independent of the time interval).

In order to adapt the result of Lemma 3.2.8, the point is to find a lower bound of (6.2.1), independent of t. We can take $t = n \in \mathbb{N}$. As before, let $\hat{\gamma}$ be the minimizer of the action on $H^{\gamma_0,\gamma_n}_{[0,n]}$. Lemma 3.2.8 is replaced by:

Lemma 6.2.3. *Let $\varepsilon > 0$, and $\delta \le \varepsilon$. Then there exists $\rho = \rho(\varepsilon)$ such that, for all n,*

$$(6.2.2)\quad \int_{||\gamma||_\infty \le \delta} e^{\frac{\beta}{2}\int_0^n V''_{\hat{\gamma}_s}.\gamma_s^2 + \beta R(\gamma)} d\mathcal{W}^{\beta,(0,0)}_{[0,n]}$$

$$\ge (1-e^{-\beta\rho^2\frac{\delta^2}{2}})^{n-1}(1+\underset{\beta}{o}(1))^{n-1}\int_{W_{[0,n]}} e^{\frac{\beta}{2}\int_0^n V''_{\hat{\gamma}_s}.\gamma_s^2 - \beta C\varepsilon\int_0^n|\gamma_s|^2 ds - \beta\varepsilon\sum_{j=1}^{n-1}|\gamma_j|^2} d\mathcal{W}^{\beta,(0,0)}_{[0,n]}$$

$$= (1-e^{-\beta\rho^2\frac{\delta^2}{2}})^{n-1}(1+\underset{\beta}{o}(1))^{n-1}\frac{1}{[\,_n\mathcal{A}''(\hat{\gamma}) + 2\varepsilon b_n]^{1/2}}$$

where b_n is the quadratic form on $H^{(0,0)}_{[0,n]}$:

$$b_n(\gamma,\gamma) = C\int_0^n |\gamma_s|^2 ds + \sum_{j=1}^{n-1}|\gamma_j|^2,$$

and C is an upper bound on the third derivative of V.

Proof. We know that $|R(\gamma)| \le C||\gamma||_3^3 \le C\varepsilon||\gamma||_2^2 \le C\varepsilon||\gamma||_2^2 + \varepsilon\sum_{j=1}^{n-1}|\gamma_j|^2 = \varepsilon b_n(\gamma,\gamma)$. Thus,

$$(6.2.3)\quad \int_{||\gamma||_\infty \le \delta} e^{\frac{\beta}{2}\int_0^n V''_{\hat{\gamma}_s}.\gamma_s^2 + \beta R(\gamma)} d\mathcal{W}^{\beta,0,0}_{[0,n]}$$

$$\ge \int_{||\gamma||_\infty \le \delta} e^{\frac{\beta}{2}\int_0^n V''_{\hat{\gamma}_s}.\gamma_s^2 - \varepsilon\beta b_n(\gamma,\gamma)} d\mathcal{W}^{\beta,0,0}_{[0,n]}$$

Let us consider the action (associated to a non-autonomous lagrangian):

$$\tilde{\mathcal{A}}(\gamma_{|[0,t]}) = \int_0^t \frac{|\dot{\gamma}_s|^2}{2} - \frac{1}{2}\int_0^t V''_{\hat{\gamma}_s}.\gamma_s^2 ds + C\varepsilon\int_0^t |\gamma_s|^2 ds$$

for $t \leq n$, and $\gamma \in H^{0,0}_{[0,n]}$. Let us also introduce the functions on $\mathbb{R}^2$:

$$Q_j(x,y) = \inf_{\gamma_j = x, \gamma_{j+1}=y} \tilde{\mathcal{A}}(\gamma_{|[j,j+1]}) + \varepsilon \frac{x^2+y^2}{2}$$

for $0 \leq j \leq n-1$. They are quadratic forms on $\mathbb{R}^2$.

If we condition the last term of (6.2.3) with respect to $\gamma_1, ..., \gamma_{n-1}$, and apply a Laplace estimate (Theorem 6.2.2) for fixed $\gamma_1, ..., \gamma_{n-1}$ and for the action $\tilde{\mathcal{A}}$, we get

$$\int_{||\gamma||_\infty \leq \varepsilon} e^{\frac{\beta}{2}\int_0^n V''_{\bar{\gamma}_s}\cdot\gamma_s^2 - \beta\varepsilon b_n(\gamma)} d\mathcal{W}^{\beta,0,0}_{[0,n]}$$
$$\geq (1 + \underset{\beta\to\infty}{o(1)})^n \int_{||\gamma||_\infty \leq \varepsilon} \frac{e^{-\frac{\beta}{2}(Q_0(0,\gamma_1)+...+Q_{n-1}(\gamma_{n-1},0))}}{\prod_{j=0}^{n-1}[\tilde{\mathcal{A}}''_j]^{1/2}} d\gamma_1 ... d\gamma_{n-1}$$

where $[\tilde{\mathcal{A}}''_j]$ is the determinant of the hessian of $\tilde{\mathcal{A}} : H^{\gamma_j,\gamma_{j+1}}_{[j,j+1]} \longrightarrow \mathbb{R}$ at a minimum, and does not depend on the endpoints γ_j, γ_{j+1}, since the action $\tilde{\mathcal{A}}$ is a quadratic form in the path. Thus, the problem is reduced to estimating integrals over finite dimensional spaces.

But now,

$$\langle Q(\gamma_1, ..., \gamma_{n-1}), (\gamma_1, ..., \gamma_{n-1})\rangle := Q_0(0,\gamma_1) + Q_2(\gamma_1,\gamma_2) ... + Q_{n-1}(\gamma_{n-1}, 0)$$

is a quadratic form in $(\gamma_1, ..., \gamma_{n-1}) \in \mathbb{R}^{n-1}$, which satisfies all the assumptions of Lemma 3.2.9.

Thus, we can use Lemma 3.2.9 and find ρ such that

$$\int_{||\gamma||_\infty \leq \delta} \frac{e^{-\beta(Q_0(0,\gamma_1)+...+Q_{n-1}(\gamma_{n-1},0)}}{\prod \tilde{\mathcal{A}}''_j} d\gamma_1 ... d\gamma_{n-1}$$
$$\geq (1 - e^{-\beta\rho^2\frac{\delta^2}{2}})^{n-1} \int_{\mathbb{R}^{n-1}} \frac{e^{-\frac{\beta}{2}(Q_0(0,\gamma_)+...+Q_{n-1}(\gamma_{n-1},0)}}{\prod \tilde{\mathcal{A}}''_j} d\gamma_1 ... d\gamma_{n-1}$$
$$= (1 - e^{-\beta\rho^2\frac{\delta^2}{2}})^{n-1} \int_{W^{0,0}_{[0,n]}} e^{-\frac{\beta}{2}\int_0^n V''_{\bar{\gamma}_s}\cdot\gamma_s^2 - \beta\varepsilon b_n(\gamma,\gamma)} d\mathcal{W}^{\beta,(0,0)}_{[0,n]} = \frac{(1 - e^{-\beta\rho^2\frac{\delta^2}{2}})^{n-1}}{[{}_n\mathcal{A}''(\gamma_0,\gamma_n) + 2C\varepsilon b_n]^{1/2}}$$

□

Lemma 3.2.10 can now be proved the same way as in Paragraph 3.2, and the estimates of the end of paragraph 3.2 can be performed the same way to yield Proposition 6.2.1.

The last point in the proof of Theorem 0.0.2 is to draw a link between the determinant of the hessian of $\mathcal{A}$, and Lyapunov exponents. We have seen in paragraph 4.2 how it works for a discrete time system. What we need is the analogue of Lemma 4.2.1 for a continuous time system. It is known as the Levit-Smilansky formula:

6.3. **The Levit-Smilansky formula.**

Theorem 6.3.1. *Let $\gamma : [0,t] \longrightarrow \mathbb{R}^d$ be a critical point of the action*

$$\mathcal{A}(\xi_{|[0,t]}) = \int_0^t \left(\frac{||\dot{\gamma}_s||^2}{2} - V(\gamma_s)\right) ds$$

on the affine Hilbert space $H^{x,y}_{[0,t]} = \{\xi \in H_{[0,t]}, \xi_0 = x, \xi_t = y\}$,*whose tangent space* $H^{0,0}_{[0,t]}$ *is endowed with the scalar product*

$$\langle \xi, \eta \rangle = \int_0^t \dot{\xi}_s.\dot{\eta}_s ds$$

Then the hessian $\mathcal{A}''(\gamma)$*, a self-adjoint operator on* $H^{0,0}_{[0,t]}$*, has a well defined determinant – the infinite product of its eigenvalues. And this determinant coincides with the determinant of the linear endomorphism of* $\mathbb{R}^d$*, sending* $y'_0 \in \mathbb{R}^d$ *to* $\frac{y_t}{t}$*, where* $y_s \in T_{\gamma_s}\mathbb{R}^d$ *(*$s \in [0,t]$*) is the solution of the linearized equation:*

$$\ddot{y}_s + V''(\gamma_s).y_s = 0$$
$$y_0 = 0,\ \dot{y}_0 = y'_0$$

We refer the reader to [LS65] for a proof; there also exists a more conceptual proof written by Colin de Verdière ([CV99]).

The Levit-Smilansky formula, combined with Oseledets' theorem, implies the following proposition, analogous to Proposition 4.2.2:

Proposition 6.3.2. *Let* μ *is a* σ*-invariant probability measure on* W*, carried by action-minimizing trajectories of the Euler-Lagrange flow. Then the limit* $\lim \frac{1}{n} \log[_n\mathcal{A}''(\gamma)]$ *exists for* μ*-almost every* γ*, and is equal to*

$$\sum_1^d \lambda_i^+(\gamma_0, \dot{\gamma}_0),$$

where the $\lambda_i^+(\gamma_0, \dot{\gamma}_0)$ *are the d-first (nonnegative) Lyapunov exponents of* $(\gamma_0, \dot{\gamma}_0)$ *under the Euler-Lagrange flow.*

Proposition 6.2.1, associated with Proposition 6.3.2, implies that μ_∞ (that we have identified to an action-minimizing measure on $\mathbb{T}^d \times \mathbb{R}^d$ at the end of paragraph 6.1), maximizes

$$\mu \mapsto h_\phi(\mu) - \frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} \sum_1^d \lambda_i^+(x,v) d\mu(x,v)$$

over the set of action-minimizing measures.

This is equivalent to Theorem 5.0.8, since μ_∞ is the action-minimizing lift of μ^0_∞.

6.4. **Proof of Corollary 0.0.3.** We can assume that $\max V = 0$. When $\omega = 0$, the Mather set $\tilde{\mathcal{M}}_0$ is the set $\{(x,0) \in \mathbb{T}^d \times \mathbb{R}^d, V(x) = 0\}$.

Action-minimizing have zero entropy since the Mather set consists in fixed point of the Euler-Lagrange flow. Besides, if we consider the expansion of V in orthonormal coordinates near a maximum x_0, in the form

$$V(x_0 + y) \sim V(x_0) - \frac{1}{2} \sum |a_i(x_0)|^2 y_i^2 + O(y^3),$$

then the Lyapunov exponents of the fixed point $(x_0, 0)$ are the $\pm|a_i|$. To prove Corollary 0.0.3, it remains to check that:

Lemma 6.4.1. *When* $\omega = 0$*, Assumptions (A2) and (A3) are automatically satisfied.*

Proof. Assumption (A2). Let $V(x_0) = 0$, let $\varepsilon > 0$, and let γ_0, γ_n satisfy $||\gamma_0 - x_0|| \leq c_n\varepsilon$ and γ_0, γ_n satisfy $||\gamma_n - x_0|| \leq c_n\varepsilon$ for some $c_n \in [0, 1]$.

We have $h_n(\gamma_0, \gamma_n) \leq c_n^2\varepsilon^2$; indeed, this last quantity is the action of a curve joining γ_0 to x_0 in the time interval $[0, 1]$, staying at x_0 in the time interval $[1, n-1]$, and going from x_0 to γ_n in the time interval $[n-1, n]$.

On the other hand, the action of any curve γ joining γ_0 to γ_n in the time interval $[0, n]$, and such that $||\gamma_t - x_0|| \geq \varepsilon$ for some t, is larger than $\frac{(1-c_n)^2\varepsilon^2}{2n}$, which is the energy needed to leave the ball $B(x_0, \varepsilon)$ in time n. Thus, if we take $c_n = n^{-2}$, the minimizer of the action joining γ_0 to γ_n must stay inside the ball $B(x_0, \varepsilon)$, and Assumption (A2) is satisfied.

Assumption (A3). Since $\max V = 0$, no change of gage is necessary to ensure that $\mathcal{L}_0 \geq 0$ and $\mathcal{L}_0$ vanishes on the Mather set. Since $V \leq 0$, $h_n(\gamma_0, \gamma_n) \geq \frac{||\gamma_0 - \gamma_n||^2}{2n}$, so that

$$\beta^{d/2} \int e^{-\beta h_n(\gamma_0, \gamma_n)} d\gamma_n \quad \leq \quad \beta^{d/2} \int e^{-\beta \frac{||\gamma_0 - \gamma_n||^2}{2n}} d\gamma_n \quad = \quad (2\pi n)^{d/2} \quad =: \quad B_n$$

□

Exactly the same way, we can prove

Lemma 6.4.2. *Let $L(\gamma_0, \gamma_1) = \frac{||\gamma_1 - \gamma_0||^2}{2} - V(\gamma_0)$ where V is a $\mathbb{Z}^d$-periodic potential. Then, the assumptions (A2), (A3) of Theorem 0.0.1 are satisfied.*

References

[AS82] M. AIZENMANN, B. SIMON, *Brownian motion and Harnack inequality for Schrödinger operators.* Comm. Pure Appl. Math. **35**, no. 2, 209–273, 1982.

[AMB92] S. AUBRY, R.S. McKAY, C. BAESENS, *Equivalence of uniform hyperbolicity for symplectic twist maps and phonon gap for Frenkel-Kontorova models*, Physica D **56**, 123–134, 1992.

[Ba94] G. BARLES, *Solutions de viscosité des équations de Hamilton-Jacobi*, Mathématiques et Applications **17**, Springer, 1994.

[Be88] G. BEN AROUS, *Méthodes de Laplace et de la phase stationnaire sur l'espace de Wiener*, Stochastics **25** no. 3, 125–153, 1988.

[Bes02] U. BESSI, *Aubry-Mather theory and Hamilton-Jacobi equations,* preprint, 2002.

[BDS93] G. BEN AROUS, J.D. DEUSCHEL, D.W. STROOCK, *Precise asymptotics in large deviations*, Bull. Sci. Math. **117** no. 1, 107–124, 1993.

[Ber02] P. BERNARD, *Connecting orbits of time dependent Lagrangian systems.* Ann. Inst. Fourier **52**, no. 5, 1533–1568, 2002.

[CV99] Y. COLIN DE VERDIERE, *Déterminants et intégrales de Fresnel.*, Symposium à la Mémoire de François Jaeger (Grenoble, 1998), Ann. Inst. Fourier **49** no. 3, 861–881, 1999.

[CLT01] G. CONTRERAS, A.). LOPES, P. THIEULLEN, *Lyapunov minimizing measures for expanding maps of the circle.* Ergodic Theory Dynam. Systems 21, no. 5, 1379–1409, 2001.

[Co65] E.T. COPSON, *Asymptotic expansions*, Cambridge U. P., 1965.

[CEL84] M.G. CRANDALL, L.C. EVANS, P.-L. LIONS, *Some properties of viscosity solutions of Hamilton-Jacobi equations.* Trans. Amer. Math. Soc. **282** no. 2, 487–502, 1984.

[CL83] M.G. CRANDALL, P.-L. LIONS, *Viscosity solutions of Hamilton-Jacobi equations.* Trans. Amer. Math. Soc. **277** no. 1, 1–42, 1983.

[DZ92] A. DEMBO, O. ZEITOUNI, *Large deviations techniques and applications*, Jones and Bartlett, 1992.

[Di68] J. DIEUDONNE, *Calcul infinitésimal.* Hermann, Paris 1968.

[Fa] A. FATHI, *Weak KAM theorem in Lagrangian dynamics*, book in preparation.

[Fa97-1] A. FATHI, *Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens*, C. R. Acad. Sci. Paris, Série I **324**, 1043-1046, 1997.

[Fa97-2] A. FATHI, *Solutions KAM faibles conjuguées et barrières de Peierls*, C. R. Acad. Sci. Paris, Série I **325**, 649-652, 1997.

[Go02] D.A. GOMES, *A stochastic analog of Aubry-Mather theory*, Nonlinearity **15** no.3, 581–603, 2002.

[Ha90] , N.T.A. HAYDN, *Gibbs' functionals on subshifts.* Comm. Math. Phys. **134**, no. 2, 217–236, 1990.

[KH95] A.B. KATOK, B. HASSELBLATT, *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its applications **54**, Cambridge University Press, 1995.

[He88] B. HELFFER, *Semi-classical analysis for the Schrödinger operator and applications.* Lecture Notes in Mathematics, 1336. Springer-Verlag, Berlin, 1988.

[Ki73] J.F.C. KINGMAN, *Subadditive ergodic theory*, Ann. Prob. **1**, 883–909, 1973.

[Ku75] H.H. KUO, *Gaussian measures in Banach spaces*, LNM **463**, 1975.

[LS65] S. LEVIT, U. SMILANSKY, *A theorem on infinite products of eigenvalues of Sturm-Liouville type operators*, Proc. A.M.S. **65** no.2, 299–302, 1977.

[Mn92] R. MANE, *On the minimizing measures of Lagrangian dynamical systems.* Nonlinearity **5** (3), 623–638, 1992.

[Mn96] R. MANE, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity **9** (2), 273–310, 1996.

[Mn97] R. MANE, *Lagrangian flows: the dynamics of globally minimizing orbits*, Int. Cong. on Dyn. Syst. in Montevideo, F. Ledrappier, J. Lewowicz, S. Newhouse ed., Pitman Res. Notes in Math. **362**, 120–131, 1996; reprinted in Bol. Soc. Bras. Mat. Vol **28**, no. 2, 141–153, 1997.

[Ma91] J. MATHER, *Action-minimizing invariant measures for positive definite Lagrangian systems*, Math. Z. **207**, 169–207, 1991.

[Ru78] D. RUELLE *Thermodynamic formalism. The mathematical structures of classical equilibrium statistical mechanics.* Encyclopedia of Mathematics and its Applications, **5**. Addison-Wesley Publishing Co., Reading, Mass., 1978

[Sc66] M. SCHILDER, *Some asymptotic formulas for Wiener integrals*, Trans. A.M.S. **125**, 63–85, 1966.

[Si79] B. SIMON, *Functional integration and quantum physics.* Pure and Applied Mathematics, 86. Academic Press, Inc. (Harcourt Brace Jovanovich, Publishers), New York-London, 1979.

[Si82] Ya.G. SINAI, *Commensurate-incommensurate phase transitions in one-dimensional chains.* J. Statist. Phys. **29**, no. 3, 401–425, 1982.

[Si91] Ya.G. SINAI, *Two results concerning asymptotic behavior of solutions of the Burgers equation with force.* J. Statist. Phys. **64**, no. 1-2, 1–12, 1991.

[Th72] D.J. THOULESS, *A relation between the density of states and range of localization for one dimensional random systems*, J. Phys. C: Solid state Phys., **5**, 77–81, 1972.

[Va67] S. R. S. VARADHAN, *Diffusion processes in small time intervals*, Comm. Pure Appl. Math. **20**, 659–685, 1967.

Unité de Mathématiques Pures et Appliquées, Ecole Normale Supérieure, 6, allée d'Italie, 69364 LYON Cedex 07, FRANCE

E-mail address: nanantha@umpa.ens-lyon.fr