Astérisque

VADIM KALOSHIN
JOHN N. MATHER
ENRICO VALDINOCI

Instability of resonant totally elliptic points of symplectic maps in dimension 4

Astérisque, tome 297 (2004), p. 79-116
<http://www.numdam.org/item?id=AST_2004__297__79_0>

Astérisque
297, 2004, p. 79–116

INSTABILITY OF RESONANT TOTALLY ELLIPTIC POINTS OF SYMPLECTIC MAPS IN DIMENSION 4

by

Vadim Kaloshin, John N. Mather & Enrico Valdinoci

Abstract. — A well known Moser stability theorem states that a generic elliptic fixed point of an area-preserving mapping is Lyapunov stable. We investigate the question of Lyapunov stability for 4-dimensional resonant totally elliptic fixed points of symplectic maps. We show that generically a convex, resonant, totally elliptic point of a symplectic map is Lyapunov unstable. The proof heavily relies on a proof of J. Mather of existence of Arnold diffusion for convex Hamiltonians in 2.5 degrees of freedom. The latter proof is announced in [**Ma5**], but still unpublished.

Résumé **(Instabilité des points totalement elliptiques résonnants d'applications symplectiques en dimension 4)**

Un théorème célèbre de Moser établit la stabilité au sens de Lyapounov des points fixes elliptiques génériques des applications qui conservent l'aire. On étudie la stabilité de Lyapounov des points fixes totalement elliptiques résonnants d'applications symplectiques en dimension 4. On montre que, génériquement, un point totalement elliptique résonnant convexe d'une application symplectique est instable au sens de Lyapounov. La démonstration s'appuie de façon essentielle sur celle donnée par J. Mather pour l'existence d'une diffusion d'Arnold pour les hamiltoniens convexes à 2,5 degrés de liberté. Celle-ci, annoncée dans [**Ma5**], n'est pas encore publiée.

1. Introduction

J. Moser investigated the smooth area-preserving diffeomorphisms f of the plane with elliptic fixed points. He showed [**Mo**] (see also [**LM**] for a simple proof) that, if the linearization $df(p_0)$ of f at a fixed point p_0 has eigenvalues $\exp(\pm 2\pi i\omega)$, which is

2000 ***Mathematics Subject Classification.*** — Primary 34D20; Secondary 34D23, 37J10, 37J25, 37J45, 37J50, 70F15.
Key words and phrases. — Lyapunov stability of periodic orbits, Arnold diffusion, variational methods, minimal measures, Mather theory.

The first author is partially supported by AIM and Sloan fellowships and NSF grants. The second author is partially supported by NSF grant DMS-0245336. The third author is partially supported by MIUR Variational Methods and Nonlinear Differential Equations.

not a small root of unity, then generically p_0 is Lyapunov stable[1]. An application of such result is the stability of the planar restricted three body problem (see *e.g.* [**MH**]).

Let $\mathbb{R}^{2n}$ be the Euclidean space $(x_1,\dots,x_n,y_1,\dots,y_n)\in\mathbb{R}^{2n}$ and Ω be the standard bilinear skew-symmetric 2-form $\Omega=\sum_{i=1}^n dx_i\wedge dy_i$. A C^s smooth map $f:\mathbb{R}^{2n}\to\mathbb{R}^{2n}$ is called *symplectic* if it preserves ω, *i.e.* $f^*\Omega=\Omega$. Let $f(0)=0$ be a fixed point. We say that a fixed point is *totally elliptic* if all the eigenvalues of the linearization $df(0)$ are pairwise complex conjugate, non-real, and of absolute value one, *i.e.* $\exp(\pm 2\pi i\omega_j)$, $2\omega_j\notin\mathbb{Z}$, $j=1,\dots,n$. A fixed point 0 is called *Lyapunov stable* if for any $\varepsilon>0$ there is $\delta=\delta(\varepsilon)>0$ such that if $|p-0|<\delta$, then $|f^n p-0|<\varepsilon$ for all n.

In the multidimensional case (*i.e.* $n>1$), totally elliptic periodic points are the only possible stable periodic points. Indeed, since $df(0)$ preserves the skew-symmetric form ω if one of eigenvalues λ of $df(0)$ is not 1 in absolute value, then λ^{-1} is also an eigenvalue. So either λ or λ^{-1} in absolute value exceeds 1, say λ. The approximation of the dynamics by linearization shows that p_0 is unstable along the eigenspace corresponding to λ.

R. Douady [**Dou**] proved that the stability or instability property of a totally elliptic point is *a flat phenomenon* for C^∞ mappings. Namely, if a C^∞ symplectic mapping f_0 satisfies certain non degeneracy hypotheses, then there are two mapping f and g such that

- f_0-f and f_0-g are flat mappings at the origin and
- the origin is Lyapunov unstable for f and Lyapunov stable for g.

In the present paper we begin an investigation of totally elliptic fixed points in dimension 4. Let $f:\mathbb{R}^4\to\mathbb{R}^4$ be a C^s smooth symplectic mapping with a fixed totally elliptic point at the origin $10\leqslant s\leqslant\infty$. Denote the eigenvalues of $df(0)$ by $\exp(\pm 2\pi i\omega_j^0)$, $j=1,2$. We assume that:

(H1, *resonance*) Let $\omega^0=(\omega_1^0,\omega_2^0)$ have a resonance of order at least 10, *i.e.* for any $k=(k_0,k_1,k_2)\in\mathbb{Z}^3$, $(k_1,k_2)\neq 0$ such that $k_0+k_1\omega_1^0+k_2\omega_2^0=0$ we have $|k_1|+|k_2|>9$ and there is at least one k with this property. Denote

$$k_{\omega^0}=\min\{|k_1|+|k_2| :\ k_0+k_1\omega_1^0+k_2\omega_2^0=0\}\quad\text{and}\quad d_{\omega^0}=\frac{1}{2}\min\{k_{\omega^0},s\}.$$

In particular, (H1) does not exclude possibility of rational $\omega^0=(p_1/q,p_2/q)$ with $q>p_1,\ p_2$ and $|p_1|+|p_2|>9$. We shall not consider low order resonances here.

Denote $\Lambda_k\subset\mathbb{R}^2$ the line of ω's in the frequency space satisfying this equation. Notice that such line passes through ω^0. As a matter of fact, we shall construct orbits diffusing "along" Λ_k.

[1] We remark that earlier a weaker result was obtained by V. Arnold [**Ar1**]. Lyapunov stability will be defined below.

Let $(x_1, x_2, y_1, y_2) \in \mathbb{R}^4$ be Euclidean coordinates. Let us introduce "canonical polar coordinates":

$$x_j = \sqrt{2r_j}\cos 2\pi\theta_j, \quad y_j = \sqrt{2r_j}\sin 2\pi\theta_j, \quad j = 1, 2,$$

where θ_j is determined modulo 1 or simply $\theta_j \in \mathbb{T}$ and $r_j \geqslant 0$. Denote $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geqslant 0\}$ the positive semi-axis. To avoid degeneracy of transformation to polar coordinates, it is convenient to introduce cones. For any $0 < \alpha < 1$, $0 < \rho$ denote

$$K^\alpha_\rho = \{(r_1, r_2) : \ 0 < \alpha r_1 < r_2 < \rho, \ 0 < \alpha r_2 < r_1 < \rho\}.$$

In the interior of K^α_ρ, the transformation from Euclidean to polar coordinates is non-degenerate. Denote by

$$\mathcal{K}^\alpha_\rho = \{(\theta_1, \theta_2, r_1, r_2) \in \mathbb{T}^2 \times \mathbb{R}^2_+ : (r_1, r_2) \in K^\alpha_\rho\}$$

the cone part of the ρ-neighborhood of the origin. Its complement contains neighborhood of the planes $\{r_j = 0\}_{j=1,2}$, where polar coordinates are degenerate.

Suppose we have a totally elliptic point at $r = 0$ satisfying (H1). Birkhoff Normal Form (BNF), *e.g.* [**Ar2**], App. 7A or [**Dou**], states that for small r the map $f(\theta_1, r_1, \theta_2, r_2) = (\Theta_1, R_1, \Theta_2, R_2)$ can be written in the form:

$$\begin{aligned} &\begin{pmatrix} \Theta_j \\ R_j \end{pmatrix} \longmapsto \begin{pmatrix} \theta_j + \omega^0_j + B_j r + \dfrac{\partial P(r)}{\partial r_j} \pmod 1 \\ r_j \end{pmatrix} + \mathrm{Rem}(\theta, r), \\ &B = \{B_j\}_{j=1,2} = \{\beta_{ij}\}_{i,j\leqslant 2}, \end{aligned} \tag{1}$$

where $B : \mathbb{R}^2 \to \mathbb{R}^2$ is a symmetric matrix, $P(r)$ is a polynomial in (r_1, r_2) having zero of order at least 3 at the origin $(r_1, r_2) = 0$. The remainder term $\mathrm{Rem} : U \to \mathbb{R}^2$ is naturally defined near the origin $0 \in U \subset \mathbb{R}^2$ and is C^s smooth away from $\{r_j = 0\}_{j=1,2}$. Since condition (H1) rules out resonances of order up to $k_{\omega^0} - 1$, the smallest term in Taylor expansion of $\mathrm{Rem}(\theta, r)$ at the origin is of order at least $2d_{\omega^0} \geqslant 10$ in (x_1, y_1, x_2, y_2). It implies that inside $\mathcal{K}^\alpha_\rho$ all partial derivative with respect to (r_1, r_2) of $\mathrm{Rem}(\theta, r)$ of order $d_{\omega^0} - 1 \geqslant 4$ (resp. $d_{\omega^0} \geqslant 5$) tend to 0 (resp. stay bounded) as $r \to 0$.

We also make the following assumption:

(H2, *positive torsion*) Let B be symmetric non-degenerate positive definite and let it map degenerate planes $\{r_j = 0\}_{j=1,2}$ transversally to the resonant line Λ_k, *i.e.* for $j = 1, 2$ the intersection $B\{r_j = 0\} \cap \Lambda_k$ is exactly one point.

Generically B is symmetric non-degenerate and satisfies image condition. However, B is not necessarily positive definite. M. Herman [**Her**] gave an example of Hamiltonian systems and symplectic maps arbitrarily close to integrable, which have elliptic fixed points with B not positive definite. The positive definiteness assumption on B is needed to recover fiber-convexity hypothesis required to apply Mather theory.

Let $\alpha > 0$ be small enough so that the image cone $\omega^0 + B(K_1^{4\alpha})$ contains an nonempty interval of Λ_k around ω^0:

$$(2) \qquad \left(\Lambda_k \cap \left\{\omega^0 + B(K_1^{4\alpha})\right\}\right) \smallsetminus \omega^0 \neq \varnothing.$$

We shall restrict consideration of the remainder terms of BNF to the cone $\mathcal{K}_\rho^\alpha$ for small $\rho > 0$.

1.1. Genericity of totally elliptic points and the main result. — Here we shall formalize the notion of a generic totally elliptic point. Let $\widehat{K}_\rho^\alpha = K_\rho^\alpha \cup \{0\}$ and $\widehat{\mathcal{K}}_\rho^\alpha = \mathbb{T}^2 \times \widehat{K}_\rho^\alpha$. We denote $C^{s,d}(\widehat{\mathcal{K}}_\rho^\alpha)$— the space of C^s functions with the natural C^s-topology having all partial derivative of order d bounded and of order $(d-1)$ tending to 0 as r tends to 0 inside $\mathcal{K}_\rho^\alpha$ and $(\theta, r) \in \mathbb{T}^2 \times \widehat{K}_\rho^\alpha = \widehat{\mathcal{K}}_\rho^\alpha$ coordinates.

Fix ω^0 satisfying (H1) and let $d = \frac{1}{2}\{k_{\omega^0}, s\}$. Consider the canonical polar coordinates. Denote

$$(3) \qquad \begin{gathered}\left(\left(\frac{\partial P(r)}{\partial r_1}, \frac{\partial P(r)}{\partial r_1}\right) \pmod 1, r_1, r_2\right) + \mathrm{Rem}(\theta, r) = \mathcal{R}(\theta, R) \in \mathbb{T}^2 \times \mathbb{R}_+^2 \\ \mathcal{R}(\theta, r) = (\Theta_1(\theta, r), \Theta_2(\theta, r), R_1(\theta, r), R_1(\theta, r)) \in \mathbb{T}^2 \times \mathbb{R}_+^2.\end{gathered}$$

Denote the space of remainder terms $\mathrm{Rem}(\theta, r)$ in BNF (1) defined on $\mathcal{K}_\rho^\alpha$ for some small $\rho > 0$ by $\mathcal{R}_{\alpha,\rho}^s$. In a view of discussion after BNF (1) we have that $\mathcal{R}_{\alpha,\rho}^s \subset C^{s,d}(\widehat{K}_\rho^\alpha)$. With the above notations BNF (1) becomes

$$(4) \qquad \begin{pmatrix}\Theta_j \\ R_j\end{pmatrix} \longmapsto \begin{pmatrix}\theta_j + \omega_j^0 + B_j r \\ r_j\end{pmatrix} + \mathcal{R}(\theta, r).$$

Let s be a positive integer. Let M be one of U, $\mathbb{T}^2 \times U$, $\mathbb{T}^2 \times U \times \mathbb{T}$, or $\mathcal{K}_\rho^\alpha$. If f is a C^s real valued function on M, the C^s *norm* $\|f\|_s$ of f

$$\|f\|_s = \sup_{m \in M, |\alpha| \leqslant s} \|\partial^\alpha f(m)\|,$$

where the supremum is taken over the absolute values of all partial derivatives ∂^α of f order $\leqslant s$. The Banach space of C^s real valued functions on M with the C^s–norm is denoted $C^s(M)$. The topology associated to the C^s–norm is called the C^s*–topology.*

Consider the space of remainders $\mathcal{R}_{\alpha,\rho}^s$. We endow it with the strong C^s–topology on the space of functions on a non-compact manifold or the C^s Whitney topology. A base for this topology consists of sets of the following type. Let $\Phi = \{\varphi_i, U_i\}_{i \in \Lambda}$ be a locally finite set of charts on $\mathbb{T}^2 \times K_\rho^\alpha$, where K_ρ^α is the open cone. Let $K = \{K_i\}_{i \in \Lambda}$ be a family of compact subsets of $\mathbb{T}^2 \times K_\rho^\alpha$, $K_i \subset U_i$. Let also $\varepsilon = \{\varepsilon_i\}_{i \in \Lambda}$ be a family of positive numbers. A strong basic neighborhood $\mathcal{N}^s(f, \Phi, K, \varepsilon)$ is given by

$$\|(f\varphi_i)(x) - (g\varphi_i)(x)\|_s \leqslant \varepsilon_i,$$

The strong topology has all possible sets of this form.

The set of C^∞ (*i.e.* infinitely differentiable) real valued functions on M is denoted $C^\infty(M)$. The C^∞*–topology* on $C^\infty(M)(= \cap_s C^s(M))$ is the topology generated by the union of C^s topologies, and it may be also described as the projective (or inverse) limit of the C^s topologies.

Definition 1.1. — We say that a totally elliptic point satisfying (H1-H2) is of *generic type* if the remainder $\mathcal{R}(\theta, r)$ belongs to a set C^s Whitney open dense set in $\mathcal{R}^s_{\alpha,\rho}$.

The main result, announced in this paper, is the following:

Theorem 1.2. — *Suppose hypotheses (H1-H2) hold true, $\alpha > 0$ and satisfies (2), $\rho > 0$ is small, $10 \leqslant s \leqslant \infty$, $\mathcal{R}$ is a remainder term of f given by (3-4). Then, for $\mathcal{R}(\theta, r) \in \mathcal{R}^s_{\alpha,\rho}$ of generic type, the elliptic fixed point 0 is Lyapunov unstable. Moreover, there is $0 < 4\delta = 4\delta\,(\alpha, \{P_j, Q_j\}_{j=1,2}) < \rho$ such that there is a pair of points $(\theta^{\pm}, r^{\pm})$ with $|r^{\pm}| > \delta$ and $f^n(\theta^{\pm}, r^{\pm}) \to 0$ as $n \to \pm\infty$, respectively, and trajectories $\{f^n(\theta^{\pm}, r^{\pm})\}_{n\in\mathbb{Z}_+}$ belong to $\mathcal{K}^{2\alpha}_{2\delta}$.*

Remark 1.3. — As a matter of fact, in Theorem 8.1 below we shall give further details about diffusing trajectories $\{f^n(\theta^{\pm}, r^{\pm})\}_{n\in\mathbb{Z}_{\pm}}$. An important point is that these trajectories diffuse along the resonant segment Λ_k (see (H1)) and, therefore, belong to $\mathcal{K}^{2\alpha}_{2\delta}$ avoiding degenerate planes $\{r_j = 0\}_{j=1,2}$.

Remark 1.4. — The above result can be viewed as a counterexample to a 4-dimensional counterpart to Moser stability theorem under hypotheses (H1-H2) of a resonance between eigenvalues.

Remark 1.5. — As the reader will see, the proof essentially relies on Mather's proof of existence of Arnold diffusion for a cusp residual set of nearly integrable convex Hamiltonian systems in 2.5 degrees of freedom [**Ma5, Ma4**]. The latter proof is highly involved, long, and extremely complicated. Since it is still unpublished, we do not find it possible to describe it here in full details. This is the main reason why this paper is an announcement of Theorem 1.2. Below we just extract necessary intermediate results from Mather's proof. The application to our result is carried out in Section 9.

Remark 1.6. — We hope to get rid of resonant hypothesis (H1) in future work. However, positive torsion (H2) is crucial to apply variational methods and Mather theory.

Assumptions of high differentiability $s \geqslant 10$ and absence of low order resonances $k_{\omega^0} \geqslant 10$ are required to extract sufficient differentiability of the remainder term $\mathcal{R}(\theta, r)$ with respect to r at $r = 0$ in "canonical polar coordinates" inside a cone $\mathcal{K}^{\alpha}_{\rho}$. More presicely, $\mathcal{R} \in C^{s,d}(\mathcal{K}^{\alpha}_{\rho})$ for $d \geqslant 5$. See representation of the remainder in the form (11).

The proof is organized as follows. In Section 2, we suspend a symplectic map f satisfying hypothesis (H1-H2) in the small cone $\mathcal{K}^{\alpha}_{\rho}$ near a totally elliptic point 0 to a time periodic fiber-convex Hamiltonian $H_f(\theta, r, t)$, *i.e.* we construct a Hamiltonian whose time 1 map equals f in $\mathcal{K}^{\alpha}_{2\delta}$. In Section 2.1, we recall how to switch from Hamiltonian equations to Euler-Lagrange equations. Section 3 is devoted to an outline of the proof of Theorem 1.2, *i.e.* the proof of existence of "diffusing" trajectories. In Section 4, we state Mather Diffusion Theorem [**Ma5**] in terms of Lagrangians. An important part of this result is an explicit list of non-degeneracy hypotheses which guarantee existence of diffusion. In Sections 5 – 7 we state these non-degeneracy hypotheses. In Section 8 we restate Mather Diffusion Theorem in terms of existence of a minima for a certain Variational Principle due to Mather [**Ma5**]. Existence of such a minimum corresponds to existence of a "diffusing" trajectory. In Section 9, we verify that for small positive δ_0 and $\{\delta_j = 2^{-j}\delta_0\}_{j\in\mathbb{Z}_+}$ in each annulus $0 < \delta_{j+1} \leqslant |r| \leqslant \delta_j \leqslant \rho \ll 1$ intersected with $\mathcal{K}^{\circ}_{2\delta}$ the symplectic map (1) (resp. the suspending Hamiltonian H_f) is a small perturbation of an integrable map (resp. an integrable Hamiltonian). Therefore, we manage to apply the above mentioned Variational Principle in each of these annuli. In the final Section, we derive the main result (Theorem 1.2) by "gluing" the annuli. Namely, show existence of a minima to the aforementioned Variational Principle and conclude that it corresponds to one of "diffusing" trajectories from Theorem 1.2. Existence of the other trajectory can be proven in the same way. This would complete the proof. For the reader's convenience, this paper is provided with two appendices: in Appendix A we introduce necessary notions and objects of Mather theory, while Appendix B contains proofs of auxiliary lemmas.

Sections 2, 3, 9, and Appendices A & B are written by the first and the third authors. Sections 4-8 are written by the first author based on the graduate class of the second author [**Ma4**].

2. Suspension of a symplectic map near totally elliptic points of a time periodic fiber-convex Hamiltonian

Moser [**Mo2**] showed how to suspend a twist map of a cylinder to a time 1 map of a time periodic *fiber-convexity* Hamiltonian, *i.e.* Hessian $\partial^2_{rr}H$ in r is positive definite everywhere. To the best of our knowledge, there is no general extension of this result to higher dimensional case, even locally. We apply the standard method of generating functions to construct a required suspension. Even though the fact we need seems quite standard we could not find an appropriate reference.

The following suspension results are known to the authors. Bialy and Polterovich (see [**Go**], sect. 41, A) proved existence of smooth suspension theorem with fiber-convexity. However, this result makes use of the restrictive assumption that a generating function $S(\theta,\Theta)$ corresponding to $f(\theta,r) = (\Theta, R)$ has to have a symmetric matrix $\partial^2_{\theta,\Theta}S(\theta,\Theta)$. Since such condition is not satisfied in general, we can not apply this result. Kuksin-Poschel [**KP**] proved existence of global analytic suspension, which does not possess fiber-convexity.

We propose here a way to construct a local suspension *keeping fiber-convexity*. Our proof, given in Appendix B, is a modification of the one by Golé [**Go**] (see sect. 41, B).

It is based on the construction of a suitable family of generating functions and on a local analysis of it.

Lemma 2.1. — *Let $f : \mathbb{R}^4 \to \mathbb{R}^4$ be a C^s smooth symplectic map with a totally elliptic point $f(0) = 0$ at the origin satisfying hypothesis (H1-H2) of positive definite torsion and $10 \leqslant s \leqslant \infty$. Then for any $0 < \alpha < 1/2$ and a small positive ρ there is a $C^{s+1,d+1}$ smooth Hamiltonian, written in BNF (θ, r)-polar coordinates (1) as*

$$
(5)\qquad \begin{aligned} H_f(\theta, r, t) &= \omega_1 r_1 + \omega_2 r_2 + \frac{1}{2}\langle Br, r\rangle + P(r) + r_f(\theta, r, t) \\ &= \omega_1 r_1 + \omega_2 r_2 + \frac{1}{2}\langle Br, r\rangle + h_f(\theta, r, t), \end{aligned}
$$

where $P(r)$ is a polynomial in r having zero of order at least 3 at $r = 0$ and $r_f(\theta, r, t)$ belongs $C^{s+1,d+1}(\mathcal{K}^{\alpha}_{2\rho} \times \mathbb{T})$, periodic in t of period 1 and such that the time 1 map of Hamiltonian flow of H equals f in the cone $\mathcal{K}^{2\alpha}_{\rho}$.

Remark 2.2. — By definition of H_f, we see that $\partial^2_{rr} H_f$ is positive definite for small $r \in K^{\alpha}_{\rho}$ and all $(\theta, t) \in \mathbb{T}^2 \times \mathbb{T}$. Notice that one can not deduce that H_f is positive definite in a full ρ-neighborhood of zero, since polar coordinates are degenerate along the planes $\{r_j = 0\}_{j=1,2}$ and the origin. This hides the degeneracy of positive definiteness. This is the reason why we restrict our suspension to the cone $\mathcal{K}^{2\alpha}_{\rho}$, which does not contain those planes.

The proof of this lemma is in Appendix B.

2.1. Hamiltonian and Euler-Lagrange flows are conjugate. — In this Section, which may be skipped by an expert, we exhibit the standard duality between Hamiltonian and Lagrangian systems given by the *Legendre* transform. More explicitly, we state that if a Hamiltonian H satisfies certain conditions, then there is a Lagrangian L such that the Hamilton flow of H corresponds to the Euler-Lagrange flow of L after a coordinate change (see *e.g.* [**Ar2**]). Because of this construction, after this Section, we may consider only Euler-Lagrangian flows.

We shall denote $(\theta, v) \in \mathbf{T}\mathbb{T}^n \simeq \mathbb{T}^n \times \mathbb{R}^n$. The Legendre transform associates to a Hamiltonian $H(\theta, r, t)$, $H : \mathbf{T}^*\mathbb{T}^n \times \mathbb{T} \to \mathbb{R}$, which is assumed to be positive-definite in r, a Lagrangian $L(\theta, v, t)$, $L : \mathbf{T}\mathbb{T}^n \times \mathbb{T} \to \mathbb{R}$, which is positive-definite in v, according to the following scheme:

$$
(6)\qquad \begin{aligned} &L(\theta, v, t) = \sup_{r \in \mathbf{T}^*_\theta \mathbb{T}^n} \{\langle r, v\rangle - H(\theta, r, t)\}, \\ &\text{where } \langle\ .\ \rangle : \mathbf{T}\mathbb{T}^n \times \mathbf{T}^*\mathbb{T}^n \longrightarrow \mathbb{R} \quad \text{is pairing between dual spaces.} \end{aligned}
$$

When L is related to H as above, we say that L is the *dual* of H. Let us consider the *Euler-Lagrange* flow associated to L. The latter is defined as a flow on the extended phase space $\mathbf{T}\mathbb{T}^n \times \mathbb{T}$ such that its trajectories $(\theta(t), \dot\theta(t), t) = (d\theta(t), t)$ are solutions of the Euler-Lagrange equation:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot\theta}\right) = \frac{\partial L}{\partial \theta} \tag{7}$$

The trajectories of the Euler-Lagrange flow can be also characterized as those which minimize locally the action $\int L(d\gamma(t), t)dt$ among absolutely continuous curves with the same boundary conditions. The standard (but crucial for our purposes) observation is that when we subtract a *closed* 1-form η, defined on $\mathbb{T}^n \times \mathbb{T}$, from the Lagrangian L, then both $L - \eta$ and L have *the same* Euler-Lagrange equations (see *e.g.* [**Fa**]).

Let us suppose the Hamiltonian $H(\theta, r, t)$ satisfies the following properties:

(1) *Positive definiteness in r:* For each $(\theta, r) \in \mathbf{T}^*\mathbb{T}^n$ and $t \in \mathbb{T}$ the restriction of H to $\mathbf{T}^*_\theta\mathbb{T}^n \times \{t\}$ is positive definite;

(2) *Super-linear Growth in r:* For each $(\theta, r) \in \mathbf{T}^*\mathbb{T}^n$ and $t \in \mathbb{T}$

$$\frac{H(\theta, r, t)}{\|r\|} \longrightarrow +\infty \quad \text{as} \quad \|r\| \longrightarrow +\infty$$

(3) *Completeness:* All the solutions of the Hamiltonian equations can be extended for all $t \in \mathbb{R}$.

We need these conditions to be satisfied in order to apply Mather theory (see Appendix A and Section 9). Notice that the Hamiltonian H of the form (5) satisfies all these properties near $r = 0$. The standard result says:

***Lemma 2.3* (see *e.g.* [Ar2], § 15).** — *If a C^{s+1} Hamiltonian $H(\theta, r, t)$ satisfies the above conditions (1–3) with $s \geqslant 1$ and $L(\theta, v, t)$ is the dual of H, then the map $\mathcal{L} : (\theta, r, t) \to (\theta, v, t)$, given by*

$$\mathcal{L}(\theta, r, t) = (\theta, \partial_r H(\theta, r, t), t), \tag{8}$$

is C^s-smooth and invertible, and it conjugates the Hamiltonian flow of H to the Euler-Lagrange flow of L, i.e. it provides a one-to-one correspondence between trajectories of both flows. Moreover, the Lagrangian L satisfies properties (1–3) above with r replaced by v and H by L.

Let H_f be given by formula (5). Let H^*_f be the integrable part $H^*_f(r) = \omega_1 r_1 + \omega_2 r_2 + \frac{1}{2}\langle Br, r\rangle + P(r)$. Namely, $P(r)$ is the polynomial in r part of $h_f(\theta, r)$. Notice that Legendre transform of H_f has the form

$$L_f(\theta, v, t) = \ell_0(v - \omega^0) + P_f(\theta, v - \omega^0, t), \tag{9}$$

where $\ell_0(v - \omega^0)$ is the Legendre transform of $H^*_f(r)$ and P_f is a C^s smooth remainder defined on $\mathbb{T}^2 \times \{\partial_r H_f(\theta, K^{2\alpha}_{2\rho}, t) - \omega^0\} \times \mathbb{T}$. The form (5) of H_f shows that $\partial_r H_f(\theta, r, t) = \omega^0 + Br + \partial_r h_f(\theta, r, t)$. Therefore, for small ρ we have that $B(K^\alpha_\rho) \subset \{\partial_r H_f(\theta, K^{2\alpha}_{2\rho}, t) - \omega^0\}$ so we could assume that L_f is well-defined on $B(K^\alpha_\rho)$. Moreover, P_f has a zero in $(v - \omega^0)$ of at least 6-th order, *i.e.* $P_f \in C^{s,d+1}(\mathbb{T}^2 \times B(K^\alpha_\rho) \times \mathbb{T})$

for $d \geqslant 5$. We shall apply Mather's technology to the Lagrangian L_f and its Euler-Lagrange flow.

3. Scheme of construction of diffusing trajectories using Mather action functional

In this Section, we outline a variational approach due to Mather [**Ma5, Ma4, Ma3**] to construct diffusing trajectories toward and outward from zero from Theorem 1.2. We concentrate on the one going toward zero first. Construction of the other one is very similar.

3.1. Rough sketch of the proof of Theorem 1.2. — Application of lemma 2 for $n = 2$ to the symplectic map f, given by (1) provides the suspension Hamiltonian $H_f(\theta, r, t)$ given in $\mathcal{K}^{\alpha}_{2\delta}$ near $r = 0$. We have that locally, *i.e.*, in $\mathcal{K}^{\alpha}_{2\delta}$, H_f satisfies fiber-convexity from hypothesis (1, Sect. 2.1) above. To meet hypotheses (2–3, Sect. 2.1), one may smoothly extend H_f for large r keeping convexity in r so that it is an integrable Hamiltonian, *e.g.* given by (36). Thus, Mather theory is applicable (see Section 9).

Let $L_f(\theta, v, t)$ be the dual of $H_f(\theta, r, t)$, given by (9). The Legendre coordinate change (8) in our case has the form $\mathcal{L}(\theta, r, t) = (\theta, \omega^0 + Br + \partial_r h_f(\theta, r, t), t)$. Let us approximate it by its linearization $T_{\omega^0, B} : r \to v = \omega^0 + Br$. Denote $K^{\alpha}_{\rho}(\omega^0, B) = T_{\omega^0, B}(K^{\alpha}_{\rho})$ the image cone, whose complement we need to avoid. By (H2), for small $\alpha > 0$ the image cone $K^{2\alpha}_{\rho}(\omega^0, B)$ contains a segment in Λ_k of length $2\rho/\|B^{-1}\|$ centered at ω^0. Fix $\alpha = \alpha(k, B) > 0$ with the above property. We shall "diffuse" inside $K^{\alpha}_{\rho}(\omega^0, B)$. Denote by e_k the unit vector parallel to Λ_k and fix a small $0 < \delta < \rho/\|B^{-1}\|$ (to be determined later).

Put $2\delta = \delta_0$ and $\delta_j = 2^{-j}\delta_0$ for each $j \in \mathbb{Z}_+$. Fix the sequence of annuli

$$A_j(\omega^0) = \left\{2^{-2}\delta_j < |v - \omega^0| < 2^2\delta_j\right\} \subset \mathbb{R}^2_+. \tag{10}$$

Denote $K^{\alpha}_{\delta, j}(\omega^0, B) = K^{\alpha}_{2\delta}(\omega^0, B) \cap A_j(\omega^0)$. By definition for each $j \in \mathbb{Z}_+$ we have $\omega^0 + \delta_{j+1}e_k, \omega^0 + \delta_j e_k, \omega^0 + \delta_{j-1}e_k \in K^{\alpha}_{\delta, j}(\omega^0, B)$ and adjacent annuli $A_j(\omega^0)$ and $A_{j+1}(\omega^0)$ overlap, *i.e.* $A_j(\omega^0) \cap A_{j+1}(\omega^0) \neq \varnothing$. We now point out the aim of our constructions and the steps needed to reach it:

The goal. Construct a diffusing trajectory $\{(\theta, \dot\theta)(t)\}_{t \geqslant 0}$ such that at time 0 its velocity $\dot\theta(t)$ is approximately $\omega^0 + \delta_0 e_k$;

Stage 1. At a time $\mathcal{T}_1 > 0$, its velocity $\dot\theta(\mathcal{T}_1)$ is approximately $\omega^0 + \delta_1 e_k$ and, in between 0 and $\mathcal{T}_1$, we have $\dot\theta(t) \in K^{\alpha}_{\delta, 1}(\omega^0, B)$;

Stage 2. At a time $\mathcal{T}_2 > \mathcal{T}_1$, its velocity $\dot\theta(\mathcal{T}_1)$ is approximately $\omega^0 + \delta_2 e_k$ and, in between $\mathcal{T}_1$ and $\mathcal{T}_2$, we have $\dot\theta(t) \in K^{\alpha}_{\delta, 2}(\omega^0, B)$ and so on;

Stage j. At a time $\mathcal{T}_j > \mathcal{T}_{j-1}$, its velocity $\dot\theta(\mathcal{T}_1)$ is approximately $\omega^0 + \delta_j e_k$ and, in between $\mathcal{T}_{j-1}$ and $\mathcal{T}_j$, we have $\dot\theta(t) \in K^{\alpha}_{\delta,j}(\omega^0, B)$, and so on for possibly infinite number of stages.

If we could construct a trajectory with these properties, we would obtain a trajectory for the symplectic map f which goes toward the origin. To construct a trajectory going out from the origin, the arguments involved are analogous. This would indeed prove Theorem 1.2. Formalization of this scheme requires some notions of Mather theory.

3.2. A naive idea of Mather's mechanism of diffusion. — We start with

A Model Example. — Suppose $f : M \to M$ be a smooth diffeomorphism of an $2n$-dimensional manifold possibly with a boundary. Let m be any positive integer and $p_1, \dots, p_m$ be a collection of hyperbolic periodic points of the same index, *i.e.* the dimensions of stable and unstable manifolds are the same. Suppose that for each $i = 1, \dots, m-1$ the unstable manifold $W^u(p_i)$ intersects the stable manifold $W^s(p_{i+1})$ transversally and both belong to M. Then, it is easy to show that $W^s(p_1)$ intersects $W^u(p_m)$.

If $M = \mathbb{T}^2 \times K^{\alpha}_{\rho} \ni (\theta, r)$ and r-coordinates of p_i are close to $\delta_i e_k$, then there exists a trajectory whose r-coordinate change from nearby $\delta_1 e_k$ to nearby $\delta_m e_k$.

As a matter of fact, in Mather's mechanism of diffusion we use the following objects:

the hyperbolic periodic points p_i's are replaced by Mather sets $\mathcal{M}_i$, whose projection onto r-component is localized near $\delta_i e_k$;

the stable and unstable manifolds $W^s(p_i)$ and $W^u(p_i)$ are replaced by the stable and unstable sets $W^s(\mathcal{M}_i)$ and $W^u(\mathcal{M}_i)$ respectively. $W^s(\mathcal{M}_i)$ and $W^u(\mathcal{M}_i)$ are not necessarily manifolds and not even continuous;

to verify the intersection of unstable and stable sets $W^s(\mathcal{M}_i)$ and $W^u(\mathcal{M}_{i+1})$, respectively, we shall use the barrier function defined in Section 8 (see formulas (31-32));

to show that the intersection of $W^s(\mathcal{M}_i)$ and $W^u(\mathcal{M}_{i+1})$ for each $i = 1, \dots, m-1$ implies existence of a connecting trajectory between $\mathcal{M}_1$ and $\mathcal{M}_m$, we define a suitable action functional (24). As it was shown by Mather [**Ma4**], under certain hypotheses, the minimum of such an action functional is achieved on a trajectory of the Euler-Lagrange equation connecting $\mathcal{M}_1$ and $\mathcal{M}_m$ (see Section 8).

In Appendix A, we define Mather sets, barrier functions, and related objects. The reader who is familiar the basics of Mather theory may go directly to next Section. Others may read Appendix A first.

3.3. Detailed scheme of the proof of Theorem 1.2. — We now continue our discussion, assuming that the reader familiar with basic notions of Mather theory (see Section 9). The diffusing trajectories we shall construct move along the resonant segment Λ_k from (H1). Consider a sequence of subsegments of Λ_k given by $\Gamma_j = [\delta_{j+1}e_k, \delta_j e_k] \subset \Lambda_k$, $j \in \mathbb{Z}_+$. Denote

$$\Gamma_\delta = \cup_{j\geqslant 1}\Gamma_j.$$

On each of the segments Γ_j's we mark a sufficiently dense finite set of points $\{\delta_{j,p}e_k\}_{p=1}^{m_j} \subset \Gamma_j$ (we determine later how dense this set has to be). Each stage of diffusion described in Section 3.1 consists of m_j sub-stages. First, we enumerate marked points in Γ_δ. We set $M_p = \sum_{i=1}^{p} m_i$ and for $M_p < i \leqslant M_{p+1}$ we set $\omega_i = \omega^0 + \delta_{p,i-M_p}e_k$. Loosely speaking, $\mathcal{M}_i$ as introduced in the previous Section, is an invariant set of trajectories with approximate rotation vector ω_i. We shall formalize this idea in Section 8. To give a precise definition of $\mathcal{M}_i$'s we need further discussion.

Let $H_f(\theta, r, t)$ be the suspension of the symplectic map f under consideration given in $\mathcal{K}^\alpha_\rho$ near $r = 0$ by lemma 2. We have that locally (*i.e.*, near $r = 0$) H_f satisfies convexity condition (1, Sect. 2.1). To meet hypotheses (2–3, Sect. 2.1), one extends H_f for large r by an integrable Hamiltonian, *e.g.* given by (36) keeping convexity in r. Thus, Mather theory is applicable to our case (see Section 9).

Consider the Legendre transform of H_f, which would lead to a Lagrangian of the form (9). The first term of L_f corresponds to an integrable Lagrangian. Moreover, we show in Section 9 that for small $\delta_0 > 0$ the second term P_f can be considered as a small perturbation. Therefore, we shall be able to apply Mather Diffusion Theorem (see Section 4). We shall write the remainder in the form (34) with $m = 3$ and r replaced by $(v - \omega^0)$

$$P_f(\theta, v-\omega^0, t) = \sum_{p=0}^{3} (v-\omega^0)_1^p (v-\omega^0)_2^{3-p} P_p(\theta, v-\omega^0, t), \tag{11}$$

where $(v - \omega^0)_i$ is i-th coordinate of $(v - \omega^0)$, $i = 1, 2$. Let us denote $\mathbf{P}_f = (P_0, P_1, P_2, P_3)$ and define the unit sphere for perturbations $\mathbf{P}_f$

$$S^{s,3} = \Big\{\mathbf{P}_f : \sum_{p=0}^{3} \|P_p\|^2_{C^3(\mathbb{T}^2\times B(K^{3\alpha}_\rho)\times\mathbb{T})} = 1\Big\}.$$

Since $P_f \in C^{s,d+1}(\mathbb{T}^2 \times B(K^{2\alpha}_\rho) \times \mathbb{T})$ with $d \geqslant 5$, by lemma B.1 we have that C^3 norm of $\mathbf{P}_f$ is well-defined.

We denote also by $\mathcal{L}_{\beta,f}$ the Fenchel-Legendre transform associated with L_f by (29). In the images of each marked frequency, we choose a cohomology class $c_i \in \mathcal{L}_{\beta,f}(\omega_i)$ for each $i = 1, \dots, M_p, \dots$ so that adjacent c_i's are sufficiently close. We are now in position to define the sets $\{\mathcal{M}_i\}_{i=1}^m$ from the previous Section to be Mather sets $\mathcal{M}^{c_i}$. We shall slightly modify the choice of c_i's in Section 8.

Definition 3.1. — We say that $\mathcal{L}_{\beta,f}$ has *channel property* with respect to a resonant segment Γ_δ if there is a smooth connected curve $\sigma_\Gamma \subset \mathcal{L}_{\beta,f}(\Gamma_\delta)$ such that for each $\omega' \in \Gamma_\delta$ the curve σ_Γ intersects $\mathcal{L}_{\beta,f}(\omega')$.

Lemma 3.2 **([Ma4])**. — *Let $10 \leqslant s \leqslant \infty$. Then, for a C^s Whitney open dense set of $\mathbf{P}_f \in S^{s,3}$, there is $\delta = \delta(\mathbf{P}_f) > 0$ such that the Fenchel-Legendre transform $\mathcal{L}_{\beta,f}$ has channel property with respect to Γ_δ. In particular, for any pair of positive integers $i < i'$ the sets $\mathcal{L}_{\beta,f}(\omega_i)$ and $\mathcal{L}_{\beta,f}(\omega_{i'})$ are connected by σ_Γ.*

Remark 3.3. — In Section 6.2 we introduce certain non-degeneracy hypothesis (C1)–(C3) and $(C4)_\omega$–$(C8)_\omega$ for perturbations of integrable Lagrangian systems and in Section 9 show how to adapt these hypothesis for remainder terms in BNF (4) of totally elliptic points. The C^s Whitney open dense set of remainders $\mathbf{P}_f \in S^{s,3}$ that satisfy adapted non-degeneracy hypotheses (C1)–(C3) and $(C4)_\omega$–$(C8)_\omega$ fulfills channel property of the lemma.

We construct trajectories that *diffuse along σ_Γ inside the channel $\mathcal{L}_{\beta,f}(\Gamma_\delta)$*. To accomplish this, roughly speaking, we vary the cohomology c in order to vary the velocity $\dot\theta$.

We shall apply the Mather method of changing Lagrangians [**Ma5**]. Mather applied this method in [**Ma3**] to show the existence of unbounded trajectories for generic time periodic mechanical systems on $\mathbb{T}^2$. We outline some of the key ideas of the method. For simplicity let $L(\theta, v, t) = \frac{1}{2}\langle v, v\rangle + \varepsilon P(\theta, v, t)$ be sufficiently smooth nearly integrable Lagrangian and $\eta^c = c\, d\theta$ be the standard closed one form on $\mathbb{T}^2 \times \mathbb{T}$ for a vector $c \in \mathbb{R}^2 \simeq T_\theta\mathbb{T}^2$, $\theta \in \mathbb{T}^2$. Then the following scheme can be exploited:

(1) Euler-Lagrange flows of L and $(L - \eta^c)$ are the same (see *e.g.* [**Fa**]).

(2) Minimization of c-action $\int_a^b (L - \widehat{\eta}^c)(d\gamma(t), t)dt$ with ε-error leads to minimization of

$$\frac{1}{2}\langle\dot\theta, \dot\theta\rangle - \langle\dot\theta, c\rangle = \frac{1}{2}\Big(\langle\dot\theta - c, \dot\theta - c\rangle - \langle c, c\rangle\Big). \tag{12}$$

Therefore, trajectories minimizing c-action have approximate velocity c. As a matter of fact, even if η^c is a closed one form with $[\eta^c]_{\mathbb{T}^2} = c$, L is close to integrable and $b - a$ is large enough, trajectories minimizing c-action still have approximate velocity c (see [**Ma4**]). From now on we consider η^c as a closed one form.

(3) Suppose we can find an action functional

$$\sum_{i=1}^{j} \int_{t_i}^{t_{i+1}} (L - \widehat{\eta}_i)(d\gamma(t), t)dt \tag{13}$$

for a sequence of closed one forms $\{\eta_i\}_{i=1}^j$ such that $[\eta_i]_{\mathbb{T}^2} = c_i$ and $[\eta_{i+1}]_{\mathbb{T}^2} = c_{i+1}$, where c_i and c_{i+1} are close for each $i = 1, \dots, j-1$ and the minimum of such integral is achieved on a trajectory $\{(d\gamma(t), t) : t \in [t_1, t_m]\}$ of the Euler-Lagrange flow of $L^{(2)}$. Standard properties of action minimization give that this is indeed true for time $t \neq t_1, \dots, t_m$, but it is a delicate problem to show that this does not happen at connection times $t = t_1, \dots, t_m$. The corresponding minimizing trajectory $\gamma(t)$ might have *corners* $\dot\gamma(t_i^-) \neq \dot\gamma(t_i^+)$. Notice now that at time t in $[t_1, t_2]$ velocity is approximately c_1 and at time t in $[t_{m-1}, t_m]$ velocity is approximately c_m. Thus, the key to the method is to find an action functional with the above property and justify absence of corners. In (13) we made only a rough attempt. This functional is defined in Section 8. Usually, this construction is quite involved and highly nontrivial [**Ma5, Ma4, Ma3**].

4. Mather diffusion theorem

In this Section, we state Mather result about existence of Arnold diffusion in a generality we use for our application. See [**Ma5**] for the most general version. As a matter of fact, to prove our main result (Theorem 1.2) in Section 8) we shall reformulate Mather Diffusion Theorem in terms of a certain variational principle and in Section 9 apply this principle to prove Theorem 1.2.

In the subject of Arnold diffusion, one studies a time periodic or autonomous Hamiltonians/Lagrangians that are perturbations of integrable Hamiltonians/Lagrangians (see *e.g.* [**AKN**]).

In the time periodic case, the Lagrangian takes the form

$$L(\theta, \dot\theta, t) = \ell_0(\dot\theta) + \varepsilon P(\theta, \dot\theta, t),$$

where ℓ_0 is a C^s smooth function on a convex closed set with smooth boundary $U \subset \mathbb{R}^2$, ε is a small positive number, P is a C^s smooth function on $\mathbb{T}^2 \times U \times \mathbb{T}$, and $s \geqslant 3$. In other words, P is periodic of period 1 in θ_1, θ_2, and t. The function ℓ_0 is called the *unperturbed integrable Lagrangian* and the function P is called the *perturbation term*.

Denote $d^2_{\dot\theta}\ell_0 = \partial^2_{\dot\theta_i\dot\theta_j}\ell_0$ the Hessian matrix of second partial derivatives of ℓ_0, *i.e.* $d^2_{\dot\theta}\ell_0 = \left(\partial^2_{\dot\theta_i\dot\theta_j}\ell_0\right)_{i,j=1,2}$. We shall assume that $d^2_{\dot\theta}\ell_0$ is everywhere positive definite on U, *i.e.* we have $\sum_{i,j=1}^2 \partial^2_{\dot\theta_i\dot\theta_j}\ell_0(\dot\theta)\dot\varphi_i\dot\varphi_j > 0$, for all $\dot\theta \in U$ and all $(\dot\varphi_1, \dot\varphi_2) \in \mathbb{R}^2 \setminus 0$. In the Hamiltonian case, the analogous assumption is that the unperturbed integrable Hamiltonian convex.

Now, we briefly discuss the problem of Arnold diffusion. For the unperturbed integrable Lagrangian $L = \ell_0$, the Euler-Lagrange (E. L. for short in the sequel) equations reduce to $d^2\theta/dt^2 = 0$. Every solution θ lies on a torus $\{\dot\theta = \omega\}$, where $\omega = (\omega_1, \omega_2) \in U$. The ω_i's are called the *frequencies* of the solution.

By a *trajectory* of L, we mean a solution of the E. L. equations associated to L. Along a trajectory of L, $\dot\theta$ is constant in the case of the unperturbed integrable

(2) Actually, summation could be over infinite number of terms, as formula (25) shows.

system and varies slowly in the case of a small perturbation of the integrable system. The problem of Arnold diffusion is whether $\dot{\theta}$ can vary a lot over long periods of time. Recently a great progress has been achieved in proving Arnold diffusion in so-called *a priori unstable* case by many different groups (see papers [**Be**], [**CY**], [**DLS**], [**T1, T2, T3**], announcements [**X1, X2**], and work in preparation [**KM**]). The result below is for the *a priori stable* case.

Recall that for a positive integer $C^s(M)$ denotes the Banach space of C^s real valued functions on M (see Section 1.1 for notations and definitions). Now let s be ∞, or an integer $\geqslant 3$. We let $\mathcal{L}^s$ denote the topological space of C^s functions $\ell_0 : U \to \mathbb{R}$, such that ℓ is positive definite in $\dot{\theta}$, *i.e.* the Hessian $d^2_{\dot\theta}\ell_0$ is positive definite. Endow $\mathcal{L}^s$ with the C^s topology. We let $\mathcal{P}^s$ denote the topological space of C^s functions $P : \mathbb{T}^2 \times U \times \mathbb{T} \to \mathbb{R}$ endowed with the C^s topology. Denote

$$\mathcal{S}^s_L = \{P \in \mathcal{P}^s : \ \|P\|_{C^3(\mathbb{T}^2\times U\times\mathbb{T})} = 1\}$$

the unit sphere in the space of perturbations. The topology in $\mathcal{S}^s$ is induced from the ambient space $\mathcal{P}^s$.

Definition 4.1. A set $W^s_\delta \subset \mathcal{P}^s$ is called *δ-cusp residual* if

A) there is a non-negative continuous function δ on $\mathcal{S}^s_L$ such that the set $U^s_\delta = \{P \in \mathcal{S}^s_L : \ \delta(P) > 0\}$ is open and dense in $\mathcal{S}^s_L$;

B) there is a cusp set $V^s_\delta = \{\varepsilon P \in \mathcal{P}^s : \ P \in U^s_\delta,\ 0 < \varepsilon < \delta(P)\}$, which is a subset of homogeneous extension of U^s_δ which is defined by $\mathbb{R}U^s = \{\lambda P \in \mathcal{P}^s : P \in U^s,\ \lambda > 0\}$;

C) there is an open and dense set W^s_δ in V^s_δ.

Definition 4.2. If $\Gamma \subset \mathbb{R}^2$ is a line segment, we shall say that it is *rational or resonant* if there is a resonance[(3)] $k = (k_0, k_1, k_2) \in \mathbb{Z}^3$ such that Γ is contained in the line Λ_k.

We say that a curve $\Gamma \subset \mathbb{R}^2$ is a *resonant piecewise linear curve* if it is a finite union of resonant segments $\Gamma = \cup_{s=1}^m \Gamma_s$ so that end points of Γ_s belong to end points of Γ_{s-1} and Γ_{s+1}, for all $s = 2, \dots, m-1$.

The following result is a modified version of the result announced by Mather [**Ma5**] for the time-periodic case:

Mather Diffusion Theorem. *Let Γ be a resonant piecewise linear curve in U and let $3 \leqslant s \leqslant \infty$. There exists a non-negative continuous function $\delta(\ell_0, \Gamma) : \mathcal{P}^s \to \mathbb{R}_+$, such that, for any perturbation εP in a $\delta(\ell_0, \Gamma)$-cusp residual set $W^s_{\delta(\ell_0,\Gamma)} \subset \mathcal{P}^s$, there is a trajectory $(\theta, \dot\theta)(t)$ of $L_\varepsilon = \ell_0 + \varepsilon P$, whose velocity moves along Γ. More precisely, there is a constant $C = C(\ell_0, P, \Gamma) > 0$ and $T = T(\ell_0, P, \Gamma) > 0$ such that* $\operatorname{dist}\big(\cup_{0<t<T}\dot\theta(t), \Gamma\big) \leqslant C\sqrt{\varepsilon}$, *where* dist *is the standard Hausdorff distance between sets in U.*

---

(3) Recall that saying that k is a *resonance*, we mean that $k \in \mathbb{Z}^3$ and $(k_1, k_2) \neq 0$.

Remark 4.3. — The function $\delta(\ell_0, \Gamma)$ (and consequently the $\delta(\ell_0, \Gamma)$-cusp residual set $W^s_{\delta(\ell_0,\Gamma)}$) depends on the choice of resonant lines $\Gamma_1, \dots, \Gamma_m$. However, they are independent of behavior of the diffusing trajectory $(\theta, \dot\theta)(t)$.

In particular, this implies the following result. Consider a finite collection of non-void open subsets $\Omega_1, \dots, \Omega_{m+1}$ of U, then there is a resonant piecewise linear curve Γ, consisting of m resonant segments $\Gamma = \cup_{s=1}^m \Gamma_s$ connecting distinct Ω_k's in any pre-assigned order. Then, by Mather Diffusion Theorem there is a trajectory which visits the sets $\Omega_1, \dots, \Omega_{m+1}$ in the pre assigned order.

We point out that existence of a "diffusing" trajectory $(\theta, \dot\theta)(t)$ moving along any prescribed resonant piecewise linear curve is a strong form of Arnold diffusion. However, existence of such a trajectory is proved only for a δ-cusp residual set of perturbations [**Ma4**].

The purpose of the next four sections is to define qualitatively the function $\delta(\ell_0, \Gamma)$ and the sets $U^s_{\delta(\ell_0,\Gamma)}$ and $W^s_{\delta(\ell_0,\Gamma)}$ mentioned in definition 4.1 and Mather Diffusion Theorem. We start by defining two averaged mechanical systems $L_{\omega,\Lambda}$ and L_ω. For sake of brevity, we shall not say precisely in what sense the trajectories of $L_{\omega,\Lambda}$ and L_ω approximate certain trajectories of L. We also give an heuristic motivation of the notion of these averaged systems $L_{\omega,\Lambda}$ and L_ω. These averaged mechanical systems are used to define a C^s open and dense set $U^s_{\delta(\ell_0,\Gamma)}$ of "good directions" of perturbations on the unit sphere $\mathcal{S}^s$. In Section 7 we define $\delta(\ell_0, \Gamma)$-cusp residual set $W^s_{\delta(\ell_0,\Gamma)}$ using barrier functions. In Section 8 we restate Mather Diffusion Theorem in terms of a certain variational principle. Finally, in Section 9 we apply this variational principle to prove our main result (Theorem 1.2).

5. Averaged mechanical systems corresponding to single and double resonances

5.1. A Single Resonance Averaged System or a First (ω, Λ)-Averaged System. — Let $L(\theta, \dot\theta, t) = \ell_0(\dot\theta) + \varepsilon P(\theta, \dot\theta, t)$ be a C^s small perturbation of an integrable Lagrangian ℓ_0 on $\mathbb{T}^2 \times U \times \mathbb{T}$, $s \geqslant 3$. Let us assume $d^2\ell_0 > 0$ on U. Consider a resonant frequency vector $\omega = (\omega_1, \omega_2) \in \mathbb{R}^2$ and its resonance $k = (k_0, k_1, k_2)$. This means that $k \in \mathbb{Z}^3$, $(k_1, k_2) \neq 0$ and $k_0 + k_1\omega_1 + k_2\omega_2 = 0$. If $\omega \in \mathbb{Q}^2$, it admits two linearly independent resonances; otherwise, it admits at most one resonance up to multiplication by scalar.

We denote by $\Lambda = \Lambda_k$ the resonant line from (H1). Thus, Λ is the set of all $\omega \in \mathbb{R}^2$ for which k is a resonance. We set

$$\mathbb{T}^2_\Lambda = \left\{(\theta_1, \theta_2, t) \in \mathbb{T}^2 \times \mathbb{T}^1 : \ k_1\theta_1 + k_2\theta_2 + k_0 t = 0 \ (\mathrm{mod}\ 1)\right\}. \tag{14}$$

If $\dot\theta(0) \in \Lambda$, the trajectory of the unperturbed Euler-Lagrange of $\ell_0(\dot\theta)$ either belongs to $\mathbb{T}^2_\Lambda$ or to its parallel translation. Thus, the 2-torus $\mathbb{T}^2_\Lambda$ can be viewed as a subgroup of $\mathbb{T}^2 \times \mathbb{T}^1$. We set $\mathbb{T}^1_\Lambda = \mathbb{T}^2 \times \mathbb{T}^1 / \mathbb{T}^2_\Lambda$ (and we refer to it as the factor space). Since the unperturbed Euler-Lagrange flow is parallel to $\mathbb{T}^2_\Lambda$, we call $\mathbb{T}^2_\Lambda$ — torus of *fast* motion and $\mathbb{T}^1_\Lambda$ — torus of *slow* motion.

Let $(\theta_1, \theta_2, t) = (\varphi^s_\Lambda, \varphi^f_\Lambda) \in \mathbb{T}^1_\Lambda \times \mathbb{T}^2_\Lambda$ denote slow and fast coordinates on $\mathbb{T}^1_\Lambda$ and $\mathbb{T}^2_\Lambda$, respectively. The product decomposition depends on an arbitrary choice. We shall specify our choice later (see lemma 6.3). Denote by $d\mathcal{H}_\Lambda$ the normalized Haar (Lebesgue) measure on the fast torus $\mathbb{T}^2_\Lambda$. Let $\widetilde{P}(\theta, t, \omega) = P(\theta, \omega, t)$. Define *the first* (ω, Λ)*-averaged potential*

$$P_{\omega,\Lambda}(\varphi^s_\Lambda) = \int_{\varphi^f_\Lambda} \widetilde{P}(\varphi^s_\Lambda, \varphi^f_\Lambda, \omega)\, d\mathcal{H}_\Lambda(\varphi^f_\Lambda). \tag{15}$$

So $P_{\omega,\Lambda} : \mathbb{T}^1_\Lambda \to \mathbb{R}$ is a real valued function on $\mathbb{T}^1_\Lambda$.

To define the first (ω, Λ)-averaged kinetic energy one needs some linear algebra. Actually, the precise form of this kinetic energy is not important for us. What really matters is that the kinetic energy is given by a constant quadratic form on $T(\mathbb{T}^1_\Lambda)$.

Consider the natural projection $\pi_\Lambda : \mathbb{T}^2 \times \mathbb{T}^1 \to \mathbb{T}^1_\Lambda$ along the fast torus $\mathbb{T}^2_\Lambda$. The definition of both slow and fast tori $\mathbb{T}^1_\Lambda$ and $\mathbb{T}^2_\Lambda$ depends only on the resonance k determining $\Gamma \subset \Lambda_k$. The projection π_Λ induces a linear map $d\pi_\Lambda : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}_\Lambda$. The restriction to $\mathbb{R}^2 \times 0$ has a null space, denoted by $N_\Lambda \subset \mathbb{R}^2$. Denote by $N^\perp_\Lambda$ the orthogonal complement of N_Λ with respect to $d^2\ell_0(\omega)$. Define

$$K_{\omega,\Lambda} = (d^2\ell_0(\omega)/2)|_{N^\perp_\Lambda}. \tag{16}$$

$K_{\omega,\Lambda}$ may be regarded as a constant quadratic form on $T\ (\mathbb{T}^1_\Lambda)$ in a view of identification of N_Λ and $T\ (\mathbb{T}^1_\Lambda)$ given by $d\pi_\Lambda$.

Let

$$L_{\omega,\Lambda} = K_{\omega,\Lambda} + P_{\omega,\Lambda} \circ \mathrm{pr}_{\omega,\Lambda} : T\ (T^1_\Lambda) \longrightarrow \mathbb{R},$$

where $pr_{\omega,\Lambda} : T\ (T^1_\Lambda) \to \mathbb{T}^1_\Lambda$. We call $L_{\omega,\Lambda}$ *the first* (ω, Λ)*-averaged system associated to* $\omega \in \Lambda$, which is an autonomous mechanical system whose kinetic energy is $K_{\omega,\Lambda}$ and whose potential energy is $-P_{\omega,\Lambda} \circ \mathrm{pr}_{\omega,\Lambda}$.

A classical idea of averaging consists in the fact that the trajectories of L with approximate frequency vector ω can be approximately described in terms of fast and slow variables. The fast variables correspond to the motion parallel to $\mathbb{T}^2_\Lambda$ and the slow variables correspond to the motion normal (in a suitable sense) to $\mathbb{T}^2_\Lambda$. If we average with respect to the fast variables, we obtain a new Lagrangian system $L_{\omega,\Lambda} = K_{\omega,\Lambda} + P_{\omega,\Lambda}$ whose trajectories approximate the trajectories of L with approximate rotation vector ω.

5.2. A Double Resonance Averaged System or a Second ω-Averaged System Associated to a Rational Frequency.

— Following the notation introduced here above, we let $L(\theta, \dot\theta, t) = \ell_0(\dot\theta) + \varepsilon P(\theta, \dot\theta, t)$ be a C^s small perturbation of an integrable Lagrangian ℓ_0 on $\mathbb{T}^2 \times U \times \mathbb{T}$, and we assume that $d^2\ell_0 > 0$ on U. Consider a rational frequency vector $\omega = (\omega_1, \omega_2) = (p_1/q, p_2/q) \in \mathbb{Q}^2$ and assume that $(p_1/q, p_2/q)$ is the reduced form, *i.e.* the greatest common divisor of integer p_1, p_2, and q is 1. For the unperturbed integrable system ℓ_0, every trajectory with rotation vector $\dot\theta \equiv \omega$ is closed and parallel to the 1-torus

$$\mathbb{T}^1_\omega = \{(\lambda p_1, \lambda p_2, \lambda q) \in \mathbb{T}^2 \times \mathbb{T}^1 : \ \lambda \in \mathbb{R}\}. \tag{17}$$

Since $\mathbb{T}^2 \times \mathbb{T}^1$ is an Abelian group, $\mathbb{T}^1_\omega$ may be considered as a subgroup. Let $\mathbb{T}^2_\omega = \mathbb{T}^2 \times \mathbb{T}^1/\mathbb{T}^1_\omega$ be the 2-torus obtained as a coset of $\mathbb{T}^1_\omega$. Similarly to the previous section, we call $\mathbb{T}^1_\omega$ — *fast* and $\mathbb{T}^2_\omega$ — *slow torus* respectively. Let $(\varphi^s_\omega, \varphi^f_\omega) \in \mathbb{T}^2_\omega \times \mathbb{T}^1_\omega$ denote slow and fast coordinates in $\mathbb{T}^2_\omega$ and $\mathbb{T}^1_\omega$, respectively. Let $\widetilde{P}(\theta, t, \omega) = P(\theta, \omega, t)$. Denote by $d\mathcal{H}_\omega$ the normalized Haar (Lebesgue) measure on the 1-torus $\mathbb{T}^1_\omega$. Recall that $\widetilde{P}(\theta, t, \omega) = P(\theta, \omega, t)$. Define *the second ω-averaged potential*

$$P_\omega(\varphi^s_\omega) = \int_{\varphi^f_\omega} \widetilde{P}(\varphi^s_\omega, \varphi^f_\omega, \omega)\, d\mathcal{H}_\omega(\varphi^f_\omega). \tag{18}$$

Note that $P_\omega : \mathbb{T}^2_\omega \to \mathbb{R}$ is a real valued function. We need now some linear algebra in order to define the second ω-averaged kinetic energy. Consider the natural projection $\pi_\omega : \mathbb{T}^2 \times \mathbb{T}^1 \to \mathbb{T}^2_\omega$ along fast torus $\mathbb{T}^1_\omega$. The definition of both $\mathbb{T}^1_\omega$ and π_ω depends only on $(q, p_1, p_2) \in \mathbb{Z}^3$, where $\omega = (p_1/q, p_2/q)$. The projection π_ω induces a linear map $d\pi_\omega : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^2_\omega$. The restriction to $\mathbb{R}^2 \times 0$ becomes an isomorphism. Since ℓ_0 is a C^2 smooth function on $\mathbb{R}^2$, its Hessian $d^2\ell_0(\omega)$ can be regarded as a quadratic form on $\mathbb{R}^2$. We define

$$K_\omega = d^2\ell_0(\omega)/2 \tag{19}$$

and we shall identify $\mathbb{R}^2$ with $T\,(\mathbb{T}^2_\omega)$ via $d\pi_\omega$. Let also define

$$L_\omega = K_\omega + P_\omega \circ \mathrm{pr}_\omega : T\,(\mathbb{T}^2_\omega) \longrightarrow \mathbb{R}, \tag{20}$$

where $\mathrm{pr}_\omega : T\,(\mathbb{T}^2_\omega) \to \mathbb{T}^2_\omega$ is the natural projection. We call L_ω *the second ω-averaged system associated with the rational frequency* $\omega = (\omega_1, \omega_2) = (p_1/q, p_2/q) \in \mathbb{Q}^2$. This is an autonomous mechanical system whose kinetic energy is K_ω and whose potential energy is $-P_\omega \circ \mathrm{pr}_\omega$.

6. Definition of $U^s_{\delta(\ell_0,\Gamma)}$

6.1. Part I: Building blocks. — In this Section, we begin the definition of the set of admissible directions $U^s_{\delta(\ell_0,\Gamma)}$ on the unit sphere of perturbations S^s_L or, equivalently, qualitative definition of $\delta(\ell_0, \Gamma)$. Later, we use this to define a $\delta(\ell_0, \Gamma)$-cusp residual set $W^s_{\delta(\ell_0,\Gamma)}$, where Mather Diffusion Theorem holds. We need it for the application of Mather Diffusion Theorem to our main result Theorem 1.2. The set $U^s_{\delta(\ell_0,\Gamma)}$ implicitly appears in Mather Diffusion Theorem and it is defined as a set where a non–negative functional $\delta(\ell_0, \Gamma)$ is positive. We shall not give here a complete definition $\delta(\ell_0, \Gamma)$.

since this would need quite a long discussion. We shall only sketch some qualitative aspects of its definition. For the discussion of the size of $\delta(\ell_0, \Gamma)$ we refer to [**Ma4**].

Step 1. — Consider a resonant piecewise linear curve $\Gamma = \cup_{s=1}^{m} \Gamma_s \subset B^2$ consisting of m resonant line segments $\Gamma_s \subset \Lambda_s = \{(\omega_1, \omega_2) \in B^2 : \; k_0^s + k_1^s\omega_1 + k_2^s\omega_2 = 0\}$, $k^s = (k_0^s, k_1^s, k_2^s) \in \mathbb{Z}^3$ as in definition 4.2.

Step 2. — For each resonant segment Γ_s we associate a non-negative function $\delta(\ell_0, \Gamma_s) : S_L^s \to \mathbb{R}_+$, defined in the next Section. Then, $\delta(\ell_0, \Gamma) = \min_{s=1}^{m} \delta(\ell_0, \Gamma_s)$.

Now we discuss the qualitative part of the definition of $\delta(\ell_0, \Gamma_s)$ for one segment. For the sake of simplicity, we omit the subindex s in the sequel, hence, in what follows, Γ will denote a single resonant segment. For one segment, we state a finite collection of non–degeneracy hypotheses of two types. Each hypothesis turns out to be fulfilled generically [**Ma4**].

Type 1. — Non–degeneracy of the 1-parameter family of the first (ω, Γ)-averaged mechanical system $\{L_{\omega,\Gamma}\}_{\omega\in\Gamma}$ on $T\,(\mathbb{T}_\Gamma^1)$.

Type 2. — Γ non–degeneracy of the second ω-averaged mechanical system L_ω on $T\,(\mathbb{T}_\omega^2)$ associated to a rational frequency $\omega \in \Gamma \cap \mathbb{Q}^2$.

There are countably many rationals ω's in any resonant segment Γ. However, we need type 2 non–degeneracy only for finitely many rational ω's. At the end of Section 6.3, we define a marginal denominator $q_0 = q_0(\ell_0, P, \Gamma_s)$ with the following meaning. Let $\omega = (p_1/q, p_2/q)$ be in the reduced form, then we need to impose non–degeneracy hypotheses of type 2 on ω only if $q < q_0$. In the next two Sections, we define the non-degeneracy hypotheses of type 1 on the family $\{L_{\omega,\Lambda}\}_{\omega\in\Lambda}$ of the first (ω, Λ)-averaged system and of type 2 on the second ω-averaged system L_ω, $\omega \in \Lambda \cap \mathbb{Q}^2$ along with $q_0(\ell_0, P, \Gamma_s)$ respectively.

6.2. Part II: Non-degeneracy of averaged systems associated to a single segment Γ. — By means of Step 2 of the last section, we see that it suffices to define $U^s_{\delta(\ell_0,\Gamma)}$ for one segment. Since ℓ_0 is fixed, we shall omit it from the notation and denote this set $U^s_{\delta(\Gamma)}$.

Let $\Lambda = \Lambda_k$ be the line that contains a bounded segment Γ. For $\omega \in \Gamma$, we write $P_{\omega,\Gamma}$ for the averaged function $P_{\omega,\Lambda}$ defined in section, and $\mathbb{T}_\Gamma^i$ for $\mathbb{T}_\Lambda^i$ ($i = 1, 2$). Thus, $\{P_{\omega,\Gamma} : \omega \in \Gamma\}$ is a C^s smooth 1-parameter family of functions defined on the circle $\mathbb{T}_\Gamma^1$. For εP to be in $U^s_{\delta(\Gamma)}$, we require that the global minima of $\{P_{\omega,\Gamma} : \omega \in \Gamma\}$ are of generic type. More precisely, we require the following three hypotheses to be fulfilled:

(C1) For each value $\omega \in \Gamma$, each global minimum of m_ω of $P_{\omega,\Gamma}$ is non–degenerate, *i.e.* $P''_{\omega,\Gamma}(m_\omega) > 0$.

(C2) For each $\omega \in \Gamma$, there are at most two global minima of $P_{\omega,\Gamma}$.

Let $\omega_0 \in \Gamma$ and suppose that $P_{\omega_0,\Gamma}$ has two global minima m_{ω_0} and m'_{ω_0}. We may continue these to local minima m_ω and m'_ω of $P_{\omega,\Gamma}$, for $\omega \in \Gamma$ near ω_0, in view of (C1). Thus, m_ω and m'_ω depend continuously on ω and they are the given global minima for $\omega = \omega_0$. In addition to (C1) and (C2), we require that the following *first transversality condition* be fulfilled:

(C3) $$\frac{dP_{\omega,\Gamma}(m_\omega)}{d\omega}\Big|_{\omega=\omega_0} \neq \frac{dP_{\omega,\Gamma}\ (m'_\omega)}{d\omega}\Big|_{\omega=\omega_0}.$$

Next, we require ℓ_0 and P to fulfill some conditions on the second ω-averaged systems L_ω associated to $\omega \in \Gamma \cap \mathbb{Q}^2$, defined in Section 5.2. Such an ω has the form $\omega = p/q = (p_1/q, p_2/q)$, where $p = (p_1, p_2) \in \mathbb{Z}^2$, and $q \in \mathbb{Z}$, $q > 0$. If p/q is in the reduced form, *i.e.* 1 is the greatest common *denominator* of p_1, p_2 and q, then we say that q is the denominator of ω. We shall require the remaining hypotheses only in the case ω has *small denominator*, *i.e.* $q \leqslant q_0$, where $q_0 = q_0(\ell_0, P, \Gamma)$ is a positive integer depending on ℓ_0, P, and Γ. The definition of q_0 is the quantitative aspect of the definition of $U^s_{\delta(\Gamma)}$ that we shall postpone to Section 6.3.

The first condition we require L_ω to fulfill is a condition on P_ω alone:

$(C4)_\omega$ The function P_ω on $\mathbb{T}^2_\omega$ has only one global minimum m_ω and it is non degenerate in the sense of Morse, *i.e.* the quadratic form $d^2P_\omega(m_\omega)$ is non singular.

To state the remaining hypotheses, we need to define a special homology element $h_{\omega,\Gamma}$ of $H_1(\mathbb{T}^2_\omega;\mathbb{R})$:

Since $\omega \in \Gamma \cap \mathbb{Q}^2$, we have $\mathbb{T}^1_\omega \subset \mathbb{T}^2_\Gamma$, so $\mathbb{T}^2_\Gamma/\mathbb{T}^1_\omega$ is a circle in $\mathbb{T}^2_\omega$, and

$$\mathbb{Z} \approx H_1(\mathbb{T}^2_\Gamma/\mathbb{T}^1_\omega;\mathbb{Z}) \subset H_1(\mathbb{T}^2_\omega;\mathbb{Z}) \subset H_1(\mathbb{T}^2_\omega;\mathbb{R}).$$

We let $h_{\omega,\Gamma}$ be a generator of $H_1(\mathbb{T}^2_\Gamma/\mathbb{T}^1_\omega;\mathbb{Z})$. In view of the inclusions above, this is an element of $H_1(\mathbb{T}^2_\omega;\mathbb{R})$. Geometrically, the above situation has the following meaning. Consider a circle $l_{\omega,\Gamma} \subset \mathbb{T}^2_\omega$ in the homology class $h_{\omega,\Gamma}$ and take $\pi_\omega^{-1}(l_{\omega,\Gamma}) \subset \mathbb{T}^2 \times \mathbb{T}$, where the projection $\pi_\omega : \mathbb{T}^2 \times \mathbb{T} \to \mathbb{T}^2_\omega$ is defined in Section 5.2. Therefore, $h_{\omega,\Gamma}$ is such that $\pi_\omega^{-1}(l_{\omega,\Gamma})$ is parallel to $\mathbb{T}^2_\Lambda \subset \mathbb{T}^2 \times \mathbb{T}$.

The Lagrangian L_ω describes a conservative mechanical system, *i.e.* it has the form *kinetic energy* – *potential energy*. Here, the kinetic energy K_ω is associated to the constant Riemannian metric $g_\omega = d^2\ell_0(\omega)$ on $\mathbb{T}^2_\omega$. The potential energy is $-P_\omega \circ pr_\omega$. By a slight abuse of terminology, we shall shorten this to $-P_\omega$.

Next conditions that we require on L_ω are easily described in terms of the Maupertuis principle:

We let $E_\omega = -P_\omega(m_\omega)$, where $m_\omega \in \mathbb{T}^2_\omega$ is the unique minimum of P_ω, as above. For any $E \geqslant E_\omega$, we let

(21) $$g_E = 2(P_\omega + E)K_\omega.$$

For $E > E_\omega$, the function $P_\omega + E$ is everywhere positive on $\mathbb{T}^2_\omega$. Hence, g_E is a C^s Riemannian metric on $\mathbb{T}^2_\omega$. For $E = E_\omega$, the function $P_\omega + E$ is positive except at m_ω, where it vanishes and has a non degenerate minimum.

The Maupertuis principle states that trajectories of L_ω having energy E are the same as geodesics of g_E, except for a time reparametrization. Carneiro [**DC**] has extended the Maupertuis principle and shown that absolute minimizers of L_ω having energy E correspond to class A geodesics of g_E (in the sense of Morse [**Mor**] and Hedlund [**Hed**]).

Pick a large energy constant $E^* = E^*(\ell_0, P, \Gamma) > 0$. The next condition that we impose on L_ω concerns the shortest closed geodesics of g_E in the homology class $h_{\omega,\Gamma}$, for $E_\omega + E^* \geqslant E \geqslant E_\omega$. Briefly, we require that these are of generic type. More explicitly, we ask that the following four hypotheses are fulfilled:

$(\mathrm{C5})_\omega$ For $E_\omega + E^* \geqslant E \geqslant E_\omega$, each shortest closed geodesic of g_E in the homology class $h_{\omega,\Gamma}$ is non degenerate in the sense of Morse.

$(\mathrm{C6})_\omega$ For $E_\omega + E^* \geqslant E \geqslant E_\omega$, there are at most two shortest closed geodesics of g_E in the homology class $h_{\omega,\Gamma}$.

Let $E_1 > E_\omega$ and suppose that there are two shortest geodesics γ and γ' of g_{E_1} in the homology class $h_{\omega,\Gamma}$. We may continue these to locally shortest geodesics γ_E and γ'_E of g_E for E near E_1, in view of $(\mathrm{C5})_\omega$. If μ is a closed curve on $\mathbb{T}^2_\omega$, we let $\ell_E(\mu)$ denote its length with respect to g_E. We require that the following *second transversality condition* be fulfilled:

$$(\mathrm{C7})_\omega \qquad \frac{d(\ell_E(\gamma_E))}{dE}\Big|_{E=E_1} \neq \frac{d(\ell_E(\gamma'_E))}{dE}\Big|_{E=E_1}.$$

These are the hypotheses that we require g_E to fulfill when $E_\omega + E^* \geqslant E \geqslant E_\omega$. Note that the case $E = E_\omega$ is somehow special, because g_{E_ω} is not a Riemannian metric, since it vanishes at m_ω[4]. Nevertheless, we may define the length of a curve with respect to g_{E_ω} just as one normally defines the length of a curve with respect to a Riemannian metric. We define a geodesic to be a curve that is the shortest distance between any two sufficiently nearby points. It is easy to see that there exists a shortest geodesic of g_{E_ω} in the homology class $h_{\omega,\Gamma}$. We require L_ω to fulfill the following condition:

$(\mathrm{C8})_\omega$ There is only one shortest geodesic γ of g_{E_ω} in the homology class $h_{\omega,\Gamma}$, and γ is non degenerate in the sense of Morse.

In saying that a g_E shortest geodesic γ is non degenerate in the sense of Morse, we mean the following:

Let μ be a transversal to γ, intersecting γ in one point, not m_ω, in the case that $E = E_\omega$. For each point $P \in \mu$, let γ_P be the g_E shortest curve through P in the homology class $h_{\omega,\Gamma}$ and let $\ell_E(\gamma_P)$ denote its g_E length. The function $P \to \ell_E(\gamma_P)$ is C^s near $\mu \cap \gamma$ and the condition that γ be non degenerate means that its second derivative is positive.

In the case that $E > E_\omega$, this is the usual notion of non degeneracy in the sense of Morse.

[4] This corresponds to a periodic trajectory for the Euler-Lagrangian flow of $L = \ell_0 + \varepsilon P$.

Definition 6.1. — $U^s_{\delta(\Gamma)}\big(= U^s_{\delta(\ell_0,\Gamma)}\big) = \{\varepsilon P : \varepsilon > 0, P \in \mathcal{P}^s$, and P satisfies hypotheses (C1)–(C3) as well as hypotheses $(C4)_\omega$–$(C8)_\omega$ for $\omega \in \Gamma \cap \mathbb{Q}^2$ with small denominator, *i.e.* such that $q \leqslant q_0(\ell_0, p, \Gamma)$, where q denotes the denominator of ω.$\}$

Remark 6.2. — This definition can be considered as an implicit definition of $\delta(\ell_0, \Gamma)$.

6.3. What denominators are small?— In this Section, we define the marginal denominator $q_0 = q_0(\ell_0, P, \Gamma)$ from the previous definition. This would answer the question for which rational ω's we need to verify the non-degeneracy hypotheses $(C4)_\omega$–$(C8)_\omega$. Recalling (14) and (17), we associate to a rational frequency ω and a resonant segment $\Gamma \ni \omega$ two decompositions of $\mathbb{T}^2 \times \mathbb{T}^1$ into (the standard) direct product, and we denote the result of this operation by $\mathbb{T}^2_\Gamma \times \mathbb{T}^1_\Gamma$ and $\mathbb{T}^1_\omega \times \mathbb{T}^2_\omega$. These decompositions can be defined by changing the basis on $\mathbb{T}^2 \times \mathbb{T}^1$. Based on the lemma below we can define the following decomposition $\mathbb{T}^1_\omega \times \mathbb{T}^1_{\omega,\Gamma} \times \mathbb{T}^1_\Gamma = \mathbb{T}^2 \times \mathbb{T}^1$ into a direct sum.

Lemma 6.3. — *There is a choice of these decompositions so that* $\mathbb{T}^1_\Gamma \subset \mathbb{T}^2_\omega$, $\mathbb{T}^1_\omega \subset \mathbb{T}^2_\Gamma$.

Proof. — It seems easiest to discuss this in terms of a short sequence of topological abelian groups.

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0.$$

Thus, A is a topological subgroup of B, and C is a quotient group of B. Denote the inclusion of A into B by i bad the projection of B onto C by j. To say that the sequence is short exact means that the kernel of j is $i(A)$.

A splitting of such a sequence is given by a continuous homomorphism k of C into B such that $j \circ k$ is the identity. Equally well, it can be given by a continuous homomorphism l of B into A. such that $l \circ i$ is the identity. The relation between k and l is that the kernel of l is $k(C)$.

Given k (resp. l) there is a unique l (resp. k) such that this relation holds. Given such a splitting, B is the direct sum of $i(A)$ and $k(C)$.

There is a splitting, in fact many, for both of the short exact sequences in the case we consider (it suffices A to be a torus). Indeed, notice that when we consider $\mathbb{T}^1_\Gamma$ (resp. $\mathbb{T}^2_\omega$) as a subgroup $\mathbb{T}^2 \times \mathbb{T}$ we choose a splitting of the appropriate one of the two exact sequences in question(5). This proves the lemma. □

Below we present a test to determine q_0. The idea of the test is to check how dense the unperturbed closed trajectory $\dot\theta = \omega$ in the 2-torus $\mathbb{T}^2_\Gamma$. A precise definition is in terms of averaged systems $L_{\omega,\Gamma} = K_{\omega,\Gamma} + P_{\omega,\Gamma}$ and $L_\omega = K_\omega + P_\omega$ on slow tori $T\ (\mathbb{T}^1_\Gamma)$ and $T\ (\mathbb{T}^2_\omega) = T\ (\mathbb{T}^1_{\omega,\Gamma} \times \mathbb{T}^1_\Gamma)$, respectively. In the notation of the previous section, let m_ω and m'_ω be global (or local continuation of global) minima of P_ω on $\mathbb{T}^1_\Gamma$ and let

(5)Note that the inclusion of $\mathbb{T}^1_\Gamma$ in $\mathbb{T}^2_\omega$ depends on the choice of splitting. One has inclusion for some splittings not for others.

$\pi_{\omega,\Gamma} : \mathbb{T}^1_{\omega,\Gamma} \times \mathbb{T}^1_\Gamma \to \mathbb{T}^1_\Gamma$ be the natural projection. For each energy $E_\omega + E^* \geqslant E \geqslant E_\omega$, we need that the shortest geodesics γ_E and γ'_E (if it exists) of g_E are contained in a small neighborhood of $\pi^{-1}_{\omega,\Gamma}(m_\omega \cup m'_\omega)$ (or $\pi^{-1}_{\omega,\Gamma}(m_\omega)$ if m'_ω does not exist). The precise definition is as follows.

On the slow torus $\mathbb{T}^1_\Gamma$, we define the 1-parameter family of first (ω, Γ)-averaged potentials $\{P_{\omega,\Gamma} : \mathbb{T}^1_\Gamma \to \mathbb{R}\}_{\omega\in\Gamma}$. Suppose hypotheses (C1)–(C3) are fulfilled. By the first transversality condition (C3), there are finitely many $\omega_0 \in \Gamma$ with $P_{\omega_0,\Gamma}$ having two global minima m_{ω_0} and m'_{ω_0}. Mark these ω_0's. By (C1) and (C2), for ω's nearby the marked ω_0's, there is a smooth continuation m_ω and m'_ω of m_{ω_0} and m'_{ω_0}, respectively, to local minima nearby. Pick a small $\eta > 0$, so that η-neighborhoods of the marked ω_0's, denoted by $\Upsilon^\eta_{\omega_0}$, are disjoint. Moreover, in each neighborhood $\Upsilon^\eta_{\omega_0}$, there is a well defined continuation m_ω and m'_ω. Such η will be called (ℓ_0, P, Γ)-*admissible.*

Consider now a small $\tau > 0$ with the following properties. For each $\omega \in \Gamma$, consider two cases. In the first case ω is in one of $\Upsilon^\eta_{\omega_0}$. Then, we define a 2-tuple of τ-neighborhoods $\widetilde{D}^\tau_\omega$ and $\widehat{D}^\tau_\omega$ in $\mathbb{T}^1_\Gamma$ are centered at m_ω and m'_ω respectively and disjoint. Denote $D^\tau_\omega = \widetilde{D}^\tau_\omega \cup \widehat{D}^\tau_\omega$. In the other case, ω is outside of neighborhoods of marked frequencies $\Upsilon^\eta_{\omega_0}$'s put D^τ_ω to be a τ-neighborhood centered at the global minimum m_ω.

Definition 6.4. — A rational frequency $\omega \in \Gamma \cap \mathbb{Q}^2$ is $(\ell_0, P, \Gamma, \eta, \tau)$-admissible with a small $\tau > 0$ if the family of first (ω, Γ)-averaged systems $\{L_{\omega,\Gamma} = K_{\omega,\Gamma} + P_{\omega,\Gamma}\}_{\omega\in\Gamma}$ satisfy hypotheses (C1)–(C3) and for an (ℓ_0, P, Γ)-admissible $\eta > 0$ and any $E_\omega + E^* \geqslant E \geqslant E_\omega$ each shortest geodesic (resp. local continuation of a shortest geodesic) γ_E (resp. γ'_E if it exists) of the Maupertuis metric g_E in the homotopy class $h_{\omega,\Gamma}$ belongs to $\pi^{-1}_{\omega,\Gamma}(D^\tau_\omega)$.[(6)]

Recall that, for each double resonance of a rational frequency $\omega = (p_1/q, p_2/q) \in \Gamma \cap \mathbb{Q}^2$ in (20), we may define the double resonant mechanical system L_ω on the slow 2-dimensional torus $\mathbb{T}^2_\omega$ and the natural projection $\pi_{\omega,\Gamma} : \mathbb{T}^2_\omega \to \mathbb{T}^1_\Lambda$ onto the slow 1-dimensional torus $\mathbb{T}^1_\Gamma \subset \mathbb{T}^2_\omega$. Then, we have the following result:

***Lemma 6.5* ([Ma4])**. — *Suppose the perturbation term $P(\theta, \dot\theta, t)$ satisfies hypotheses (C1)–(C3). Then, for any $\tau > 0$, there is an integer $q_0 = q_0(\ell_0, P, \Gamma, \tau)$, such that, for any rational frequency ω with denominator $q > q_0$, we have that ω is $(\ell_0, P, \Gamma, \eta, \tau)$-admissible. Namely, a corresponding shortest geodesic (resp. local continuation of a shortest geodesic) γ_E (resp. γ'_E if exists) of the Maupertuis metric g_E, defined in (21), belongs to the strip $\pi^{-1}_{\omega,\Gamma}(D^\tau_\omega) \subset \mathbb{T}^2_\omega$.*

---

(6) As a matter of fact the proof in [**Ma4**] requires a stronger form of admissibility which still fits into the proof of our main result (Theorem 1.2).

For further reference, we need to give a definition of η and τ-neighborhoods for double resonances. Let L_ω and $\mathbb{T}^2_\omega$ be the mechanical Lagrangian on the 2-torus corresponding to a rational frequency $\omega = (p_1/q, p_2/q) \in \Gamma \cap \mathbb{Q}^2$ as above. Let $\{g_E\}_{E\in[E_\omega+E^*,E_\omega]}$ be the 1-parameter family of Maupertuis metrics defined by (21). Suppose hypotheses $(C4)_\omega$–$(C8)_\omega$ are fulfilled. Mark parameters E_0 where g_E has two shortest geodesics in the homology class $h_{\omega,\Gamma}$. By the second transversality condition $(C7)_\omega$, there are finitely many $E_0 \in [E_\omega + E^*, E_\omega]$ with metrics g_{E_0} having two shortest geodesics γ_{E_0} and γ'_{E_0}. By $(C8)_\omega$, there is a smooth continuation γ_{E_1} and γ'_{E_1} to locally shortest geodesics. Pick a small $\eta_\omega > 0$ so that η_ω-neighborhoods of marked E_0's (denoted by $\Upsilon^{\eta_\omega}_{E_0}$) are disjoint. Moreover, in each neighborhood $\Upsilon^{\eta_\omega}_{E_0}$ there is a well defined continuation γ_E and γ'_E. Such η_ω is called *$(\ell_0, P, \Gamma, \omega)$-admissible.*

Pick a small $\tau_\omega > 0$ with the following properties. For each $E\in[E_\omega, E_\omega + E^*]$ consider two cases. Either ω is in one of $\Upsilon^{\eta_\omega}_{E_0}$. Then, we define a 2-tuple of τ-neighborhoods $\widetilde{D}^\tau_E$ and $\widehat{D}^\tau_E$ in $\mathbb{T}^2_\omega$ of the locally shortest geodesics γ_{E_0} and γ'_{E_0} respectively so that these neighborhoods and disjoint. Denote $D^\tau_E = \widetilde{D}^\tau_E \cup \widehat{D}^\tau_E$. In the other case E is outside of these neighborhoods of marked energies, then D^τ_E is a τ-neighborhood of the shortest geodesic γ_E.

7. Definition of $W^s_{\delta(\ell_0,\Gamma)}$ using type 2 non-degeneracy (of Barrier functions)

In this Section, we define the non-degeneracy hypotheses of the second type. They are formulated in terms of minima of certain barrier functions, restricted to what we call *Poincaré screens.* First, we explain the meaning of Poincaré screens and we define them. Later on, we define required barrier functions and state the non-degeneracy hypotheses that we need to define $\delta(\ell_0, \Gamma)$-cusp residual set $W^s_{\delta(\ell_0,\Gamma)}$.

As mentioned in Section 3.2 diffusing trajectories stay most of the time close to the corresponding Mather sets $\mathcal{M}_i$ and from time to time make almost heteroclinic excursions along stable and unstable sets $W^s(\mathcal{M}_i)$ and $W^u(\mathcal{M}_{i+1})$ from one set $\mathcal{M}_i$ to the next one $\mathcal{M}_{i+1}$. In order to keep track of those excursions, we pose a smooth hypersurface (Poincaré screen) "in between" $\mathcal{M}_i$ and $\mathcal{M}_{i+1}$. To give a precise definition we need further discussion.

Recall that $\widetilde{\mathcal{M}}_i = \pi\mathcal{M}_i \subset \mathbb{T}^2 \times \mathbb{T}$ is the projected Mather sets. Suppose hypotheses (C1)–(C3) and $(C4)_\omega$–$(C8)_\omega$ for rational ω's with small denominator are fulfilled. Consider two different cases:

(1) ω is $C\sqrt{\varepsilon}$-close to a rational $(p_1/q, p_2/q)$ with small denominator $q < q_0$, where C is some positive constant depending only on ℓ_0, P, Γ, τ and is closely related to the energy constant E^*.

(2) the opposite case.

Recall that, for any frequency $\omega \in \Gamma$, we associate homology class in $H_1(\mathbb{T}^2, \mathbb{R}) \simeq \mathbb{R}^2$ equal to ω. Each Lagrangian satisfying conditions (1–3) of section 2.1 has Fenchel-Legendre transform $\mathcal{L}_\beta$ associated to it by (29). Using $\mathcal{L}_\beta$, we associate to each homology class ω any cohomology class c_ω inside $\mathcal{L}_\beta(\omega)$.

It turns out that, in the first case, for a sufficiently small $\varepsilon \neq 0$ and a cohomology $c \in \mathcal{L}_\beta(\omega)$, there is rescaling which relates c and $E \in [E_\omega, E_\omega + E^*]$, such that the projected Mather set $\widetilde{\mathcal{M}}^c$ belongs to $\pi_\omega^{-1}(D_E^{\tau_\omega})$. In the second case, the projected Mather set $\widetilde{\mathcal{M}}^c$ belongs to $\pi_\Gamma^{-1}(D_\omega^\tau)$. In both cases, the projected Mather sets are localized in a τ-neighborhood of one or two hypersurfaces on the base $\mathbb{T}^2 \times \mathbb{T}$. We shall distinguish these cases.

Definition 7.1. — Let $\omega \in \Gamma$ and $c \in \mathcal{L}_\beta(\omega)$ be a cohomology class. Distinguish two cases: $D_E^{\tau_\omega}$ (resp. D_ω^τ) has one or two components.

In the one component case, let us define $S^c \subset \mathbb{T}^2 \times \mathbb{T}$ to be a smooth hypersurface (*i.e.*, a codimension one closed smooth submanifold) topologically parallel $\pi_\omega^{-1}(\gamma_E)$, which is transversal to class A geodesics with respect to γ_E, and disjoint from its τ-neighborhood $\pi_\omega^{-1}(D_E^{\tau_\omega})$ in the first case and topologically parallel to $\pi_\Gamma^{-1}(m_\omega)$, transversal to class A geodesics with respect to γ_E and γ'_E, and disjoint from its τ-neighborhood $\pi_\Gamma^{-1}(D_\omega^\tau)$ in the second case.

In the two component case: let us define $S_-^c, S_+^c \subset \mathbb{T}^2 \times \mathbb{T}$ to be a pair of smooth hypersurfaces parallel and separating $\pi_\omega^{-1}(\gamma_E)$ and $\pi_\omega^{-1}(\gamma'_E)$ in the double resonance case. "Separating" means that S_-^c and S_+^c cut $\mathbb{T}^2 \times \mathbb{T}$ into two disjoint parts each containing either $\pi_\omega^{-1}(\gamma_E)$ or $\pi_\omega^{-1}(\gamma'_E)$. In the single resonance case define S_-^c, $S_+^c \subset \mathbb{T}^2 \times \mathbb{T}$ to be of smooth hypersurfaces parallel and separating $\pi_\Gamma^{-1}(m_\omega)$ and $\pi_\Gamma^{-1}(m'_\omega)$. Call S^c (resp. S_-^c and S_+^c) *Poincaré screen (resp. screens)* associated with cohomology class $c \in \mathcal{L}_\beta(\omega)$.

Hypotheses (C1-C8) we impose do not imply that the geodesic γ_E and the minimum m_ω vary continuously with E and $\omega \in \Gamma$. Points of discontinuity are usually call *bifurcations*. However, (C1) and (C4$_\omega$) imply that γ_E and m_ω vary piecewise continuously. Therefore, we can choose Poincaré screens so that they are piecewise constant with respect to c. In other words, one could divide Γ into a finite number of subintervals, so that for all ω in a subinterval the Poincaré screen is the same.

By construction, all S^c are topologically parallel. This property essentially relies on the fact that we have only *one* resonant segment Γ under consideration. Since S^c is piecewise constant in c, we shall treat the case of one Poincaré screen S^c for each c. The other case is analogous.

Denote by $S_i = S^{c_i}$ Poincaré screens corresponding to the frequencies ω_i, $i = 1, 2, \ldots$ related to c_i, $i = 1, 2, \ldots$ by Fenchel-Legendre transform respectively. We marked these frequencies ω_i's in Section 3.3. Consider a cyclic cover $\mathbb{T}_\Gamma^2 \times \mathbb{R}$ over $\mathbb{T}_\Gamma^2 \times \mathbb{T}_\Gamma^1$ obtained by cutting along a Poincaré screen S and unrolling. Fix one representative of $\mathbb{T}^2_\Gamma \times \mathbb{T}^1_\Gamma$ in $\mathbb{T}^2_\Gamma \times \mathbb{R}$ and denote it by $\mathbb{T}^2_\Gamma \times \mathbb{T}^1_\Gamma(0)$. For each integer k, we denote by $\iota_k : \mathbb{T}^2_\Gamma \times \mathbb{R} \to \mathbb{T}^2_\Gamma \times \mathbb{R}$ the deck transformation along $\mathbb{T}^1_\Gamma$-direction, and by $\mathbb{T}^2_\Gamma \times \mathbb{T}^1_\Gamma(k) = \iota_k(\mathbb{T}^2_\Gamma \times \mathbb{T}^1_\Gamma(0))$ its k-th shift. Denote by S_i^0 an image of S_i in $\mathbb{T}^2_\Gamma \times \mathbb{R}$ under the natural embedding so that $S_i^0 \cap (\mathbb{T}^2_\Gamma \times \mathbb{T}^1_\Gamma(0)) \neq \varnothing$. By construction, for each $i = 1, 2, \dots$ we have that the corresponding S_i is topologically parallel to $\mathbb{T}^2_\Gamma$ and that $\iota_k(S_i) \cap \iota_{k'}(S_i)$ are disjoint for any $k \neq k'$. Denote $S_i^k = \iota_k(S_i)$ for k, $i \in \mathbb{Z}$ and $\widehat{S}_i = S_i^i$. Now we define the $\delta(\ell_0, \Gamma)$-cusp residual set $W^s_{\delta(\ell_0,\Gamma)}$.

Consider a closed one form η, with $[\eta]_{\mathbb{T}^2} = c$ and $c \in \mathcal{L}_\beta(\Gamma)$. Define the barrier function on S_i

$$H_{\eta,T}((\theta,t),(\theta',t')) = \inf\left\{ \int_a^b (L - \widehat{\eta})(d\gamma(t),t)dt \right\}, \tag{22}$$

where the infimum is taken over all absolutely continuous curves $\gamma : [a,b] \to \mathbb{T}^2_\Gamma \times \mathbb{R}$ such that $\gamma(a) = \theta$, $\gamma(b) = \theta'$, $a \equiv t' \pmod 1$, $b \equiv t \pmod 1$, $b - a \geqslant T$, $(\theta,t) \in S_i^0$, $(\theta',t') \in S_i^1$.

For next definition, we need to introduce suitable curves $\iota_1(\theta,t) = (\theta',t')$, which correspond to closed curves on $\mathbb{T}^2_\Gamma \times \mathbb{T}^1_\Gamma$ connecting a point on S_i with itself and making only one turn in $\mathbb{T}^1_\Gamma$-direction. Notice that, in this case, $H_{\eta,T}$ is independent of the choice of η in $[\eta]_{\mathbb{T}^2} = c$ and $[\eta]_{\mathbb{T}}$, because such curves are closed. For a Mañé critical or subcritical form, the barrier function $H_{\eta,T}$ is finite and continuous [**Ma5**]. Let us consider

$$H_c(\theta,t) = \liminf_{T\to+\infty} H_{\eta,T}((\theta,t),(\theta,t)). \tag{23}$$

This definition is a particular case of the definition of barrier function (32). In [**Ma2**], Mather proved that the limit exists.

Lemma 7.2 **([Ma4])**. — *Let $P \in U^s_{\delta(\ell_0,\Gamma)}$, ε be sufficiently small and positive, and Γ be a resonant line segment in U. Then, the Fenchel-Legendre transform $\mathcal{L}_\beta$ associated with $L = \ell_0 + \varepsilon P$ by (29) has the channel property with respect to Γ.*

By lemma 7.2, there is a closed connected curve $\sigma_\Gamma \subset \mathcal{L}_\beta(\Gamma)$ with the channel property. We could parameterize this curve by a smooth parameter, say τ, *i.e.* $\widehat{\sigma}_\Gamma : [0,1] \to \sigma_\Gamma \subset \mathcal{L}_\beta(\Gamma)$. Thus, we can define a family of barrier functions

$$\{H_\tau : S_\tau \to \mathbb{R}\}_{\tau\in[0,1]}$$

by $c_\tau = \widehat{\sigma}_\Gamma(\tau)$, $S_\tau = S^{c_\tau}$ and $H_\tau(\theta,t) = H_{c_\tau} : S_\tau \to \mathbb{R}$. It turns out that, under our hypotheses, H_τ is continuous in τ and even satisfies certain modulus of continuity (see [**Ma4**]). As we pointed out above, hypersurfaces S_τ can be chosen to be smooth in τ. Recall that a closed subset D of a torus $\mathbb{T}^d$ is called *acyclic in* $\mathbb{T}^d$ if there is a neighborhood V of D in $\mathbb{T}^d$ such that the inclusion map $H_1(V,\mathbb{R}) \subset H_1(\mathbb{T}^d,\mathbb{R})$ is the zero map. Since the ambient manifold $\mathbb{T}^d$ is a torus, the above inclusion map is the zero map if and only if any closed curve in V is contractible. Let

$$D_\tau = \left\{(\theta, t) \in S_\tau : \ H_\tau(\theta, t) = \min_{(\theta', t') \in S_\tau} H_\tau(\theta', t')\right\}$$

the set where minimum of the barrier H_τ on S_τ is achieved. Recall that S_τ is diffeomorphic to the 2-dimensional torus. The last non–degeneracy hypothesis we require is the following:

(C9) For each $\tau \in [0, 1]$ the set $D_\tau \subset S_\tau$ is acyclic.

Suppose that there is a curve $\sigma_\Gamma \in \mathcal{L}_\beta(\Gamma)$ with channel property such that the family of barrier functions $\{H_\tau\}_{\tau \in [0,1]}$ satisfies hypothesis (C9); then, we denote the set of perturbation terms εP with this property by $W^s_{\delta(\ell_0,\Gamma)} \subset \mathbb{R}U^s_{\delta(\ell_0,\Gamma)}$. The following result is not trivial to prove:

***Lemma 7.3* ([Ma4]).** — *The set $W^s_{\delta(\ell_0,\Gamma)}$ is C^s open and dense in $V^\delta_{\delta(\ell_0,\Gamma)}$.*

The application of Mather theory to the instability of elliptic points requires the following lemmas about the localization of the velocity of the minimizers. Recall that $L(\theta, v, t) = \ell_0(v) + \varepsilon P(\theta, v, t)$ is C^s smooth nearly integrable Lagrangian, defined on $\mathbb{T}^2 \times U \times \mathbb{T}$. Let $\mathcal{L}_\beta$ be Fenchel-Legendre transform associated with L. Denote by $\pi_v : \mathbb{T}^2 \times U \times \mathbb{T} \to U$ the natural projection.

***Localization Lemma I*.** — *There is $C = C(\ell_0, P) > 0$ such that for any frequency $\omega \in U$ and any cohomology class $c \in \mathcal{L}_\beta(\omega)$ the Mañé set $\mathcal{N}^c(\supset \mathcal{M}^c)$ is contained in $\pi_v^{-1}(B^2_{C\sqrt{\varepsilon}}(\omega))$.*

In other words, velocity of minimizers with approximate velocity ω may differ from ω at most by $C\sqrt{\varepsilon}$.

Let the perturbation direction $P \in U^s_{\delta(\ell_0,\Gamma)}$, then, for any frequency $\omega \in \Gamma$ and any cohomology class $c \in \mathcal{L}_\beta(\omega)$, the Poincaré screen S^c, the barrier function H_c, and its minimum set $D^c \subset S^c$ are well defined.

***Localization Lemma II*.** — *The property D^c being acyclic depends only on the values of L inside $\pi_v^1(B^2_{C\sqrt{\varepsilon}}(\omega))$ with the same C as in the Localization Lemma I.*

These lemmas are a restatement of Lemma 3 in [**Ma5**] and they follow from it. Their proof is based on a careful perturbation analysis. First, one proves that, with the standard identification of $H_1(\mathbb{T}^2, \mathbb{R}) \simeq \mathbb{R}^2$, $H^1(\mathbb{T}^2, \mathbb{R}) \simeq \mathbb{R}^2$, Fenchel-Legendre transform $\mathcal{L}_\beta$ is $C\sqrt{\varepsilon}$-close to the map $\nabla_v \ell_0 : U \to \mathbb{R}^2$. Then, using a generalization of (12) to the case of arbitrary C^s smooth convex unperturbed integrable Lagrangian $\ell_0(\dot\theta)$ and the remark that $\langle \dot\theta + c, \dot\theta + c\rangle$ is non-negative, one shows that, if C is too large, c-minimality would be contradicted. See also [**BK**] for similar results.

8. Variational principle and restatement of Mather diffusion theorem

In this Section, we introduce a variational principle of Mather [**Ma4**]. We shall use the notation of the previous section. By lemma 7.2, we have that $\mathcal{L}_\beta(\Gamma)$ has a smooth connected curve $\sigma_\Gamma \subset \mathcal{L}_\beta(\Gamma)$ having channel property. Fix an orientation on σ_Γ toward $\mathcal{L}_\beta(\omega^0)$ and a sufficiently dense ordered set of cohomology classes $\mathfrak{C} = \{c_i\}_{i\in\mathbb{Z}_+} \subset \sigma_\Gamma$ so that they are monotonically oriented along σ_Γ and in between any two c_{i-1} and c_{i+1} on σ_Γ there is only c_i from $\mathfrak{C}$. How dense this set needs to be will depend on how close the family of barriers $\{H_\tau\}_{\tau\in[0,1]}$ defined above to fail hypothesis (C9). This collection of c_i's plays the role of the collection of ω_i's from Section 3.3. For each positive integer i, denote Poincaré screens by $S_i = S^{c_i}$ on $\mathbb{T}^2_\Gamma \times \mathbb{T}^1_\Gamma$ and by $\widehat{S}_i = \widehat{S}^{c_i}$ on $\mathbb{T}^2_\Gamma \times \mathbb{R}$, Mather sets by $\mathcal{M}_i = \mathcal{M}^{c_i} \subset \mathbb{T}^2 \times \mathbb{R}^2 \times \mathbb{T}$. Fix a sequence of closed one forms η_i such that $[\eta]_{\mathbb{T}^2} = c_i$ and positive numbers T_i. For $(\theta, t) \in \widehat{S}_i$, $(\theta', t') \in \widehat{S}_{i+1}$ and $T > 0$ define

$$H_{i,T_i}((\theta,t),(\theta',t')) = \inf\left\{\int_a^b (L-\widehat{\eta}_i)(d\gamma(t),t)dt\right\}, \tag{24}$$

where the infimum is taken over all absolutely continuous curves $\gamma : [a,b] \to \mathbb{T}^2_\Gamma \times \mathbb{R}$ such that $\gamma(a) = \theta$, $\gamma(b) = \theta'$, $a \equiv t \pmod 1$, $b \equiv t' \pmod 1$, $b - a \geqslant T_i$. This leads to a variational principle

$$\sum_{i\in J'} H_{i,T_i}((\theta_i,t_i),(\theta_{i+1},t_{i+1}))\,. \tag{25}$$

Here above, we understood the following notation: if J is a set of consecutive integers, we denote by J' the index set J without its largest element (provided it exists).

If all one forms η_i are critical or subcritical, then each H_i is finite and continuous. Therefore, we can define a *minimizer* of the variational principle to be a sequence $\{(\theta_i, t_i) : i \in J\}$ such that if $a < b$ and $\{(\theta'_i, t'_i) : i \in J\}$ is any sequence satisfying $(\theta'_i, t'_i) = (\theta_i, t_i)$ for $i \leqslant a$ and $i \geqslant b$, then

$$\sum_{i\in J'} H_{i,T_i}((\theta_i,t_i),(\theta_{i+1},t_{i+1})) \leqslant \sum_{i\in J'} H_{i,T_i}\left((\theta'_i,t'_i),(\theta'_{i+1},t'_{i+1})\right).$$

Since each H_i is finite and continuous, an elementary compactness argument shows the existence of a minimizer.

Theorem 8.1. — *Let $\varepsilon P \in W^s_{\delta(\ell_0,\Gamma)}$. Then, for any index set J, there are sequences of positive numbers $\{\varepsilon_i\}_{i\in J}$ and subcritical closed one forms $\{\eta_i\}_{i\in J}$, such that $[\eta_i]_{\mathbb{T}} - \alpha_L([\eta_i]_{\mathbb{T}^s}) < \varepsilon_i$, and large positive numbers $\{T_i\}_{i\in J}$ satisfying the following property: there exists a minimizer $\{d\gamma(t) : t \in \mathbb{R}\}$ of the variational principle (25), which provides a smooth solution of the Euler-Lagrange equation (7). In other words, the minimizer $\{d\gamma(t) : t\in\mathbb{R}\}$ has no corners.*

9. Application

In this Section, we describe how to apply Theorem 8.1 in order to prove Theorem 1.2. It does not seem possible to simplify arguments, because of degeneracy of polar coordinates at the planes $r_j = 0$, $j = 1, 2$.

Consider the rough sketch of the proof from Section 3.1 and Lagrangian L_f defined in (9). We shall modify it by restricting L_f to the annuli $\{A_j(\omega^0)\}_{j\in\mathbb{Z}_+}$ and applying Theorem 8.1 to each of these restrictions. This will allow us to construct a modification of the variational principle (25) appropriate for our problem.

Write the remainder P_f of L_f in the form (11). If $\Gamma \subset \Lambda_k$ is contained in one of the axes, some terms might vanish, but not all of them. For a unit vector $e_k = (e_k^1, e_k^2)$ parallel to Γ, denote

$$P_{f,\Gamma}(\theta, v-\omega^0, t) = \sum_{p=0}^{3} (e_k^1)^p (e_k^2)^{3-p} P_p(\theta, v-\omega^0, t).$$

To apply Theorem 8.1 to L_f in each of the annuli, we need to verify the following hypotheses

– (C1)–(C3) for all ω's in Γ,

– $(C4)_\omega$–$(C8)_\omega$ for ω's with small denominators in Γ, and

– (C9) for all c's in $\mathcal{L}_\beta(\sigma_\Gamma)$.

First we shall verify all hypotheses except (C9). For this purpose, we define the following Lagrangian

$$L_{f,\Gamma}(\theta, v, t) = \frac{1}{2}\langle B^{-1}(v-\omega^0), (v-\omega^0)\rangle + \varepsilon P_{f,\Gamma}(\theta, v-\omega^0, t), \tag{26}$$

where ε is nonzero and small. For this perturbation term $P_{f,\Gamma}$ using (15) (resp. (18)) define the first (ω,Γ)-averaged potential, denoted by $P_{f,\omega,\Gamma} : \mathbb{T}^1_\Gamma \to \mathbb{R}$ (resp. the second ω-averaged one, denoted by $P_{f,\omega} : \mathbb{T}^2_\omega \to \mathbb{R}$).

Rescale the annulus $A_j(\omega^0)$ to the unit size. Denote $T^\lambda_{\omega^0} : v \to \lambda(v-\omega^0)+\omega^0$ rescaling centered at ω^0. We have that $T^{2^j/\delta_0}(A_j(\omega^0)) = A(\omega^0) = \{1/4 < |r| < 4\}$. Notice that Λ_k is invariant under rescaling $T^\lambda_{\omega^0}$ for any $\lambda > 0$.

Restrict the Lagrangian L_f to $\mathbb{T}^2 \times A_j(\omega^0) \times \mathbb{T}$. Consider the rescaling $T^{2^j/\delta_0}_{\omega^0}$ in $(v-\omega^0)$ of $\mathbb{T}^2 \times A_j(\omega^0) \times \mathbb{T}$ to $\mathbb{T}^2 \times A(\omega^0) \times \mathbb{T}$. it gives the new "rescaled" Lagrangian

$$\begin{aligned} L^j_f(\theta, v, t) = {}& \frac{1}{2}\langle B^{-1}(v-\omega^0), (v-\omega^0)\rangle \\ & + \delta_0 2^{-j} \sum_{p=0}^{3} (v-\omega^0)_1^p (v-\omega^0)_2^{3-p} P^j_p(\theta, v-\omega^0, t), \end{aligned} \tag{27}$$

where $\{P^j_p(\theta, v-\omega^0, t) = P_p(\theta, T^{2^j/\delta_0}_{\omega^0} v - \omega^0, t)\}_{p=0}^3$ is $C^{s,d-3}$ smooth and defined in $\mathbb{T}^2 \times K^\alpha_(\omega^0, B) \times \mathbb{T}$. Denote the remainder

$$\textstyle\sum_{p=0}^{3} (v-\omega^0)_1^p (v-\omega^0)_2^{3-p} P^j_p(\theta, T^{2^j/\delta_0}_{\omega^0} v - \omega^0, t)$$

by $P^j_f(\theta, v-\omega^0, t)$.

The definitions of all hypotheses except (C9) involves averaged kinetic and potential energies. Fix any positive integer j. Notice that the unperturbed integrable parts are the same for both L_f and $L^j_{f,\Gamma}$. These parts define the averaged kinetic energies (see (16) and (19)). Thus, the averaged kinetic energies of L_f and $L^j_{f,\Gamma}$ coincide. Now consider the perturbation terms. A direct calculation based on (15) and (18) shows that up to a constant the first ω-averaged (resp. the second (ω, Γ)-averaged) potentials of L_f and $L_{f,\Gamma}$ are coincide respectively. Therefore, for each $j \in Z_+$ we have that up to a constant

Γ-averaged mechanical systems associated to L_f and $L^j_{f,\Gamma}$ coincide.

Denote the first ω-averaged and the second (ω, Γ)-averaged mechanical systems by $L_{\omega,f}$ and $L_{\omega,f,\Gamma}$ respectively. The definition of a small denominator q_0 involves only averaged mechanical systems. After q_0 is determined, notice that choosing δ_0 small enough we need to verify $(C4)_\omega$–$(C8)_\omega$ only for at most one ω^0 in the case ω^0 is a rational with small denominator. Suppose $L_{f,\Gamma}$ in (26) satisfies hypotheses (C1)–(C3) and $(C4)_\omega$–$(C8)_\omega$ (if the latter is necessary). Then, there exists $\delta = \delta(\ell_0, \Gamma, P_{f,\Gamma}) > 0$ such that the variational principle (25) is well-defined for the Lagrangian $L_{f,\Gamma}$ and each $0 < |\varepsilon| < \delta$. In notations of Section 3.1, let δ be given by $2\delta = \delta_0$. Consider also the rescaling of the original Lagrangian L_f in $(v - \omega^0)$, in order to see that the above arguments is applicable to L_f too.

The definition of all hypotheses except (C9) involves averaged kinetic and potential energies. The above verification shows that if L^1_f satisfies hypotheses (C1)–(C3) and $(C4)_\omega$–$(C8)_\omega$ (if the latter is necessary) on Γ_δ, then L^j_f satisfies these hypotheses on Γ_δ too. The only difference is that the constant in front of P^j_f decreases as j increases. This implies that *if $P^1_f \in U^s_{\delta(\ell_0,\Gamma)}$, then $P^j_f \in U^s_{\delta(\ell_0,\Gamma)}$*.

In notations of Section 3.1 we now verify that, for a C^s Whitney open and dense set of remainders (11), the restriction of L_f to any of $\mathbb{T}^2 \times K^\alpha_{\delta,j}(\omega^0, B) \times \mathbb{T}$ satisfies hypotheses (C1)–(C3) on Γ_δ and $(C4)_\omega$–$(C8)_\omega$ (if the latter is necessary). This implies that the variational principle (25) is well-defined. What is left to verify is hypothesis (C9), and the fact that velocity of minimizers of the variational principle (25) belong to the corresponding cones $K^\alpha_{\delta,j}(\omega^0, B)$.

Suppose the first potential satisfies hypotheses (C1)–(C3) on $\Gamma_\rho = [\omega^0, \omega^0 + \rho e_k]$, where ρ is the radius of the ball such that (9) is defined on $\mathbb{T}^2 \times K^\alpha_\rho(\omega^0, B) \times \mathbb{T}$.

Consider now the rescaling L^j_f of the restriction of L_f on the annulus $A_j(\omega)$. By lemma 3.2, there is a smooth curve σ_Γ with channel property. Denote by σ^j_γ a part of this curve which connects $\mathcal{L}_\beta(\omega^0 + \delta_j e_k)$ and $\mathcal{L}_\beta(\omega^0 + \delta_{j+1} e_k)$. We can apply Theorem 8.1 with the curve σ^j_γ as the curve with channel property. According to the variational principle (25) its minimizers velocity moves along Γ with a certain error.

The Localization Lemma I shows that minimizers of the variational principle (25) for $L = L^j_f$ have velocity $C2^{-j/2}$-close to Γ for some constant C. Therefore, after the backward rescaling, the velocity has to be $C2^{-3j/2}$-close to Γ_j. It implies that minimizing trajectory of (25) does not leave the cone $\mathbb{T}^2 \times K^{\alpha}_{\delta,j}(\omega^0, B) \times \mathbb{T}$.

The Localization Lemma II shows that non-degeneracy hypothesis (C9) holds for L^j_f taking into account that the velocity value is $C2^{-j/2}$-close to Γ. By lemma 7.3, the set of restrictions of L^j_f onto $\mathbb{T}^2 \times K^{\alpha}_{\delta,j}(\omega^0, B) \times \mathbb{T}$, where hypothesis (C9) holds, is C^s Whitney open and dense. Therefore, there is a C^s Whitney open and dense set of remainders P_f in (9) such that for all positive integers j the corresponding P^j_f fulfills hypothesis (C9). This completes the proof of Theorem 1.2. □

Appendix A. Mather minimal sets

In this Appendix, we discuss basic objects of Mather's theory of minimal or action-minimizing measures [**Ma**]. This theory can be considered as an extension of KAM theory. Namely, it provides a large class of invariant sets for a Hamiltonian (or dual Euler-Lagrange) flow. KAM invariant tori and Aubry-Mather sets are examples of these sets. We need to define these notions to give the detailed scheme of the proof of Theorem 1.2 (see Section 3.1).

We start with a positive integer n, a smooth n-dimensional torus $\mathbb{T}^n$, and a C^s-smooth time periodic Lagrangian $L : \mathbf{T}\mathbb{T}^n \times \mathbb{T} \to \mathbb{R}$, $(\theta, v) \in \mathbf{T}\mathbb{T}^n$, $s \geqslant 2$ which *satisfies hypotheses* (1–3) of section 2.1. Note that all definitions and results of this section can be given for any smooth compact manifold instead of $\mathbb{T}^n$. Later we apply it for $n = 2$ and the Lagrangian L given by (9) near the zero section and extended outside to keep fiber-convexity.

We say that μ is a *probability measure*, if it is a Borel measure of total mass one. Let $\mathcal{P}_L$ be the space of probability measures on $\mathbf{T}\mathbb{T}^n \times \mathbb{T}$ invariant with respect to Euler-Lagrange flow (7). We shall consider probability measures only from $\mathcal{P}_L$. If η is a closed one-form on $\mathbb{T}^n \times \mathbb{T}$, we may associate to it a real valued function $\widehat{\eta}$ on $\mathbf{T}\mathbb{T}^n \times \mathbb{T}$ as follows: express η in the form

$$\eta = \eta_{\mathbb{T}^n} d\theta + \eta_\tau d\tau,$$

where $\eta_{\mathbb{T}^n}$ is the restriction of η to $\mathbb{T}^n$ and $\eta_\tau : \mathbb{T} \to \mathbb{R}$ and set

$$\widehat{\eta} = \eta_{\mathbb{T}^n} + \eta_\tau \circ \pi,$$

where $\pi : \mathbf{T}\mathbb{T}^n \times \mathbb{T} \to \mathbb{T}^n \times \mathbb{T}$ denotes the natural projection. This function has the property

$$\int_a^b \widehat{\eta}(d\gamma(t), t)dt = \int_{(\gamma,\tau)} \eta,$$

for every absolutely continuous curve $\gamma : [a, b] \to \mathbb{T}^n$ with the right hand being usual integral over the curve $(\gamma, \tau) : [a, b] \to \mathbb{T}^n \times \mathbb{T}$ defined by $(\gamma, \tau)(t) = (\gamma(t), t \bmod 1)$.

If μ is an invariant probability measure on $\mathbf{T}\mathbb{T}^n\times\mathbb{T}$, its *average action* is defined as

$$A(\mu)=\int L(\theta,v,t)\ d\mu(\theta,v,t).$$

Since L is bounded below, this integral is well defined, although it may be equal to $+\infty$. Next step is to define an appropriate notion which generalizes the rotation vector of a periodic trajectory. If $A(\mu)<\infty$, one can define *a rotation vector* $\rho(\mu)\in H_1(\mathbb{T}^n,\mathbb{R})$ *of a probability measure* μ by

$$\langle\rho(\mu),[\eta]_{\mathbb{T}^n}\rangle+[\eta]_{\mathbb{T}}=\int\widehat{\eta}(\theta,v,t)\ d\mu(\theta,v,t) \tag{28}$$

for every C^1 closed one form η on $\mathbb{T}^n\times\mathbb{T}$, where

$$[\eta]=([\eta]_{\mathbb{T}^n},[\eta]_{\mathbb{T}})\in H^1(\mathbb{T}^n\times\mathbb{T},\mathbb{R})=H^1(\mathbb{T}^n,R)\times\mathbb{R}$$

denotes the de Rham cohomology class and $\langle\,,\,\rangle$ denotes the dual pairing $H_1(\mathbb{T}^n,\mathbb{R})\times H^1(\mathbb{T}^n,\mathbb{R})\to\mathbb{R}$. The idea of a rotation vector is classical and goes back to Schwartzman's asymptotic cycles (see [**Ma**] in the time independent case), but in the time dependent case definitions and arguments are the same. In [**Ma**], by using a Krylov-Bogoliuboff type argument, Mather proved the following result:

Lemma A.1. — *For every homology class* $h\in H_1(\mathbb{T}^n,\mathbb{R})$ *there exists a probability measure* $\mu\in\mathcal{P}_L$ *such that* $A(\mu)<\infty$ *and* $\rho(\mu)=h$.

Such a probability measure $\mu\in\mathcal{P}_L$ is called *minimal* or *action-minimizing* if

$$A(\mu)=\min\{A(\nu)\ :\ \rho(\nu)=\rho(\mu)\},$$

where ν ranges in $\mathcal{P}_L$ and $A(\nu)<\infty$. If $\rho(\mu)=h$, we also say that μ is *h-minimal.* Denote by $\mathcal{M}_h$ closure of the union of supports of all h-minimal measures from $\mathcal{P}_L$. This set $\mathcal{M}_h\subset T\mathbb{T}^n\times\mathbb{T}$ is called *Mather set.* By the above lemma $\mathcal{M}_h$ is always nonempty.

A probability measure $\mu\in\mathcal{P}_L$ is *c-minimal* for $c\in H^1(\mathbb{T}^n,\mathbb{R})$, if it minimizes

$$A_c(\mu)=A(\mu)-\langle\rho(\mu),c\rangle$$

over all invariant probability measures. $A_c(\mu)$ as above is called c-action of a measure. Mather [**Ma**] also proved the following result:

Lemma A.2. — *For every cohomology class* $c\in H^1(\mathbb{T}^n,\mathbb{R})$ *there exists a* c-minimal *probability measure* $\mu\in\mathcal{P}_L$ *such that* $A(\mu)<\infty$.

Denote by $\mathcal{M}^c$ closure of supports of the union of all c-minimal measures from $\mathcal{P}_L$. $\mathcal{M}^c\subset T\mathbb{T}^n\times\mathbb{T}$ is also called *Mather set.* By the above lemma $\mathcal{M}^c$ is always nonempty. Mather [**Ma**] proved that

$$\cup_{h\in H_1(\mathbb{T}^n,R)}\mathcal{M}_h=\cup_{c\in H^1(\mathbb{T}^n,R)}\mathcal{M}^c.$$

It turns out that $\mathcal{M}^c$ can be "nicely" projected onto the base $\mathbb{T}^n\times\mathbb{T}$.

Graph Theorem. — *Let $\pi : \mathbf{T}\mathbb{T}^n \times \mathbb{T} \to \mathbb{T}^n \times \mathbb{T}$ be the natural projection onto the base. Then, for any $c \in H^1(\mathbb{T}^n, \mathbb{R})$, the corresponding Mather set $\mathcal{M}^c$ is a Lipschitz graph over the base $\mathbb{T}^n \times \mathbb{T}$, i.e. $\pi^{-1}|_{\pi\mathcal{M}^c} : \pi\mathcal{M}^c \to \mathcal{M}^c$.*

Call $\pi\mathcal{M}^c$ *projected Mather set* and denote $\widetilde{\mathcal{M}}^c = \pi\mathcal{M}^c$.

Definition A.3. — The function

$$\beta_L : \ H_1(\mathbb{T}^n, \mathbb{R}) \longrightarrow \mathbb{R}, \quad \beta_L(h) = A(\mu),$$

where μ is an $h-$minimal probability measure, is called *Mather's β-function.* The function

$$\alpha_L \ : \ H^1(\mathbb{T}^n, \mathbb{R}) \longrightarrow \mathbb{R}, \quad \alpha_L(c) = \sup_{h \in H_1(\mathbb{T}^n, R)} \{\langle h, c\rangle - \beta_L(h)\}$$

is called *Mather's α-function.*

***Lemma A.4* ([Ma])**. — *Both α-function and β-function are convex and conjugate by the Legendre transform.*

By definition,

$$\beta_L(h) + \alpha_L(c) \geqslant \langle h, c\rangle, \quad h \in H_1(\mathbb{T}^n, \mathbb{R}), \ c \in H^1(\mathbb{T}^n, \mathbb{R}).$$

To distinguish from the standard Legendre transform (6) the map

$$(29) \quad \mathcal{L}_\beta : H_1(\mathbb{T}^n, \mathbb{R}) \longrightarrow \{\text{compact, convex, non-empty subsets of } H^1(\mathbb{T}^n, \mathbb{R})\},$$

defined by letting $\mathcal{L}_\beta(h)$ be the set of $c \in H^1(\mathbb{T}^n, \mathbb{R})$ for which the inequality in (9) becomes equality, is called *Fenchel-Legendre transform.* In what follows, we shall identify each $h-$minimal invariant probability measure with a $c-$minimal invariant probability measure, provided that $c \in \mathcal{L}_\beta(h)$.

For an absolutely continuous curve $\gamma : [a, b] \to \mathbb{T}^n$, let us denote $d\gamma(t) = (\gamma(t), \dot\gamma(t))$. The above is well defined for a.e. t. For such γ and a closed one form η with $[\eta]_{\mathbb{T}^n} = c$, we call *$c$-action*

$$(30) \qquad A_c(\gamma) = \int_a^b (L - \widehat{\eta})(d\gamma(t), t)dt,$$

where $\widehat{\eta}(\theta, \dot\theta, t) = \eta_{T^n}(\theta, t)\dot\theta + \eta_T(\theta, t)$ if $\eta = {\eta_T}^n d\theta + \eta_T dt$. Notice that c-action does not depend on a choice of η in the cohomology class c. A closed one form η on $\mathbb{T}^n \times \mathbb{T}$ is called *Mañé critical* if and only if

$$\min_{\mu \in \mathcal{P}_L} \left\{ \int (L - \widehat{\eta})d\mu \right\} = 0.$$

Since each closed one form can be written as $[\eta] = ([\eta]_{\mathbb{T}^n}, [\eta]_{\mathbb{T}})$, by the definition of α-function for Mañé critical one form we have $[\eta]_{\mathbb{T}} = -\alpha([\eta]_{\mathbb{T}^n})$. We also say that η is *Mañé supercritical* if $[\eta]_{\mathbb{T}} > -\alpha_L([\eta]_{\mathbb{T}^n})$ and *Mañé subcritical* if $[\eta]_{\mathbb{T}} < -\alpha_L([\eta]_{\mathbb{T}^n})$. We shall explain geometric meaning of sub and super criticality in the next section.

We say that an absolutely continuous curve $\gamma : \mathbb{R} \to \mathbb{T}^n$ is an absolute *c-minimizer* if, for any interval $[a, b]$ and any absolutely continuous curve $\gamma_1 : [d, e] \to \mathbb{T}^n$ such that $d \equiv a \pmod 1$ and $e \equiv b \pmod 1$, we have

$$\int_a^b (L - \widehat{\eta})(d\gamma(t), t)dt \leqslant \int_d^e (L - \widehat{\eta})(d\gamma_1(t), t)dt,$$

where η is a Mañé critical closed one form on $\mathbb{T}^n \times \mathbb{T}$ such that $[\eta]_{\mathbb{T}^n} = c$. Notice that the time intervals $b - a$ and $e - d$ are not necessarily the same. Completeness of the Euler-Lagrange flow (see property 3 of Lagrangian) implies that every c-minimal curve is C^1-smooth and, therefore, as smooth as L is. So it is C^{s-1}-smooth. Denote the union of all sets of c-minimizers $\{(d\gamma(t), t) : \ t \in \mathbb{R}\} \subset T\mathbb{T}^n \times \mathbb{T}$ by $\mathcal{N}^c$ and call it *Mañé set.* This set is certainly a closed set.

We now introduce the notion of barrier function and we deal with another set of trajectories associated to a cohomology class $c \in H^1(\mathbb{T}^n, \mathbb{R})$. The barrier function is introduced in [**Ma2**] and is a generalization of Peierl's barrier. Let $\theta_1, \theta_2 \in \mathbb{T}^n$, $\tau_1, \tau_2 \in \mathbb{T} \geqslant 0$, and η is a Mañé critical one closed one form on $\mathbb{T}^n \times \mathbb{T}$ such that $[\eta]_{\mathbb{T}^n} = c$. Define

$$h_{\eta,T}((\theta_1, \tau_1), (\theta_2, \tau_2)) = \inf \int (L - \widehat{\eta})(d\gamma(t), t)dt, \tag{31}$$

where the infinum is taken over all absolutely continuous curves $\theta : [a, b] \to \mathbb{T}^n$ such that $a \equiv \tau_1 \pmod 1$, $b \equiv \tau_2 \pmod 1$, $\theta(a) = \theta_1$, $\theta(b) = \theta_2$, and $b - a \geqslant T$. Define the barrier function

$$h_\eta((\theta_1, \tau_1), (\theta_2, \tau_2)) = \liminf_{T\to+\infty} h_{\eta,T}((\theta_1, \tau_1), (\theta_2, \tau_2)). \tag{32}$$

In [**Ma2**], Mather has proved that the limit exists. He also introduced a pseudo-metric

$$\rho_c((\theta_1, \tau_1), (\theta_2, \tau_2)) = h_\eta((\theta_1, \tau_1), (\theta_2, \tau_2)) + h_\eta((\theta_2, \tau_2), (\theta_1, \tau_1)). \tag{33}$$

It turns out that this construction is independent of η, provided η is Mañé critical and $[\eta]_{\mathbb{T}^n} = c$. One can show that $\rho_c \geqslant 0$, satisfies the triangle inequality, and is independent of the choice of a Mañé critical closed one form η. The set

$$\mathcal{A}^c = \{(\theta, \tau) : \ \rho_c((\theta, \tau), (\theta, \tau)) = 0\}$$

is called *Aubry set.* One can show [**Ma2**] that

$$\mathcal{M}^c \subset \mathcal{A}^c \subset \mathcal{N}^c \subset T\mathbb{T}^n \times \mathbb{T}$$

for all $c \in H^1(\mathbb{T}^n, \mathbb{R})$ and $\mathcal{A}^c$ also satisfies the Graph Theorem stated above.

Appendix B. Proofs of auxiliary lemmas

Lemma B.1. — *For positive ρ and $0 < \alpha < 1$ consider the space of $C^{s,d}$ smooth function on $\widehat{\mathcal{K}}^\alpha_\rho \ni (\theta, r)$ with the natural C^s Whitney topology in $\mathcal{K}^\alpha_\rho$, where $m \leqslant d \leqslant$*

$s \leqslant \infty$, $d, m \in \mathbb{Z}_+$. Then for any C^s Whitney open dense set $\mathcal{D}^s_{d-m}$ of $(m+1)$-tuples $(g_0, \dots, g_m)$ of $C^{s,d-m}$ smooth functions on $\mathcal{K}^\alpha_\rho$, the set of functions of the form

$$g(\theta, r) = \sum_{p=0}^{m} r_1^p r_2^{m-p} g_p(\theta, r) \tag{34}$$

with $(m+1)$-tuples $(g_0, \dots, g_m) \in \mathcal{D}^s_{d-m}$ intersected with $C^s(\mathcal{K}^{2\alpha}_\rho)$ is C^s Whitney open dense.

Proof of Lemma B.1. — Pick $r = (r_1, r_2) \in K^\alpha_\rho$, *i.e.* $0 < \alpha r_1 < r_2 < \rho$ and $0 < \alpha r_2 < r_1 < \rho$. We find two functions $f_{11}(\theta, r)$ and $f_{21}(\theta, r)$ defined on $\mathcal{K}^\alpha_\rho$ satisfying

$$r_1 f_{11}(\theta, r) - r_2 f_{21}(\theta, r) = f(\theta, r). \tag{35}$$

Consider two functions equalities:

$$f(\theta, r) - f(\theta, r_1, \frac{\alpha}{2} r_1) = r_2 g_2(\theta, r), \quad f(\theta, r_1, \frac{\alpha}{2} r_2) = r_1 g_1(\theta, r).$$

To define f_1 and f_2 by explicit formulas inside K^α_ρ consider the coordinate change: $\widetilde{r}_1 = (r_1, \alpha r_1)$, $r_2 = (0, r_2 - \alpha r_1)$, and $\widetilde{f}(\theta, \widetilde{r}_1, \widetilde{r}_2) = f(\theta, r_1, r_2)$. By Hadamard-Torricelli's lemma

$$f_1(\theta, \widetilde{r}_1, 0) = \int_0^1 \frac{\partial \widetilde{f}}{\partial \widetilde{r}_1}(\theta, t\widetilde{r}_1, 0) dt = \frac{\widetilde{f}(\theta, \widetilde{r}_1, 0)}{r_1}$$

$$r_1 g_1(\theta, \widetilde{r}_1, 0) = \widetilde{f}(\theta, \widetilde{r}_1, 0) - \widetilde{f}(\theta, 0) = f(\theta, r_1, \frac{\alpha}{2} r_1)$$

$$g_2(\theta, \widetilde{r}_1, \widetilde{r}_2) = \int_0^1 \frac{\partial \widetilde{f}}{\partial \widetilde{r}_2}(\theta, \widetilde{r}_1, t\widetilde{r}_2) dt = \frac{\widetilde{f}(\theta, \widetilde{r}_1, \widetilde{r}_2) - \widetilde{f}(\theta, \widetilde{r}_1, 0)}{r_2 - \frac{\alpha}{2} r_1}.$$

This implies that for $f_{11} = g_1 + \frac{\alpha}{2} g_2$ and $f_{21} = g_2$ (35) holds true.

Notice that f_{11} and f_{21} have zero of order $(m-1)$ in r in the sense that they are $C^{s,d-1}$ smooth. Application of Hadamard-Torricelli's lemma to f_{11} and f_{21} gives explicit formulas for functions f_{02}, f_{12}, and f_{22}, which have zero of order $(m-2)$ in r, namely, functions belong to $C^{s,d-2}(\mathcal{K}^\alpha_\rho)$. After m steps we get explicit formulas for functions $f_{0m}, \dots, f_{mm}$. Denote $f_{pm} = f_p$ for $p = 0, \dots, m$. These functions satisfy (34) which completes the proof. □

Proof of Lemma 2.1. — We start with an integrable truncation of f. Let $\widehat{f}$ be defined by (1) with $P_j, Q_j \equiv 0$. Then, the time 1 map of the Hamiltonian

$$H_0(\theta, r) = \omega_1 r_1 + \omega_2 r_2 + \frac{1}{2}\langle Br, r\rangle \tag{36}$$

coincides with f_0, as an easy calculation shows. Construct a time periodic deformation $\{H(\cdot, t)\}_{t\in\mathbb{T}}$ of H_0 so that the time 1 map of $H(\cdot, t)$ coincides with f and so that $H(\cdot, t)$ is C^{s+1} close to $H_0(\cdot)$ near $r = 0$ for all $t \in \mathbb{T}$. Since H_0 is convex in r, $s \geqslant 2$, and we are interested in small r, it implies the desired convexity of $H(\cdot, t)$ in r for each $t \in \mathbb{T}$.

The construction of $\{H(\cdot, t)\}_{t\in\mathbb{T}}$ is done using generating functions. We recall a standard fact from Hamiltonian system (see *e.g.* Arnold [**Ar2**] sect. 48) for a C^s

smooth symplectic map $g(\theta,r)=(\Theta,R)$, $\theta,\Theta\in\mathbb{T}^n$, $r,R\in\mathbb{R}^n_+$: one can define a C^{s+1} smooth generating function $S_g(\theta,\Theta)$ so that

$$\begin{cases} r=-\partial_\theta S_g(\theta,\Theta),\\ R=\partial_\Theta S_g(\theta,\Theta).\end{cases} \tag{37}$$

The function $S_g(\theta,\Theta)$ above is defined up to a constant. Direct calculation for $\widehat{f}$ and f in a small r-neighborhood of zero show that

$$\begin{aligned} S_{\widehat{f}}(\Theta,\theta)&=\frac{1}{2}\langle B^{-1}(\Theta-\theta-\omega),(\Theta-\theta-\omega)\rangle \text{ and}\\ S_f(\Theta,\theta)&=S_{\widehat{f}}(\Theta,\theta)+0(|(\Theta-\theta-\omega)|^3).\end{aligned} \tag{38}$$

Consider a smooth family of generating functions $\{\widetilde{S}_t\}_{t\in[0,1]}$ given by

$$\widetilde{S}_t(\theta,\Theta)=\begin{cases}\dfrac{h(t)}{2}\big\langle B^{-1}(\Theta-\theta-\omega/h(t)),(\Theta-\theta-\omega/h(t))\big\rangle & \text{for } t\in(0,\frac{1}{2}]\\ h(t)\widetilde{S}_{1/2}(\theta,\Theta)+(1-h(t))S_f(\theta,\Theta) & \text{for } t\in[\frac{1}{2},1],\end{cases}$$

where h is a smooth positive function away from zero, $h(1)=h'(1/2)=0$, $h(1/2)=1$, and $1/h(t)=t$ near $t=0+$. The choice of h is designed so that $\widetilde{S}_t$ is sufficiently smooth with respect to t for $t\in(0,1]$. By construction, $\widetilde{S}_t$ generates a smooth family $\{\widetilde{f}_t\}_{t\in(0,1]}$ of exact symplectic twist maps (see [**Go**] sect 26 or [**McS**] sect 9.3). More precisely, $\widetilde{f}_t(\theta,r)=\big(\theta+(h(t))^{-1}(\omega+Br),r\big)$ for $0\leqslant t\leqslant 1/2$ and $\lim_{t\to 0+}\widetilde{f}_t=\mathrm{Id}$. Define $s_t(\theta,r)=\widetilde{S}_t(\theta,\theta+(h(t))^{-1}(\omega+Br))$ for $0\leqslant t\leqslant 1/2$. It becomes

$$s_t(\theta,r)=(2h(t))^{-1}\langle Br,r\rangle.$$

By assumption, $1/h(t)=t$ near zero s_t can be smoothly continued for all $t\in[0,1]$ with $s_0\equiv 0$ and $s_1(\theta,r)=S_f(\theta,\Theta(\theta,r))$, where $\Theta(\theta,r)$ is given by $f(\theta,r)=(\Theta,R)$ for some R. Now we can write:

$$\widetilde{f}_t^*(rd\theta)-rd\theta=ds_t,\quad t\in[0,1]. \tag{39}$$

It shows that $\{\widetilde{f}_t\}_{t\in[0,1]}$ is a Hamiltonian isotopy. By standard results from symplectic geometry, obtained by combining homotopy formula and (39) (see *e.g.* Prop. 9.18 in [**McS**] or Thm. 58.9 in [**Go**]) this family generates Hamiltonian functions $\{H(\cdot,t)\}_{t\in[0,1]}$ as follows. Denote by $\widetilde{X}_t$ vector fields generated by isotopy $\widetilde{f}_t$, *i.e.* given by $\widetilde{X}_t(\theta,r)=(d\widetilde{f}_t/dt)\big((\widetilde{f}_t)^{-1}(\theta,r)\big)$, by $i_{\widetilde{X}_t}\alpha$ the interior derivative of a 1-form α by $\widetilde{X}_t$, and by $s_t(\theta,r)$ a form of generating function, given by (39). Then

$$\widetilde{H}_f(\cdot,t)=i_{\widetilde{X}_t}rd\theta-(\widetilde{f}_t^{-1})^*\left(\frac{d}{dt}s_t\right).$$

One can check that $d\widetilde{H}_f(\cdot,t)=-i_{\widetilde{X}_t}dr\wedge d\theta$.

By the construction, the time map of $\widetilde{H}_f(\cdot,t)$ from time $t=0$ to $t=1$ equals f. However, $\{H(\cdot,t)\}_{t\in[0,1]}$ is not necessarily periodic in t. To attain periodicity we slightly modify the above construction. For small t, say $t\in[0,\delta]$, we have $f_t(\theta,r)=$

$(\theta + tBr, r)$ and, therefore, $\widetilde{H}_f(\cdot, t) = H_0(\theta, r)$. Let us define $f_{1-\tau} = f_\tau^{-1} \circ f$ for $\tau \in [0, \delta]$. Let $\widehat{S}_{1-\tau}(\theta, \Theta)$ be the generating function of $f_{1-\tau}$ with $\widehat{S}_{1-\tau}(\theta, \theta + \omega) = 0$. Consider the following family of generating functions

$$S_t(\theta, \Theta) = \begin{cases} \widetilde{S}_t(\theta, \Theta) & \text{for } t \in (0, 1-\delta] \\ (1 - g(t))\widetilde{S}_t(\theta, \Theta) + g(t)\widehat{S}_{1-t}(\theta, \Theta)) & \text{for } t \in [1-\delta, 1], \end{cases}$$

where $g(t)$ is a C^{s+1} smooth function on $[1-\delta, 1]$, with $g(1-\delta) = 0$, $g(1) = 1$, and $g^{(p)}(1) = g^{(p)}(1-\delta) = 0$ for $p = 0, 1, \ldots, s+1$. By construction, S_t defines a Hamiltonian isotopy $\{f_t\}_{t\in(0,1]}$ with $f_1 = \widetilde{f}_1 = f$. By the same token as above $\{f_t\}_{t\in(0,1]}$ defines a Hamiltonian function $H_f(\cdot, t)$ which is C^{s+1} smooth and periodic in time t.

In order to verify positive definiteness, we consider two cases: $t \in [0, 1/2]$ and $t \in [1/2, 1]$. In the first case, near $t = 0+$, we have $f_t(\theta, r) = (\theta + tBr, r)$ and, therefore, $H_f(\theta, r, t) = H_0(\theta, r)$. Similarly, for $0 \leqslant t \leqslant 1/2$, but not near zero, we have that $H_f(\theta, r, t) = m(t)H_0(\theta, r)$, where $m(t)$ is a smooth strictly positive function (explicitly computable from $h(t)$). Definition (36) of H_0 and hypothesis (H2) of positive definiteness of B implies positive definiteness of $\partial^2_{rr}H_f$ for $0 \leqslant t \leqslant 1-\delta$.

In the case $t \in [1-\delta, 1]$, by definition, $S_t(\theta, \Theta) = S_{\widetilde{f}}(\theta, \Theta) + 0(|(\Theta - \theta - \omega)|^3)$. Explicit calculation gives that the underlying Hamiltonian has the form

$$H(\theta, r, t) = \omega_1 r_1 + \omega_2 r_2 + \frac{1}{2}\langle Br, r\rangle + 0(|r|^3). \tag{40}$$

It implies the Hessian $\partial^2_{rr}H(\cdot, t)$ is close to B and, therefore, positive definite. This proves the lemma. □

Acknowledgment. — The first author would like to thank Sergei Bolotin for useful discussions. The third author would like to thank Caltech for an invitation during which he had the pleasant opportunity to start working on this topic.

References

[Ar1] V. Arnold – "On the Stability of Positions of Equilibrium of a Hamiltonian System of Ordinary Differential Equations in the General Elliptic Case", *Dokl. Akad. Nauk SSSR* **137** (1961), no. 2, p. 255–257, *Sov Math Dokl* **2** (1961), p. 247-279.

[Ar2] ______, *Mathematical Methods in Classical Mechanics*, Graduate Texts in Math., vol. 60, Springer-Verlag, 1989.

[AKN] V. Arnold, V. Kozlov & A. Neidstadt – *Mathematical aspects of classical and celestial mechanics*, Encyclopedia Math. Sci., vol. 3, Springer, Berlin, 1993, Translated from the 1985 Russian original by A. Iacob.

[Be] P. Bernard – "The dynamics of pseudographs in convex hamiltonian systems", preprint, 56 pp., `http://www-fourier.ujf-grenoble.fr/~pbernard/textes/PG.pdf`, 2004.

[BK] D. Bernstein & A. Katok – "Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonian", *Invent. Math.* **88** (1987), p. 225–241.

[CY] Ch-Q. Cheng & J. Yan – "Existence of Diffusion Orbits in a priori Unstable Hamiltonian systems", to appear in *J. Differential Geometry*, 53 pp.

[DLS] A. Delshams, R. de la Llave & T. Seara – "A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristic and rigorous verification on a model", *Electron. Res. Announc. Amer. Math. Soc.* **9** (2003), p. 125–134, electronic.

[DC] M. Dias Carneiro – "On minimizing measures of the action of autonomous Lagrangians", *Nonlinearity* **8** (1995), p. 1077–1085.

[Dou] R. Douady – "Stabilité ou Instabilité des Points Fixes Elliptiques", *Ann. scient. Éc. Norm. Sup.* 4^e *série* **21** (1988), p. 1–46.

[Fa] A. Fathi – "Weak KAM Theorem in Lagrangian Dynamics", preprint of a forthcoming book, 139 pp., october 2003.

[Go] Ch. Golé – *Symplectic Twist Maps, Global Variational Techniques*, Advanced Series in Nonlinear Dynamics, vol. 18, 2001.

[Hed] G.A. Hedlund – "The dynamics of geodesic flows", *Bull. Amer. Math. Soc. (N.S.)* **45** (1939), p. 241–260.

[Her] M. Herman – "Dynamics connected to indefinite normal torsion", IMA, vol. 44, Springer-Verlag.

[KM] V. Kaloshin & J. Mather – "Instabilities of nearly integrable a priori unstable Hamiltonian systems", in preparation.

[KP] S. Kuksin & J. Pöschel – "On the inclusion of analytic symplectic maps in analytic Hamiltonian flows and its applications", in *Seminar on Dynamical Systems (St. Petersburg, 1991)*, Progr. Nonlinear Differential Equations Appl., vol. 12, Birkhäuser, Basel, 1994, p. 96–116.

[LM] M. Levi & J. Moser – "A Lagrangian proof of the invariant curve theorem for twist mappings", in *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, Proc. Sympos. Pure Math., vol. 69, American Mathematical Society, Providence, RI, 2001, p. 733–746.

[Ma] J. Mather – "Action minimizing invariant measures for positive Lagrangian systems", *Math. Z.* **207** (1991), no. 2, p. 169–207.

[Ma2] ______, "Variational construction of connecting orbits", *Ann. Inst. Fourier (Grenoble)* **43** (1993), p. 1349–1386.

[Ma3] ______, "Existence of unbounded orbits for generic mechanical systems on 2-torus", preprint, 1996.

[Ma4] ______, Graduate class at Princeton, 2002-2003.

[Ma5] ______, "Arnold diffusion, I: Announcement of results", *Kluwer Academic Plenum Public. ser. Journ. of Math. Sciences* (2004).

[McS] D. McDuff & D. Salamon – *Introduction to Symplectic Topology*, Oxford Mathematical Monographs, 1995.

[MH] K. Meyer & G. Hall – *Introduction to Hamiltonian dynamical systems and the N-body problem*, Springer-Verlag, New York, 1995.

[Mor] M. Morse – "A fundamental class of geodesics on any closed surface of genus greater than one", *Trans. Amer. Math. Soc.* **26** (1924), p. 25–60.

[Mo] J. Moser – "On Invariant curves of Area-Preserving Mappings of an Annulus", *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* **1** (1962).

[Mo2] ______, "Monotone twist mappings and the calculus of variations", *Ergodic Theory Dynam. Systems* **6** (1986), no. 3, p. 401–413.

[T1] D. Treschev – "Multidimensional symplectic separatrix maps", *J. Nonlinear Sci.* **12** (2002), no. 1, p. 27–58.

[T2] ______, "Trajectories in a neighborhood of asymptotic surfaces of a priori unstable Hamiltonian systems", *Nonlinearity* **15** (2002), no. 6, p. 2033–2052.

[T3] ______, "Evolution of slow variables in a priori unstable Hamiltonian systems", preprint, 34pp.

[X1] Z. Xia – "Arnold diffusion: a variational construction", in *Proc. of ICM, vol. II, (Berlin, 1998)*, 1988.

[X2] ______, "Arnold diffusion and instabilities in hamiltonian dynamics", preprint.

V. Kaloshin, Current: Institute for Advanced Studies, Einstein Drive, Princeton, NJ, 08540 (USA)
On leave: Department of Mathematics 253-37, Caltech, Pasadena, CA 91125 (USA)
E-mail : kaloshin@its.caltech.edu

J.N. Mather, Department of Mathematics, Princeton University, Princeton, NJ, 08544 (USA)
E-mail : jnm@math.princeton.edu

E. Valdinoci, Dipartimento di Matematica, Università di Roma Tor Vergata, Via della Ricerca Scientifica, Roma, I-00133 (Italy) • *E-mail :* valdinoci@mat.uniroma2.it