The anti-integrable limit

S. V. Bolotin and D. V. Treschev

Abstract. The anti-integrable limit is one of the convenient and relatively simple methods for the construction of chaotic hyperbolic invariant sets in Lagrangian, Hamiltonian, and other dynamical systems. This survey discusses the most natural context of the method, namely, discrete Lagrangian systems, and then presents examples and applications.

Bibliography: 75 titles.

Keywords: Lagrangian systems, Hamiltonian systems, chaos, hyperbolic sets, topological Markov chain, topological entropy.

Contents

1. Introduction 976
2. The standard map 978
3. Discrete Lagrangian systems 984
   3.1. The simplest case 984
   3.2. Multivalued Lagrangians 986
   3.3. DLS generated by a symplectic map 987
   3.4. DLS generated by a Lagrangian flow 988
4. Anti-integrable limit in DLS 990
   4.1. Main theorem 990
   4.2. Proof of Theorem 5 993
   4.3. $G$-equivariant DLS 995
5. Examples: systems with discrete time 997
   5.1. Light particle and periodic kicks 997
   5.2. Billiards in a wide strip 998
   5.3. Billiard systems with small scatterers 999
   5.4. Birkhoff–Smale–Shilnikov theorem 1002
   5.5. Shadowing a chain of invariant tori 1004
6. The AI limit in continuous Lagrangian systems 1008
   6.1. Turaev–Shilnikov theorem for Hamiltonian systems 1009
   6.2. $n$-centre problem with small masses 1012
   6.3. Lagrangian systems with slow time dependence 1016

This work is supported by the Russian Science Foundation under grant 14-50-00005.

AMS 2010 Mathematics Subject Classification. Primary 37D45; Secondary 37B10, 37B40.
1. Introduction

The development of mathematics has been uneven and unpredictable. Unexpected and surprising breakthroughs have alternated with strange delays. There was such a delay, difficult to explain, with the emergence of the general concept of the anti-integrable (AI) limit. The theory of dynamical systems was technically ready in part for such a concept about 100 years ago and completely ready 50 years ago. The main ideas are contained in works of Poincaré, Birkhoff, Hedlund, Morse, and other creators of symbolic dynamics. Observations about the hyperbolicity of the invariant set that appears in the AI limit are more or less straightforward. For this, one just needs the general concept of hyperbolicity and standard technical tools such as the cone criterion for hyperbolicity. However, the anti-integrable limit as a universal approach appeared only at the end of the last century.

A starting point for the formation of the ideology of the AI limit was the paper by Aubry and Abramovici [6], where the term ‘anti-integrable limit’ was introduced and the main features of the method were fixed: the context (discrete Lagrangian systems), the main technical tool (the contraction principle), and the language for presentation of the results (symbolic dynamics).

We recall that the standard map is the area-preserving map of the cylinder \( \{(x, y) : x \in \mathbb{T}, y \in \mathbb{R}\} \), onto itself given by the formula

\[
\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x + y + \lambda \sin x \\ y + \lambda \sin x \end{pmatrix}.
\]

If compactness is needed, then one can take the torus \( \mathbb{T}^2 \) as the phase space.

For \( \lambda = 0 \) the map is integrable: the variable \( y \) remains constant on trajectories. The case of small \( |\lambda| \) has been studied extensively in the physical and mathematical literature. The limit as \( \lambda \to \infty \) is called the AI limit [6]. The main result in [6] is the construction for large \( |\lambda| \) of an uncountable set of trajectories which are in one-to-one correspondence with a certain set of quasi-trajectories (codes).

The AI limit appeared as a realization of the simple idea that it is natural to expect chaotic properties to become more pronounced when the ‘distance to the set of integrable systems’ (whatever this means) increases.

The absence of an analytic first integral in systems close to the AI limit is intuitively obvious. However, a formal proof of this fact requires some work.\(^1\) On the other hand, one should not think that an unbounded growth of chaos in the anti-integrable limit leads to ergodicity of the system. Indeed, the number of elliptic periodic points can be arbitrarily large for large values of \( \lambda \), and moreover, for an increasing sequence of values of the parameter these points are asymptotically dense on the phase torus [41]. Since in a general situation elliptic periodic trajectories

\(^1\)In the case of a two-dimensional phase space non-integrability follows from the existence of transversal homoclinic trajectories to periodic solutions.
are surrounded by stability islands, the standard map does not become ergodic for large $|\lambda|$. Although numerically the dynamics looks more and more chaotic for large values of the parameter $\lambda$, in the strict sense it is not even known whether the measure of the chaotic set is positive. For example, is the metric entropy of the standard map positive for at least one value of $\lambda$?

It is easy to show that the trajectories constructed in the AI limit form a uniformly hyperbolic Cantor set. Although this set is uncountable, its measure is zero. A standard way of studying chaos quantitatively is to compute (or estimate) the topological entropy, which, unlike the metric entropy, admits relatively simple positive lower estimates. These estimates are based on the standard fact that if a system has an invariant set with dynamics conjugate to a topological Markov chain, then the topological entropy of the system is not less than that of the latter Markov chain. In the case of the standard map (1.1) this argument gives us that for large $\lambda$ the topological entropy is greater than $c \log \lambda$ for some positive constant $c$ (cf. [54]).

The plan of the paper is as follows. In §2 we give the main ideas, methods, and results about the AI limit in the simplest non-trivial example: the standard map. We define the corresponding set of quasi-trajectories (codes) and prove that for each code there is a unique trajectory of the standard map shadowing this quasi-trajectory. The trajectories obtained in this way form a hyperbolic set. We show that topological entropy supported by this set is of order $\log \lambda$. In §3 we define the class of systems for which the methods of the AI limit will be developed. These systems are called discrete Lagrangian systems (DLS). We start from globally defined DLS (§3.1) and then explain how discrete Lagrangian systems can be generated by so-called ambient systems. In §3.3 the ambient systems are symplectic maps.

In §4 we present results on the AI limit for DLS of general form, which underlie the applications of the method that follow. Then we give examples of systems which can be studied by methods of the AI limit. In each example we have a large parameter, an analogue of $\lambda$ in the standard map. We tried to arrange the examples in order of increasing complexity. The first set of examples (§5) contains systems with discrete time. We start with the problem of the motion of a light particle in the field of a potential with $\delta$-like dependence on $t$. Then we consider a billiard system in a strip with walls formed by the graphs of periodic functions, the large parameter being the distance between the walls. Finally, we consider billiards in a domain with small scatterers, where the large parameter is the inverse of the scatterers’ size. Then in §5 and §6 we use the AI limit to prove several well-known (and less well-known) shadowing theorems, starting with L. P. Shilnikov’s description of symbolic dynamics near a transversal homoclinic orbit and ending with applications to celestial mechanics. In the last section, §7, we discuss applications of the AI limit to the problem of Arnold diffusion.

Many examples presented in this survey are classical and go back to the dynamical folklore, so we do not always give references to the original results.
2. The standard map

Let us rewrite the map (1.1) in the ‘Lagrangian form’. To this end suppose that

\[
\begin{pmatrix} x_- \\ y_- \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x_+ \\ y_+ \end{pmatrix}.
\]

Then \(x_-, x, x_+\) satisfy the equation

\[
\lambda^{-1}(x_+ - 2x + x_-) = \sin x.
\]

(2.1)

The standard map SM written in this form is defined on the cylinder \(\mathcal{Z} = \mathbb{R}^2/\sim\), where the equivalence relation \(\sim\) is as follows:

\[
(x'_1, x'_2) \sim (x_1, x_2) \iff x'_1 - x_1 = x'_2 - x_2 \in 2\pi\mathbb{Z}.
\]

(2.2)

In other words, the cylinder \(\mathcal{Z}\) is the quotient space of the plane \(\mathbb{R}^2\) with respect to the action of the group of shifts

\[
(x_1, x_2) \mapsto (x_1 + 2\pi l, x_2 + 2\pi l), \quad l \in \mathbb{Z}.
\]

The map SM acts as follows:

\[
(x_-, x) \mapsto SM(x_-, x) = (x, x_+), \quad \text{where } x_-, x, x_+ \text{ satisfy (2.1)}.
\]

(2.3)

Infinite sequences \(x = (x_k)_{k \in \mathbb{Z}}\) such that the triple \((x_-, x, x_+) = (x_{k-1}, x_k, x_{k+1})\) satisfies (2.1) for any integer \(l\) are called trajectories of the standard map.

The Lagrangian form of the standard map admits a variational formulation. Namely, trajectories of the system are critical points (extremals) of the formal sum

\[
A(x) = \sum_{k \in \mathbb{Z}} L(x_k, x_{k+1}), \quad L(x', x'') = \frac{1}{2\lambda} (x' - x'')^2 - \cos x''.
\]

(2.4)

This means that \(x^0 = (x^0_k)_{k \in \mathbb{Z}}\) is a trajectory if and only if for any integer \(n\)

\[
\frac{\partial}{\partial x_n} \sum_{k=-\infty}^{\infty} L(x_k, x_{k+1}) = 0 \quad \text{at the point } x = x^0.
\]

(2.5)

In the limit as \(\lambda \to \infty\) the standard map becomes dynamically meaningless because \(x_+\) cannot be found in terms of \(x\) and \(x_-\) from equation (2.1) when \(\lambda^{-1} = 0\). However, the corresponding variational problem remains well defined. Its solutions are sequences

\[
a = (a_j)_{j \in \mathbb{Z}}, \quad a_j \in \pi\mathbb{Z}.
\]

(2.6)

For large values of the parameter \(\lambda\) the standard map has many trajectories close to sequences (2.6).

More precisely, let

\[
\mathcal{A}_\lambda = \{a = (a_k)_{k \in \mathbb{Z}} : a_k \in \pi\mathbb{Z}, |a_{k-1} - 2a_k + a_{k+1}| \leq \Lambda \text{ for any } k \in \mathbb{Z}\}.
\]
For any \( a \in \mathcal{A} \) we define the complete metric space \( \Pi = \Pi(a) \) of sequences
\[
x = \{x_k\}_{k \in \mathbb{Z}}, \quad \sup_{k \in \mathbb{Z}} |x_k - a_k| \leq \frac{\pi}{2}.
\]
The metric on \( \Pi \) is defined as follows:
\[
\rho(x', x'') = \sup_{k \in \mathbb{Z}} |x'_k - x''_k|, \quad x', x'' \in \Pi.
\]

**Theorem 1** [6]. Let \( \Lambda > 0, \sigma \in (0, \pi/2) \), and
\[
\lambda_0 = \lambda_0(\Lambda, \sigma) = \max \left\{ \frac{\Lambda + 4\sigma}{\sin \sigma}, \frac{8}{\cos \sigma} \right\}.
\]
For any \(|\lambda| \geq \lambda_0\) and any \( a \in \mathcal{A} \) the standard map has a unique trajectory \( x \) with \( \rho(a, x) < \sigma \).

Sequences in \( \mathcal{A} \) can be regarded as codes of the corresponding trajectories. This possibility to code trajectories by elements of a sufficiently large set is typical for chaotic systems.

Theorem 1 means that for large values of \( \lambda \) there is an invariant set \( \mathcal{K}_\Lambda = \mathcal{K}_\Lambda(\lambda, \sigma) \subset \mathcal{Z} \) such that the trajectories in \( \mathcal{K}_\Lambda \) are in a one-to-one correspondence with the elements in \( \mathcal{A} \). The formal definition of \( \mathcal{K}_\Lambda \) is as follows. Any code \( a \in \mathcal{A} \) determines a unique orbit \( x = x(a) \). Consider the map
\[
\mathcal{A} \ni a \mapsto \zeta(a) = (x_0, x_1)/\mathbb{Z} \in \mathcal{Z},
\]
where \((x_0, x_1)/\mathbb{Z}\) means the identification (2.2). Then by definition \( \mathcal{K}_\Lambda = \zeta(\mathcal{A}) \).

All trajectories in the set \( \mathcal{K}_\Lambda \) are hyperbolic (Theorem 2 below). Taking another \( \mathbb{Z} \)-quotient, we make the phase space of SM compact: \( \mathcal{Z}/\mathbb{Z} \) is diffeomorphic to the 2-torus \( \mathbb{T}^2 \). Then we prove that \( K_\Lambda = \mathcal{K}_\Lambda/\mathbb{Z} \subset \mathbb{T}^2 \) is compact (Theorem 3) and therefore hyperbolic for the quotient map SM\(_0\). Finally, we estimate the topological entropy of the corresponding dynamical system (Theorem 4).

**Proof of Theorem 1.** The proof is based on the contraction principle in the metric space \((\Pi(a), \rho)\), where \( a \in \mathcal{A} \) is fixed. First we rewrite equations (2.1) in the form
\[
x_k = \arcsin_k \left( \frac{x_{k+1} - 2x_k + x_{k-1}}{\lambda} \right), \quad (2.8)
\]
where \( \arcsin_k \) is the branch of \( \sin^{-1} \) such that \( \arcsin_k(0) = a_k \). Thus, \( \arcsin_k \) maps the interval \((-1, 1)\) to the interval \((a_k - \pi/2, a_k + \pi/2)\). In other words, \( \arcsin_k x = a_k \pm \arcsin x \), where \( \arcsin = \arcsin_0 : (-1, 1) \to (-\pi/2, \pi/2) \) is the standard branch of \( \sin^{-1} \) and the sign \( (+) \) is taken when \( a_k/\pi \) is even (odd).

The sequence \( x = a \) is a trajectory of equations (2.8) for \( \lambda = \infty \).

Consider the map with \( x \mapsto y = \Phi(x) \) such that
\[
y_k = \arcsin_k \left( \frac{x_{k+1} - 2x_k + x_{k-1}}{\lambda} \right).
\]
Then any fixed point of \( \Phi \) is a trajectory of SM.

Let \( B_\sigma(a) \subset \Pi \) be the closed ball with centre \( a \) and radius \( \sigma \).
Lemma 1. Suppose that $|\lambda| > \lambda_0$. Then:
1) $\Phi$ is defined on $B = B_\sigma(a)$ and $\Phi(B) \subset B$;
2) $\Phi$ is a contraction on $B$, that is,

$$\rho(\Phi(x'), \Phi(x'')) < \frac{1}{2} \rho(x', x'') \quad \text{for all } x', x'' \in B.$$ (2.9)

By the contraction principle, Theorem 1 follows from Lemma 1. Now we turn to the proof of the lemma. Recall that $\sigma < \pi/2$. To check that $\Phi(B) \subset B$, it is sufficient to show that for any $x \in B$

$$\left| \frac{x_{k+1} - 2x_k + x_{k-1}}{\lambda} \right| < \sin \sigma.$$ (2.10)

Since $\rho(x, a) < \sigma$, we have

$$|x_{k+1} - 2x_k + x_{k-1}| \leq \Lambda + 4\sigma.$$ Hence (2.10) holds for

$$\lambda_0 > \frac{\Lambda + 4\sigma}{\sin \sigma}.$$ Note that for any pair of real numbers $u', u'' \in (-\sin \sigma, \sin \sigma)$

$$|\arcsin u' - \arcsin u''| \leq \frac{1}{\cos \sigma} |u' - u''|.$$ Here

$$\frac{1}{\cos \sigma} = \sup_{u \in (-\sin \sigma, \sin \sigma)} \left| \frac{d}{du} \arcsin u \right|.$$ We put $y' = \Phi(x')$, $y'' = \Phi(x'')$. Then for any $k \in \mathbb{Z}$

$$|y'_k - y''_k| = \left| \arcsin \left( \frac{x'_{k+1} - 2x'_k + x'_{k-1}}{\lambda} \right) - \arcsin \left( \frac{x''_{k+1} - 2x''_k + x''_{k-1}}{\lambda} \right) \right|$$

$$\leq \frac{1}{\cos \sigma} \left| \frac{x'_{k+1} - 2x'_k + x'_{k-1}}{\lambda} - \frac{x''_{k+1} - 2x''_k + x''_{k-1}}{\lambda} \right|$$

$$\leq \frac{|x'_{k+1} - x''_{k+1}| + 2|x'_k - x''_k| + |x'_{k-1} - x''_{k-1}|}{\lambda \cos \sigma}$$

$$\leq \frac{4}{\lambda \cos \sigma} \rho(x', x'').$$ (2.11)

Consequently, the inequality (2.9) holds if $\lambda_0 > 8/\cos \sigma$. □

Theorem 2. Let $\Lambda > 0$, $\sigma \in (0, \pi/2)$, and $|\lambda| > \lambda_0(\Lambda, \sigma)$. Then any orbit $x(a)$ with $a \in \mathcal{A}_\Lambda$ is hyperbolic.

Proof. Recall that the standard definition of a hyperbolic orbit $x = x(a)$ is given in terms of an invariant decomposition of the tangent spaces at any point into expanding and contracting subspaces [52]. Here we use instead the cone criterion of hyperbolicity due to V.M. Alexeyev (Alekseev), which is also contained in [52].
In this proof we use the notation $f = SM$ for brevity. By (2.1) the Jacobi matrix of the map $f$ is as follows:

$$Df(x, x) = \begin{pmatrix} 0 & 1 \\ -1 & 2 + \lambda \cos x \end{pmatrix}.$$  

At each point $q = (x, x) \in K_\Lambda$ we define the cones

$$H_q = \{(u, u) \in T_q \mathcal{D}: \|u\| \leq \alpha_H \|u\|\},$$

$$V_q = \{(u, u) \in T_q \mathcal{D}: \|u\| \leq \alpha_V \|u\|\},$$  

where $(u, u)$ are the coordinates on $T_q \mathcal{D}$ associated with the coordinates $(x, x)$ on $\mathcal{D}$.

By the cone criterion, to prove the hyperbolicity of trajectories in the invariant set $K_\Lambda$ it is sufficient to check that there exists a $\mu > 1$ such that for all $q \in K_\Lambda$ the following hold:

$$Df_q H_q \subset \text{Int} H_f(q), \quad Df_q^{-1} V_f(q) \subset \text{Int} V_q,$$

$$\|Df_q \xi\| \geq \mu \|\xi\| \quad \text{for} \; \xi \in H_q, \quad \text{and} \quad \|Df_q^{-1} \xi\| \geq \mu \|\xi\| \quad \text{for} \; \xi \in V_f(q).$$  

We take $\alpha_H = \alpha_V = 1/2$. For any $q \in K_\Lambda(\lambda, \sigma)$ and $(u, u) \in H_q$

$$Df(q) \begin{pmatrix} u \\ u_+ \end{pmatrix} = \begin{pmatrix} u \\ u_+ \end{pmatrix}, \quad u_+ = -u_- + (2 + \lambda \cos x)u.$$  

We see that

$$\frac{|u|}{|u_+|} \leq \frac{1}{\lambda \cos \sigma - 5/2}.$$  

Therefore,

$$\begin{pmatrix} u \\ u_+ \end{pmatrix} \in \text{Int} H_f(q), \quad \text{provided that} \; \lambda > \frac{3}{\cos \sigma}.$$  

This implies the first inclusion in (2.13). We also have

$$\left\| \begin{pmatrix} u \\ u_+ \end{pmatrix} \right\|^2 = u^2 + \left(\lambda \cos x - \frac{5}{2}\right) u_+^2 \geq 4 \left\| \begin{pmatrix} u \\ u_+ \end{pmatrix} \right\|^2,$$

provided that $\lambda > 9/(2 \cos \sigma)$. Hence, the first inequality in (2.14) holds with $\mu = 2$.

The second inclusion (2.13) and the second inequality (2.14) can be checked similarly. It remains to note that $\lambda_0 > 9/(2 \cos \sigma)$.  

To give the chaotic properties of the map $SM$ a quantitative character, we estimate its topological entropy. Since in the case of a non-compact phase space topological entropy is usually infinite, we pass to a quotient system with a compact phase space. To this end we note that equations (1.1) can be considered modulo $2\pi$. This generates a quotient system

$$SM_0: \mathbb{T}^2 \to \mathbb{T}^2, \quad \mathbb{T}^2 = \mathbb{R}^2/(2\pi \mathbb{Z}^2),$$
such that the diagram
\[
\begin{array}{ccc}
\mathcal{X} & \xrightarrow{\text{SM}} & \mathcal{X} \\
\text{pr} & & \text{pr} \\
\mathbb{T}^2 & \xrightarrow{\text{SM}_0} & \mathbb{T}^2
\end{array}
\]
is commutative.

Consider on \(\mathcal{A}_\Lambda\) and on the space \(\mathcal{O}\) of orbits of \(\text{SM}\) the following action of the group \(\mathbb{Z}\): for any \((l_1, l_2)\in \mathbb{Z}^2\)
\[
a = (a_i)_{i\in \mathbb{Z}} \mapsto l(a) = (a_i + 2\pi l_1 + 2\pi i l_2)_{i\in \mathbb{Z}},
\]
\[
\mathcal{O} \ni x = (x_i)_{i\in \mathbb{Z}} \mapsto l(x) = (x_i + 2\pi l_1 + 2\pi i l_2)_{i\in \mathbb{Z}}.
\]
We define the quotient spaces
\[
A_\Lambda = \mathcal{A}_\Lambda / \mathbb{Z}^2, \quad O = \mathcal{O} / \mathbb{Z}^2, \quad K_\Lambda = \text{pr } \mathcal{K}_\Lambda.
\]
Then \(A_\Lambda\) can be regarded as the space of codes, and \(O\) as the corresponding space of orbits for \(\text{SM}_0\). The space \(A_\Lambda\) can be identified with
\[
B_\Lambda = \{ b = (b_i)_{i\in \mathbb{Z}} : b_i \in \pi \mathbb{Z}, |b_i| \leq \Lambda \}
\]
via the bijection
\[
a = (a_i)_{i\in \mathbb{Z}} \mapsto b = (b_i)_{i\in \mathbb{Z}}, \quad b_i = a_{i-1} - 2a_i + a_{i+1},
\]
which respects the shift operator \(\mathcal{F} : (a_i)_{i\in \mathbb{Z}} \mapsto (a_{i+1})_{i\in \mathbb{Z}}\).

We recall the definition of the Bernoulli shift. Let \(J\) be a finite set of \(q\) symbols and \(\Sigma_q = J^\mathbb{Z}\) the set of sequences \(a = (a_i)_{i\in \mathbb{Z}}\) of elements in \(J\). We equip \(\Sigma_q\) with the direct-product (Tychonoff) topology: the base consists of cylinders
\[
U_I(a) = \{ a' : a'_i = a_i \text{ for } i \notin I \}, \quad \text{where } I \subset \mathbb{Z} \text{ is a finite set.} \tag{2.15}
\]
It is well known \([52]\) that \(\Sigma_q\) is a compact totally disconnected space with no isolated points, and hence is homeomorphic to the standard Cantor set. The Bernoulli shift with \(q\) symbols is the shift \(\mathcal{F} : \Sigma_q \to \Sigma_q\).

Let
\[
q = \#(\pi \mathbb{Z} \cap [-\Lambda, \Lambda]) = 1 + 2 \left\lfloor \frac{\Lambda}{\pi} \right\rfloor, \tag{2.16}
\]
where \(\lfloor \cdot \rfloor\) is the integer part of a real number.

**Theorem 3.** For any \(\lambda > \lambda_0\) the set \(K_\Lambda\) is a compact hyperbolic invariant set for the standard map, and \(\text{SM}_0 : K_\Lambda \to K_\Lambda\) is topologically conjugate to the Bernoulli shift on the space of \(q\) symbols.

**Proof.** We only need to show that \(\zeta : A_\Lambda \to K_\Lambda\) is continuous. The product topology on \(B_\Lambda\) has the base of open sets (2.15), where \(J = B_\Lambda\). In particular, for any \(a \in B_\Lambda\) the cylinder
\[
C_n(a) = \{ a' \in B_\Lambda : a'_i = a_i \text{ for } |i| \leq n \}
\]
is an open set containing \(a\). The continuity of \(\zeta\) follows from the continuity of the map \(a \mapsto x(a)\), which is given by the following lemma. \(\square\)
Lemma 2. Suppose that $|\lambda| \geq \lambda_0$ and $a' \in C_n(a)$. Then the orbits $x = x(a)$ and $x' = x(a')$ satisfy the estimates

$$|x_k' - x_k| \leq 5^{|k| - n} \cdot 2\sigma \quad \text{for any } |k| \leq n. \quad (2.17)$$

Proof. We obtain the sequences $x$ and $x'$ as the limits

$$x = \lim_{j \to \infty} x^{(j)}, \quad x' = \lim_{j \to \infty} x'^{(j)},$$

$$x^{(0)} = a, \quad x'^{(0)} = a', \quad x^{(j+1)} = \Phi x^{(j)}, \quad x'^{(j+1)} = \Phi x'^{(j)},$$

where $\Phi$ is the operator in Lemma 1.

We use induction on $j$. For $j = 0$ the inequalities (2.17) hold because $x^{(0)} = x'^{(0)}$ for $|k| \leq n$. If $k = n$, then (2.17) follows for all $j > 0$ from assertion 1) of Lemma 1.

Suppose that (2.17) holds for some $j = s$. To prove it for $j = s + 1$ we will use the estimate

$$|x^{(j+1)}_k - x'^{(j+1)}_k| \leq \frac{|x^{(j)}_{k+1} - x'^{(j)}_{k+1}| + 2|x^{(j)}_k - x'^{(j)}_k| + |x^{(j)}_{k-1} - x'^{(j)}_{k-1}|}{\lambda \cos \sigma},$$

which follows from (2.11). For any $|k| \leq n - 1$, by the induction assumption we have

$$|x^{(s+1)}_k - x'^{(s+1)}_k| \leq \frac{(5^{-n+|k|+1} + 2 \cdot 5^{-n+|k|}) + 5^{-n+|k|}-1 \cdot 2\sigma}{\lambda \cos \sigma}.$$ 

The right-hand side of this inequality does not exceed $5^{-n+|k|} \cdot 2\sigma$ because $\lambda_0 \cos \sigma \geq 8$ by (2.7). \qed

In [45] there is another proof of a similar estimate.

Theorem 4. There exists a constant $c > 0$ such that for sufficiently large $\lambda$ the topological entropy of the map $SM_0$ satisfies the estimate\(^2\)

$$h_{\text{top}}(SM_0) \geq c \log \lambda.$$ 

Proof. We already proved that for sufficiently large $\lambda$ the map $SM_0$ has an invariant set $K_{\Lambda}$ such that the restriction $SM_0|_{K_{\Lambda}}$ is conjugate to the Bernoulli shift $\mathcal{T}: \Sigma_q \to \Sigma_q$ on the space of $q$ symbols with (2.16). Hence, $h_{\text{top}}(SM_0)$ is not less than the topological entropy $h_q$ of the Bernoulli shift [52]. The quantity $h_q$ can be computed, for example, from the following equality (which holds also for topological Markov chains) [75]:

$$h_q = \lim_{n \to \infty} \frac{1}{n} \log \theta_n, \quad (2.18)$$

where $\theta_n$ is the cardinality of the set of all admissible $n$-sequences. In this case $\theta_n = q^n$. Therefore, $h_q = \log q$.

It remains to note that according to (2.7) the quantity $\Lambda$ can be chosen greater than a positive constant multiplied by $\lambda$. \qed

\(^2\)Explicit estimates for $c$ and the minimal $\lambda$ can easily be obtained from the proof (see also [54]). However, we certainly do not expect that these estimates are close to optimal.
3. Discrete Lagrangian systems

The anti-integrable limit is usually discussed in Lagrangian systems. We will concentrate on Lagrangian systems with discrete time, which are called discrete Lagrangian systems (DLS). The case of continuous time will be reduced to the case of discrete time (see §6).

3.1. The simplest case. Let $M$ be an $m$-dimensional manifold and $L$ a smooth\(^3\) function on $M^2 = M \times M$. By definition a sequence $x = (x_i)_{i \in \mathbb{Z}}$ with $x_i \in M$ is a trajectory of the discrete Lagrangian system with Lagrangian $L$ if $x$ is an extremal of the action functional, which is the formal sum

$$A(x) = \sum L(x_i, x_{i+1})$$

in the same sense as in the case of the standard map (see (2.5)). Equivalently, for any $i \in \mathbb{Z}$

$$\partial_{x_i}(L(x_{i-1}, x_i) + L(x_i, x_{i+1})) = 0. \tag{3.1}$$

The trajectories remain the same upon multiplication of the Lagrangian by a constant, after addition of a constant to $L$, and after the gauge transformation

$$L(x, y) \mapsto L(x, y) + f(x) - f(y)$$

with an arbitrary smooth function $f$ on $M$. The gauge transformation does not change the action functional.

If the equation

$$\partial_{x}(L(x-, x) + L(x, x_+)) = 0$$

can be globally solved for $x_- = x_-(x, x_+)$ as well as for $x_+ = x_+(x-, x)$, then the map

$$(x_, x) \mapsto T(x_, x) = (x, x_+) \tag{3.2}$$

is a diffeomorphism and determines a discrete dynamical system on $M \times M$. However, the map $T$ is usually only locally defined.

To present conditions for the local existence and smoothness of the map $T$, we define\(^4\)

$$B(x, y) = \partial_{x} \partial_{y} L(x, y),$$

or, in local coordinates,

$$B(x, y) = \left( \frac{\partial^2 L}{\partial x_i \partial y_j} \right). \tag{3.3}$$

In invariant terms, $B(x, y)$ is a linear operator $T_x M \to T^*_y M$, or a bilinear form on $T_x M \times T_y M$. We say that $L$ is a twist Lagrangian if it satisfies the following condition.

**Twist condition.** $B(x, y)$ is non-degenerate for all $x, y \in M$.

\(^3\)C\(^3\) is more than enough.

\(^4\)Sometimes $B$ is defined as $-\partial_{x} \partial_{y} L$, and then in many natural DLS $B$ it will be positive definite.
In this case the map $T$ is locally defined and smooth. It is easy to check (see, for example, [74]) that $T$ is symplectic with respect to the symplectic 2-form $\omega = B(x, y) \, dx \wedge dy$,

$$\omega(u, v) = \langle B(x, y)u_1, v_2 \rangle - \langle B(x, y)v_1, u_2 \rangle, \quad u = (u_1, u_2), \quad v = (v_1, v_2). \quad (3.4)$$

If the twist condition holds, then the Legendre transform $S: M^2 \to T^* M$, which acts according to the rule

$$(x, y) \mapsto (x, p_x), \quad p_x = -\partial_x L(x, y),$$

is locally invertible, and we can represent $T$ by a locally defined map $F = STS^{-1}: T^* M \to T^* M$. The map $F$ is symplectic with respect to the standard symplectic form $dp_x \wedge dx$ on $T^* M$, and $L$ is the generating function of $F$:

$$F(x, p_x) = (y, p_y) \iff p_x = -\partial_x L(x, y), \quad p_y = \partial_y L(x, y). \quad (3.5)$$

Such a symplectic map $F$ is called a twist map (see [47]). Usually it is only locally defined.

The twist condition imposes strong topological restrictions on $M$, and it is rarely satisfied on all of $M \times M$. Here are two canonical examples of DLS.

1. The multidimensional standard map

$$L(x, y) = \frac{1}{2} \langle B(x - y), x - y \rangle - \frac{1}{2} \langle V(x) + W(y), \rangle, \quad x, y \in \mathbb{R}^m, \quad (3.6)$$

where $B$ is a symmetric constant non-singular matrix. In this case $L$ defines a symplectic twist map $F: \mathbb{R}^{2m} \to \mathbb{R}^{2m}$.

Usually $V$ is assumed to be $\mathbb{Z}^m$-periodic. Then the phase space $\mathbb{R}^{2m}$ can be factorized with respect to the $\mathbb{Z}^m$-action $(x, y) \mapsto (x + k, y + k), \ k \in \mathbb{Z}^m$, and $F$ is a symplectic twist self-map of $\mathbb{T}^m \times \mathbb{R}^m$.

We discuss such Lagrangians in more detail in §4.3.

2. Consider a domain $D \subset \mathbb{R}^m$ bounded by a smooth convex hypersurface $M$. The billiard system in $D$ is a DLS with the Lagrangian $L(x, y) = |x - y|$ on $M \times M$. Let us check that $L$ satisfies the twist condition on $(M \times M) \setminus \Delta$, where $\Delta = \{(x, x): x \in M\}$. Let $\langle B(x, y)v, W \rangle$ be the bilinear form on $\mathbb{R}^m \times \mathbb{R}^m$ corresponding to the operator $B(x, y) = \partial_x \partial_y L(x, y)$. A direct computation gives us that

$$\langle B(x, y)v, W \rangle = \frac{-\langle v, w \rangle + \langle v, e \rangle \langle w, e \rangle}{|x - y|}, \quad e = \frac{x - y}{|x - y|}. \quad (3.7)$$

Obviously, $e$ lies in both the left and the right kernel: $B(x, y)e = B^*(x, y)e = 0$. Hence if $e \notin T_x M$ and $e \notin T_y M$, then the restriction of the bilinear form $B(x, y)$ to $T_x M \times T_y M$ is non-degenerate and therefore $L$ is a twist Lagrangian. This always holds if the boundary is strictly convex. All this is well known (for a recent reference see [12]).

In fact, we can identify $T_x M$ and $T_y M$ by the isomorphism $\Pi(x, y): T_x M \to T_y M$ which is the parallel projection in $\mathbb{R}^m$ along the segment $[x, y]$: $\Pi v = v \mod e$. Then

$$\langle B(x, y)v, \Pi(x, y)v \rangle = \frac{-|v|^2 + \langle v, e \rangle^2}{|x - y|} < 0, \quad v \in T_x M \setminus \{0\}.$$
We orient $M$ like the boundary. Since $\Pi(x, y)$ changes the orientation, we get that $(-1)^m \det B(x, y) > 0$. Note that since the image and domain of $B(x, y)$ are different, $\det B(x, y)$ is not invariantly defined (at least as a number), but its sign is.

If $D$ is not convex, then the billiard map $T$ is defined on

$$\{(x, y) \in (M \times M) \setminus \Delta : \text{the line segment } (x, y) \text{ is contained in } D\}$$

and it is singular when the segment $(x, y)$ is tangent to $M$.

In [74] the reader can find other examples of DLS (mostly integrable), including multivalued ones.

### 3.2. Multivalued Lagrangians

In applications the discrete Lagrangian $L$ of a DLS with configuration space $M$ is usually multivalued. A common situation is when $L$ is a function on the universal covering of $M \times M$. For example, for the standard map (3.6) we can assume that $M = T^m$ and $L$ is a function on $\mathbb{R}^{2m}$. One can think of such an $L$ as a multivalued function, that is, a collection of functions on open sets in $M$. To cover the most possible applications we will use the following definition of a DLS.

A DLS with $m$ degrees of freedom is defined by a finite or countable collection $\mathcal{L} = \{L_\kappa\}_{\kappa \in J}$ of functions (Lagrangians) on $m$-dimensional manifolds $U_\kappa^- \times U_\kappa^+$. Usually, the $U_\kappa^\pm$ are open sets in the configuration space $M$. A trajectory of the DLS $\mathcal{L}$ corresponding to a code $k$ is a sequence

$$x = (x_i)_{i \in \mathbb{Z}}, \quad x_i \in U^+_{\kappa_{i-1}} \cap U^-_{\kappa_i}, \quad (3.8)$$

which is a critical point of the formal discrete action functional

$$A_k(x) = \sum_{i \in \mathbb{Z}} L_{\kappa_i}(x_i, x_{i+1}), \quad k = (\kappa_i)_{i \in \mathbb{Z}}. \quad (3.9)$$

More precisely,

$$\partial_{x_i}(L_{\kappa_{i-1}}(x_{i-1}, x_i) + L_{\kappa_i}(x_i, x_{i+1})) = 0, \quad i \in \mathbb{Z}. \quad (3.10)$$

Thus, a trajectory is a pair $(k, x)$, where $k \in J^\mathbb{Z}$. The code $k$ can be regarded as a path on an oriented graph with vertex set $J$. Vertices $\kappa, \kappa' \in J$ are joined by an edge $\gamma = (\kappa, \kappa')$ if $U^-_\kappa \cap U^+_\kappa' \neq \emptyset$.

In this paper we do not use global methods, so by choosing local coordinates we can assume that the $U_\kappa^\pm$ are open sets in $\mathbb{R}^m$.

If the Lagrangians satisfy the twist condition

$$\det B_\kappa(x, y) \neq 0, \quad B_\kappa(x, y) = \partial_x \partial_y L_\kappa(x, y),$$

then an edge $\gamma = (\kappa, \kappa')$ defines a local diffeomorphism

$$f_\gamma : \mathcal{D}_\gamma^- \to \mathcal{D}_\gamma^+$$

of open sets $\mathcal{D}_\gamma^\pm$ in $\mathbb{R}^{2m}$ by means of the equality

$$f_\gamma(x, y) = (y, z) \iff \partial_y(L_\kappa(x, y) + L_{\kappa'}(y, z)) = 0.$$
Then the trajectory \((k, x)\) corresponds to the composition
\[
(x_n, x_{n+1}) = f_{\gamma_n} \circ f_{\gamma_{n-1}} \circ \cdots \circ f_{\gamma_1}(x_0, x_1), \quad \gamma_i = (\kappa_i, \kappa_{i-1}).
\]
It is customary to represent such a dynamics by a single map \(\mathcal{P}\) by taking the skew product \([52]\) of the maps \(\{f_\gamma\}\). But this is often not necessary: we can represent the dynamics by a single self-map of a smooth manifold \(P\).

We say that a dynamical system \(F: P \to P\) is an ambient system of a DLS \(\mathcal{L}\) if to every trajectory \((k, x)\) of \(\mathcal{L}\) there corresponds a trajectory \(z = (z_i)_{i \in \mathbb{Z}}\) of \(F\). A formal definitions is as follows.

**Definition 1.** A dynamical system \(F: P \to P\) is said to be ambient for the DLS \(\mathcal{L}\) if there exist two systems of open sets \(\mathcal{P}_\gamma^\pm \subset P\) and diffeomorphisms \(\beta_\gamma^\pm: \mathcal{D}_\gamma^\pm \to \mathcal{P}_\gamma^\pm\) which conjugate \(f_\gamma\) with \(F\); in other words, if the diagram

\[
\begin{array}{ccc}
\mathcal{D}_\gamma^- & \xrightarrow{f_\gamma} & \mathcal{D}_\gamma^+ \\
\beta_\gamma^- & \downarrow & \beta_\gamma^+
\end{array}
\quad
\begin{array}{ccc}
\mathcal{P}_\gamma^- & \xrightarrow{F|_{\mathcal{P}_\gamma^-}} & \mathcal{P}_\gamma^+
\end{array}
\]

commutes. Thus, the \(\beta_\gamma^\pm\) are charts on \(P\), and \(f_\gamma\) is a coordinate representation of \(F\).

An ambient system exists under weak conditions on the DLS, but in general it is not unique. For the globally defined DLS in §3.1 the ambient system is the map \(T: M \times M \to M \times M\). The problem of the construction of an ambient system is not important for our purposes, because we usually start with a dynamical system \(F: P \to P\) and the corresponding DLS \(\mathcal{L}\) is a technical tool for studying the dynamics of \(F\). A standard example is the Lagrangian representation of a symplectic map in the next subsection.

### 3.3. DLS generated by a symplectic map

A DLS determines a local (in general) symplectic map. Conversely, one can represent the dynamics of a symplectic map by a DLS, but the Lagrangian will be defined only locally. Let

\[
F: P \to P, \quad F^*\omega = \omega, \quad (3.11)
\]

be a symplectic self-map of a symplectic 2\(m\)-dimensional manifold \((P, \omega)\) \([4]\).

Let \((x^-, y^-)\) and \((x^+, y^+)\) be symplectic coordinates in small neighbourhoods \(D^-\) and \(D^+\) of the points \(z^-\) and \(z^+ = F(z^-)\):

\[
\omega|_{D^-} = dy^- \wedge dx^- \quad \text{and} \quad \omega|_{D^+} = dy^+ \wedge dx^+.
\]

If \(F(D^-) \cap D^+ \neq \emptyset\), then the map \(F\) can be represented as

\[
y^+ = y^+(x^-, y^-), \quad x^+ = x^+(x^-, y^-), \quad dy_+ \wedge dx_+ = dy_- \wedge dx_- . \quad (3.12)
\]

The coordinates can be chosen so that \(\det(\partial y^+/\partial x^-) \neq 0\). Then \(F: D^- \to D^+\) is locally determined by a generating function \(L(x^-, x^+)\) on an open set \(U^- \times U^+\):

\[
y^- = \frac{\partial L}{\partial x^-} \quad \text{and} \quad y^+ = - \frac{\partial L}{\partial x^+}.
\]
Since the construction is local, we obtain a collection $\mathcal{L} = \{L_\kappa\}_{\kappa \in J}$ of generating functions defined on open sets $U^-_\kappa \times U^+_\kappa$. More precisely, take a collection of symplectic charts on $P$, that is, open sets $\{D_k\}_{k \in I}$ and symplectic maps

$$\phi_k: D_k \to U_k \times \mathbb{R}^m, \quad U_k \subset \mathbb{R}^m, \quad \phi_k^*(dy \wedge dx) = \omega|_{D_k}.$$ 

Let $J$ be the set of $\kappa = (\kappa_-, \kappa_+) \in I^2$ such that $F(D_{\kappa_-}) \cap D_{\kappa_+} \neq \emptyset$. For any $\kappa \in J$ the map $F: D_{\kappa_-} \cap F^{-1}(D_{\kappa_+}) \to F(D_{\kappa_-}) \cap D_{\kappa_+}$ satisfies

$$F(x^-, y^-) = (x^+, y^+) \iff y^+ dx^+ - y^- dx^- = dL_\kappa$$

for some function $L_\kappa$. Changing coordinates if needed, we can express $L_\kappa$ locally as a smooth function $L_\kappa(x^-, x^+)$ on a subset of $U^-_{\kappa_-} \times U^+_{\kappa_+}$. We obtain a DLS $\mathcal{L} = \{L_\kappa\}_{\kappa \in J}$. The corresponding graph has vertex set $J$, and there is an edge from $\kappa$ to $\kappa' \in J$ if $\kappa_+ = \kappa'_-$. An orbit $z = (z_i)$ of $F$ such that $z_i \in D_{k_i}$ for all $i$ corresponds to a trajectory $(\kappa, x)$ of the DLS. Here $x_i \in U_{k_i}$ and $\kappa_i = (k_i, k_{i+1})$. Any trajectory of $\mathcal{L}$ gives a unique trajectory of $F$.

If the charts $\{D_k\}_{k \in I}$ cover $P$, then any trajectory of $F$ corresponds to a trajectory (maybe not unique) of the DLS $\mathcal{L}$. However, $\bigcup_{k \in I} D_k \neq P$ in most of the examples below.

The map $F: P \to P$ determines an ambient dynamical system for the DLS $\mathcal{L}$. A concrete example is given in §5.4.

**Remark 1.** There are other ways to represent trajectories of a symplectic map $F$ by a DLS. If equations (3.12) can be solved for $y^-(x^-, x^+)$, then $F: D^- \to D^+$ can be locally represented by a generating function $S(x^-, y^+)$:

$$F(x^-, y^-) = (x^+, y^+) \iff dS(x^-, y^+) = y^- dx^- + x^+ dy^+.$$ 

We obtain a collection of functions $S_\kappa$ which generate symplectic maps $F: D^-_\kappa \to D^+_\kappa$ of open sets in $P$. Define the discrete Lagrangian by

$$L_\kappa(z^-, z^+) = \langle x^-, y^- \rangle - S_\kappa(x^-, y^+), \quad z^\pm = (x^\pm, y^\pm).$$

Then the orbits $z = (z_i)$ of $F$ are critical points of the Poincaré discrete action functional

$$\mathscr{A}_k(z) = \sum L_{\kappa_i}(z_i, z_{i+1}).$$ 

(3.13)

Thus, $z$ is a trajectory of a DLS with $2m$ degrees of freedom. This representation is standard in symplectic topology (see, for example, [59]).

### 3.4. DLS generated by a Lagrangian flow.

Discrete Lagrangian systems arise in a natural way in Lagrangian systems with continuous time (CLS). We consider a CLS with the Lagrangian $L(q, \dot{q}, t)$ on $TM \times \mathbb{R}$, where $M$ is a smooth manifold (the configuration space). We assume that $L$ satisfies the Legendre condition: the Legendre transform $\dot{q} \mapsto p = \partial_\dot{q}L$ is a diffeomorphism. Then $L$ defines a Lagrangian flow on $TM$ or a Hamiltonian flow $\phi^t$ with Hamiltonian $H(q, p, t) = \langle p, \dot{q} \rangle - L$ on $T^*M$. 


If the Lagrangian is $T$-periodic in the time, then the dynamics is described by the monodromy (Poincaré) map $\phi^T: T^*M \to T^*M$. This map is symplectic and can be represented by a DLS. For Lagrangian systems this can be done more explicitly.

Define the Hamilton action function as the action of a trajectory $\gamma: [0, T] \to M$ joining two points $x_-, x_+ \in M$:

$$S(x_-, x_+) = \int_0^T L(\gamma(t), \dot{\gamma}(t), t) \, dt.$$  

If the endpoints $x_\pm$ are non-conjugate along $\gamma$, then $S$ is locally well defined and smooth. By Hamilton’s first-variation formula, the initial and final momenta $p_- = p(0)$ and $p_+ = p(T)$ satisfy

$$p_+ \, dx_+ - p_- \, dx_- = dS(x_-, x_+). \quad (3.14)$$

Therefore, $S$ is the local generating function of the monodromy map $(q_-, p_-) \mapsto (q_+, p_+)$. 

In general there can be several trajectories joining $x_-$ and $x_+$, so the function $S$ is multivalued. Thus, we have a collection of generating functions $S_\kappa$ defined on open sets $U^{-}_\kappa \times U^{+}_\kappa \subset M^2$. Let $\gamma: \mathbb{R} \to M$ be a trajectory of the CLS, and let $x_i = \gamma(t_i)$ and $t_i = Ti$. If the points $x_i$ and $x_{i+1}$ are non-conjugate along $\gamma$, then $x = (x_i)$ is a trajectory of the DLS $\mathcal{L} = \{S_\kappa\}$. Under the Legendre condition, the converse is also true: a trajectory $x$ of the DLS determines a trajectory $\gamma$ of the CLS with momentum satisfying

$$p(t_i + 0) = -\partial_{x_i} L_{\kappa_i}(x_i, x_{i+1}) = \partial_{x_i} L_{\kappa_{i-1}}(x_{i-1}, x_i) = p(t_i - 0).$$

Indeed, then the momentum determines the velocity, and thus $\Delta p(t_i) = 0$ implies that $\Delta \dot{\gamma}(t_i) = 0$. Hence, $\gamma$ is smooth at $t_i$ and so is a trajectory of the CLS. Moreover, the discrete action functional corresponds to the action functional: $\int_\gamma L \, dt = A_K(x)$.

Remark 2. This construction can be generalized. Take hypersurfaces $\{\Sigma_k\}_{k \in I}$ in $M \times \mathbb{R}$. For $\kappa = (\kappa_-, \kappa_+) \in I^2$ define the discrete Lagrangian $L_{\kappa}(x_-, x_+)$ with $x_\pm = (q_\pm, t_\pm) \in \Sigma_{\kappa_\pm}$ as the action of a trajectory $\gamma: [t_-, t_+] \to M$ joining a pair of non-conjugate points $q_-, q_+$. In general there can be several such trajectories, so we obtain a collection of Lagrangians defined on open sets in $\Sigma_{\kappa_-} \times \Sigma_{\kappa_+}$. We use this construction in § 6.3.

For autonomous systems, any $T > 0$ is a period and can be used to define the monodromy map $\phi^T$. Then $\phi^T$ has the energy integral $H$ and the symmetry group defined by the Hamiltonian flow $\phi^t$. The corresponding DLS will have the energy integral and a (local) symmetry group. Hence, this DLS cannot be anti-integrable.

We can avoid this difficulty by replacing the monodromy map $\phi^T$ by the Poincaré map. Fix the energy and consider a local cross-section $N$ (transversal to the flow) on the energy level $H = E$. We can do this by choosing a hypersurface $\Sigma \subset M$ (in applications $\Sigma = \bigcup \Sigma_k$ can be a union of several hypersurfaces) and setting

$$N = \{(q, p): q \in \Sigma, \, H(q, p) = E\}. \quad (3.15)$$
This is a symplectic manifold, and the Poincaré map \( F: U \rightarrow N \) is a symplectic map on an open set \( U \subset N \).

We represent \( F \) by a DLS with \( m - 1 \) degrees of freedom as follows. Define the discrete Lagrangian as the Maupertuis action

\[
S(x_-, x_+) = \int_{\gamma} p \, dq, \quad p = \partial_\dot{q} L,
\]

of a trajectory \( \gamma: [0, T] \rightarrow M \) with \( H = E \) joining the given points \( x_\pm \in M \). Here \( T = T(\gamma) > 0 \) is arbitrary. The action \( S \) is smooth if \( x_\pm \) are non-conjugate along \( \gamma \) for the Maupertuis action functional. The action satisfies (3.14), but the Hamilton–Jacobi equation \( H(x_+, \partial x_+ S) = E \) implies that the twist condition always fails.

A trajectory \( \gamma \) does not always exist, and there may be several of them, so we obtain a collection \( S_\kappa \) of discrete Lagrangians on open sets \( U^-_\kappa \times U^+_\kappa \subset M^2 \). Take a hypersurface \( \Sigma \subset M \) and let \( \Sigma^\pm_\kappa = \Sigma \cap U^\pm_\kappa \). Then \( L_\kappa = S_\kappa|_{\Sigma^-_\kappa \times \Sigma^+_\kappa} \) is a local generating function of the Poincaré map of the cross-section (3.15). Thus, \( \mathcal{L} = \{ L_\kappa \} \) is a DLS describing trajectories of the continuous Lagrangian system on the energy level \( H = E \). If \( \gamma: \mathbb{R} \rightarrow M \) is a trajectory with \( H = E \) such that \( \gamma(t_i) \in \Sigma \) and the points \( x_i \) and \( x_{i+1} \) are non-conjugate along \( \gamma \) (for fixed energy), then \( \mathbf{x} = (x_i) \) is a trajectory of the DLS.

We remark that not all trajectories of the DLS correspond to trajectories of the continuous Lagrangian system: some of them correspond to billiard trajectories with elastic reflection from \( \Sigma \). Indeed, for given \( x \in \Sigma \) the equations

\[
H(x, p) = H(x, p') \quad \text{and} \quad \Delta p = p' - p - \perp T_x \Sigma
\]
do not imply that \( p' = p \). If \( H \) is convex in \( p \), there are two solutions for \( p' \): one describes elastic reflection from \( \Sigma \) and the other \((p' = p)\) describes a smooth trajectory.

However, in the local problems discussed in our paper this difficulty does not appear: locally, the equations above have a unique solution.

4. Anti-integrable limit in DLS

4.1. Main theorem. The original idea of the AI limit [6] has been extended in various directions. The case of the multidimensional standard map was considered in [57] and other papers (see the references in [42]). The general equivariant case was presented in [70]. Billiard systems with small convex scatterers were studied in [30]. The case of a particle of small mass in a potential force field as a continuous Lagrangian system near the AI limit was treated in [18]. Other examples and references can be found in §6. In this section we present a general approach to the AI limit in DLS which covers essentially all the known examples.

Consider a DLS \( \mathcal{L} = \{ L_\kappa \}_{\kappa \in J} \) as defined in §3.2. In the AI limit the Lagrangians are small perturbations of twistless Lagrangians \( L_0^\kappa \):

\[
L_\kappa(x, y) = L_0^\kappa(x, y) + u_\kappa(x, y), \quad \kappa \in J,
\]
where the function $u_\kappa$ is small in the $C^2$ norm and $L^0_\kappa$ has identically zero twist: $B^0_\kappa(x, y) \equiv 0$. Then

$$L^0_\kappa(x, y) = V^-_\kappa(x) + V^+_\kappa(y),$$

where the functions $V^\pm_\kappa$ are defined on the respective sets $U^\pm_\kappa$.

We define an oriented graph $\Gamma$ with vertex set $J$ and edge set $E$ as follows. Vertices $\kappa, \kappa' \in J$ are joined by an edge $\gamma = (\kappa, \kappa')$ if $W_\gamma = U^+_\kappa \cap U^-_{\kappa'} \neq \emptyset$ and the function $\Psi_\gamma = V^+_\kappa + V^-_{\kappa'}$ on $W_\gamma$ has a non-degenerate critical point. Then to each edge $\gamma = (\kappa, \kappa') \in E$ joining vertices $\kappa, \kappa' \in J$ there corresponds a non-degenerate critical point $a_\gamma$ of the function $\Psi_\gamma$.

**Remark 3.** In many applications the critical point $a_\gamma$ will be unique for given $\kappa$ and $\kappa'$. If there are several non-degenerate critical points, then we obtain a richer symbolic dynamics by joining $\kappa$ and $\kappa'$ by several edges $\gamma$, and to each of them there corresponds a non-degenerate critical point $a_\gamma$. In this case $\Gamma$ is not a simple graph: there can be several edges $\gamma$ joining two vertices $\kappa, \kappa'$. Theorem 5 below and its proof still work, but a path in the graph will mean a sequence of edges, not vertices.

According to the traditional definition, a path in the graph $\Gamma$ is a sequence of vertices $k = (\kappa_i)_{i \in \mathbb{Z}}$. It defines a sequence of edges $(\gamma_i)_{i \in \mathbb{Z}}$, where $\gamma_i$ joins the vertex $\kappa_{i-1}$ with $\kappa_i$. The corresponding sequence of critical points

$$a(k) = (a_{\gamma_i})_{i \in \mathbb{Z}}$$

is a trajectory of the twistless Lagrangian system $\mathcal{L}^0 = \{L^0_\kappa\}$, that is, a critical point of the uncoupled action functional:

$$A^0_k(x) = \sum_{k \in \mathbb{Z}} \Psi_{\gamma_i}(x_i). \tag{4.1}$$

The set $\Pi \subset J^{\mathbb{Z}}$ of all paths is invariant under the shift $\mathcal{T}: \Pi \to \Pi$, $(\kappa_i) \mapsto (\kappa_{i+1})$. For simplicity suppose that the graph $\Gamma$ is finite. The general case is discussed later in §4.3. Under this assumption the set $\Pi$ of all paths is a compact, shift-invariant subset of the Cantor set $J^{\mathbb{Z}}$. The dynamical system $\mathcal{T}: \Pi \to \Pi$ is called a topological Markov chain or a subshift of finite type [52]. When the graph is complete, that is, $\Pi = J^{\mathbb{Z}}$, this is the Bernoulli shift with $\#J$ symbols.

**Theorem 5.** Suppose that $\#\Gamma < \infty$. Then there exist $C, \varepsilon_0, \sigma > 0$ such that if

$$\varepsilon = \max_\kappa \|u_\kappa\|_{C^2} < \varepsilon_0, \tag{4.2}$$

then the following are true.

(a) For any path $k = (\kappa_i)_{i \in \mathbb{Z}}$ in $\Gamma$ there exists in the DLS $\mathcal{L}$ a unique orbit $x(k) = (x_i)$ which shadows the sequence $a = a(k)$:

$$\rho(x, a) = \sup_i \text{dist}(x_i, a_{\gamma_i}) < \sigma.$$

Moreover, $\rho(x, a) < C\varepsilon$. 

The orbit $\mathbf{x}$ is hyperbolic with Lyapunov exponent $\lambda \geq C |\log \varepsilon|$.

(c) The map $\mathbf{k} \mapsto \mathbf{x}$ is continuous, so that $\Lambda = \{(\mathbf{k}, \mathbf{x}(\mathbf{k})) : \mathbf{k} \in \Pi\}$ is a compact set in $\mathbb{Z}^2 \times (\mathbb{R}^m)^2$.

Remark 4. The assumption that $\Psi_\gamma$ has a non-degenerate critical point can be weakened. For example, for the existence of a shadowing trajectory it is sufficient that $\Psi_\gamma$ has a compact set of minimum points in $W_\gamma$. However, then the shadowing orbit may fail to be unique or hyperbolic.

If the graph $\Gamma$ is branched, that is, there are 2 cycles passing through the same vertex, then $\mathcal{T} : \Lambda \to \Lambda$ has positive topological entropy. If the DLS admits an ambient system $F : \mathcal{P} \to \mathcal{P}$, then we obtain a homeomorphism $\Pi \to \mathcal{K}$ onto a compact hyperbolic set $\mathcal{K} \subset \mathcal{P}$, and $F : \mathcal{K} \to \mathcal{K}$ is conjugate to the topological Markov chain $\mathcal{T} : \Pi \to \Pi$. In this case we have the following.

**Corollary 1.** If the graph $\Gamma$ is branched, then the set $\mathcal{K}$ has positive topological entropy satisfying (2.18).

Remark 5. For our purposes only the local existence of an ambient system in a neighbourhood of the hyperbolic set is needed. Thus, only the local twist condition

$$\det B_\kappa(a_\gamma, a_{\gamma'}) \neq 0$$

is required for any two edges such that $\gamma$ ends at $\kappa$ and $\gamma'$ starts at $\kappa$.

Theorem 5 does not require the twist condition. Hence, we have to generalize the usual definition of hyperbolicity. In fact, without the twist condition the present state of a trajectory does not determine the past and the future, so the usual definition of the stable (contracting) and unstable (expanding) subspaces does not work. We reformulate the cone hyperbolicity criterion due to Alexeyev [52].

Linearization of equation (3.10) at $\mathbf{x}$ yields the variational equation of the trajectory $(\mathbf{k}, \mathbf{x})$:

$$G_i - u_{i-1} + G_i u_i + G_i u_{i+1} = 0,$$  \hspace{1cm} (4.3)

where $G_i$ and $G_{i\pm}$ are linear operators. We define the cones

$$H_i = \{(u_i, u_{i+1}) : \|u_i\| \leq \alpha_H \|u_{i+1}\|\},$$  \hspace{1cm} (4.4)

$$V_i = \{(u_i, u_{i+1}) : \|u_{i+1}\| \leq \alpha_V \|u_i\|\}.$$  \hspace{1cm} (4.5)

We say that $(\mathbf{k}, \mathbf{x})$ is hyperbolic if there exists a $\mu > 1$ such that for any $i \in \mathbb{Z}$ and for any $u_{i-1}$, $u_i$, and $u_{i+1}$ satisfying (4.3)

$$(u_{i-1}, u_i) \in H_{i-1} \quad \text{implies that} \quad (u_i, u_{i+1}) \in H_i \quad \text{and} \quad \|(u_i, u_{i+1})\| \geq \mu \|(u_{i-1}, u_i)\|,$$  \hspace{1cm} (4.6)

$$(u_i, u_{i+1}) \in V_i \quad \text{implies that} \quad (u_{i-1}, u_i) \in V_{i-1} \quad \text{and} \quad \|(u_{i-1}, u_i)\| \geq \mu \|(u_i, u_{i+1})\|.$$  \hspace{1cm} (4.7)

The cone definition has this form only for a good choice of the metric. But this is not important for us because we use it only as a sufficient condition for hyperbolicity.

If $(u_i, u_{i+1}) \in H_i$, then $\|(u_j, u_{j+1})\| \geq \mu^{j-i} \|(u_i, u_{i+1})\|$ for $j \geq i$, so that the Lyapunov exponent is at least $\log \mu$.  

We say that a compact \( \mathcal{T} \)-invariant set \( \Lambda \) of trajectories \((k, x)\) is hyperbolic if every trajectory \((k, x)\) in \( \Lambda \) satisfies the conditions above with the same constants. If a DLS admits an ambient system, then the corresponding compact invariant set \( \mathcal{K} \subset P \) is hyperbolic in the traditional [52] sense.

**Remark 6.** Another hyperbolicity condition for \((k, x)\) is that the Hessian operator \( A''_k(x) : L_\infty(\mathbb{R}^m) \to L_\infty(\mathbb{R}^m) \) defined by the right-hand side of the variational equation has a bounded inverse in the \( L_\infty \) norm. The equivalence of these definitions was proved in [7].

### 4.2. Proof of Theorem 5.

The proof of Theorem 5 is almost the same as for the standard map. By the implicit function theorem, for any edge \( \gamma = (\kappa, \kappa') \in E \) joining \( \kappa, \kappa' \in J \) the map \( g_\gamma = D\Psi_\gamma \) is a diffeomorphism from a neighbourhood \( W_\gamma \) of the point \( a_\gamma \) to a neighbourhood of the origin in \( \mathbb{R}^m \). We can assume that

\[
W_\gamma = \{ x \in \mathbb{R}^m : |x - a_\gamma| < r \}.
\]  

Here \( r > 0 \) can be taken independent of \( \gamma \) because the graph \( \Gamma \) is finite. Let \( \phi_\gamma \) denote the inverse map: \( \phi_\gamma = g_\gamma^{-1} \). Then \( \phi_\gamma(0) = a_\gamma \).

Let \( k = (\kappa_i) \) be a path in the graph \( \Gamma \), \( (\gamma_i) \) be the corresponding sequence of edges, and consider the metric space \((X, \rho)\), where \( X = \prod_{i \in \mathbb{Z}} W_{\gamma_i} \) and

\[
\rho(x, x') = \sup_{i \in \mathbb{Z}} |x_i - x'_i|
\]

for any \( x, x' \in X \). A trajectory \( x \) is a critical point of the functional \( A_k \) in a neighbourhood of \( a = a(k) \) in \( X \) if and only if

\[
x_i = \phi_{\gamma_i} (\partial_{x_i} (u_{\kappa_{i-1}}(x_{i-1}, x_i) + u_{\kappa_i}(x_i, x_{i+1})))\). \tag{4.9}
\]

It remains to apply the contraction principle in a neighbourhood of the point \( a \) in \( X \). Let \( \mathcal{B} \subset X \) be the ball

\[
\mathcal{B} = \{ x \in X : \rho(x, a) \leq \sigma \}, \quad \sigma < r.
\]

Consider the operator \( \Phi : \mathcal{B}_\sigma \to X \) acting according to the rule

\[
x \mapsto y = \Phi(x), \quad y_i = \phi_{\gamma_i} (\partial_{x_i} (u_{\kappa_{i-1}}(x_{i-1}, x_i) + u_{\kappa_i}(x_i, x_{i+1})))\).
\]

Any fixed point of \( \Phi \) is a trajectory of the DLS \( \mathcal{L} \).

As in §2, it is easy to show that if \( \varepsilon \) is sufficiently small, then \( \Phi(\mathcal{B}_\sigma) \subset \mathcal{B}_\sigma \) and \( \Phi \) is contracting. Indeed,

\[
\rho(y, a) = \sup_{i \in \mathbb{Z}} |\phi_{\gamma_i} (\partial_{x_i} (u_{\kappa_{i-1}}(x_{i-1}, x_i) + u_{\kappa_i}(x_i, x_{i+1}))) - \phi_{\gamma_i}(0)|
\]

\[
\leq \lambda \sup_{i \in \mathbb{Z}} |\partial_{x_i} (u_{\kappa_{i-1}}(x_{i-1}, x_i) + u_{\kappa_i}(x_i, x_{i+1}))| \leq 2\lambda \varepsilon.
\]

Hence, \( y \in \mathcal{B}_\sigma \) if we set \( \sigma = 2\lambda \varepsilon \).
We put \( y = \Phi(x) \) and \( y' = \Phi(x') \), \( x, x' \in \mathcal{B}_\sigma \). Then

\[
\rho(y, y') = \sup_{i \in \mathbb{Z}} |\phi_{\gamma_i}(\partial_{x_i}(u_{\kappa_{i-1}}(x_{i-1}, x_i) + u_{\kappa_i}(x_i, x_{i+1})))
- \phi_{\gamma_i}(\partial_{x_i}(u_{\kappa_{i-1}}(x'_{i-1}, x'_i) + u_{\kappa_i}(x'_i, x'_{i+1})))|
\leq \lambda \sup_{i \in \mathbb{Z}} |\partial_{x_i}(u_{\kappa_{i-1}}(x_{i-1}, x_i) + u_{\kappa_i}(x_i, x_{i+1}))
- \partial_{x_i}(u_{\kappa_{i-1}}(x'_{i-1}, x'_i) + u_{\kappa_i}(x'_i, x'_{i+1})))|
\leq 2 \lambda \varepsilon \rho(x, x').
\]

Therefore, \( \Phi \) is contracting for \( \varepsilon < 1/(2\lambda) \).

Now we prove assertion (b). In (4.4)–(4.5) we take \( \alpha_H = \alpha_V = 1/2 \). Let \( x \) be the orbit corresponding to \( \mathbf{k} \in \Pi \) by assertion (a). Differentiating (4.9), we have

\[
u_i = P_i u_{i-1} + Q_i u_i + R_i u_{i-1}. \tag{4.10}
\]

This equation is equivalent to (4.3).

The linear operators \( P_i \), \( Q_i \), and \( R_i \) have small norms:

\[
\|P_i\| \leq \lambda \varepsilon, \quad \|Q_i\| \leq \lambda \varepsilon, \quad \|R_i\| \leq \lambda \varepsilon.
\]

The equations

\[
R_i u_{i+1} = (I - Q_i) u_i - P_i u_{i-1}, \quad (u_{i-1}, u_i) \in H_{i-1},
\]

imply that

\[
\frac{|u_i|}{|u_{i+1}|} \leq \frac{|(I - Q_i) u_i - P_i u_{i-1}|}{(1 - \|Q_i\| - \|P_i\|/2)|u_{i+1}|} \leq \frac{|R_i u_{i+1}|}{(1 - 3\varepsilon \lambda/2)|u_{i+1}|} \leq \frac{\varepsilon \lambda}{1 - 3\varepsilon \lambda/2}. \tag{4.11}
\]

Therefore,

\[
(u_i, u_{i+1}) \in H_i \quad \text{provided that} \quad \frac{\varepsilon \lambda}{1 - 3\varepsilon \lambda/2} < \frac{1}{2}.
\]

This implies the first assertion in (4.6). We also have

\[
\|(u_{i-1}, u_i)\|^2 \leq |u_{i-1}|^2 + \frac{|(I - Q_i) u_i - P_i u_{i-1}|^2}{(1 - \|Q\| - \|P\|/2)^2}
\leq \frac{|u_i|^2}{4} + \frac{\varepsilon^2 \lambda^2 |u_{i+1}|^2}{(1 - 3\varepsilon \lambda/2)^2} \leq \frac{1}{4} \|(u_i, u_{i+1})\|^2
\]

if \( \varepsilon \lambda/(1 - 3\varepsilon \lambda) < 1/2 \). Hence, we have the second assertion in (4.6) with \( \mu = 2 \). The assertions (4.7) can be checked similarly.

Now we prove (c). If two codes \( \mathbf{k} \) and \( \mathbf{k'} \) are close in the product topology in \( J^\mathbb{Z} \), then they have a long identical segment: \( \kappa_i = \kappa'_i \) for \( |i| \leq n \). And then the corresponding trajectories \( x \) and \( x' \) satisfy

\[
|x_i - x'_i| \leq C \alpha |i|^{-n}, \quad 0 < \alpha < 1, \quad |i| \leq n.
\]

This follows from (4.9) as in the proof of Lemma 2. If \( n \) is large, then \( x \) is close to \( x' \) in the product topology. Therefore, the map \( \mathbf{k} \mapsto x \) is continuous. \( \square \)
Another way to check hyperbolicity is to show that the Hessian $A''_k(x)$ has a bounded inverse in the $l^\infty(\mathbb{R}^m)$-norm. Since it is tridiagonal with an invertible diagonal and small off-diagonal terms, this is almost obvious. We can also describe the stable and unstable subspaces of the hyperbolic trajectory.

**Proposition 1.** For any $u \in \mathbb{R}^m$ there exists a unique trajectory $u = (u_j)_{j \geq 0}$ of the variational system such that $u_0 = u$ and $\|u_j\|$ is bounded as $j \to \infty$. The trajectory $u$ tends exponentially to zero: $\|u_j\| \leq \mu^{-j}\|u\|$, $\mu > 1$. Thus, $u$ belongs to the stable subspace of the trajectory $(k, x)$. The case of the unstable subspace is similar.

The proof of Proposition 1 is obtained by the same contraction principle argument applied to the map

$$(u_i)_{i>0} \mapsto (v_i)_{i>0}, \quad v_i = P_iu_{i-1} + Q_iu_i + R_iu_{i-1}, \quad u_0 = u$$

(see (4.10)), on the space of bounded sequences $(u_j)_{j>0}$.

**4.3. $G$-equivariant DLS.** In many applications one uses the anti-integrable limit to construct trajectories going to infinity. Then it is necessary to consider infinite graphs $\Gamma$. In the case of an infinite $\Gamma$ a certain uniformity is required.

**Condition U** (Uniform anti-integrability). There exist positive constants $r$ and $\lambda$ such that the following hold for any $\gamma \in E$:

(a) $g_\gamma: W_\gamma \to B_\gamma \subset \mathbb{R}^m$ is a diffeomorphism, where $B_\gamma$ is a neighbourhood of the origin and $W_\gamma$ is defined by (4.8) with $r > 0$ independent of $\gamma$;

(b) the map $\phi_\gamma = g_\gamma^{-1}: B_\gamma \to W_\gamma$ is Lipschitz with Lipschitz constant $\lambda$;

(c) $\varepsilon = \sup_{\kappa \in J} \|u_\kappa\|_{C^2}$ is finite and sufficiently small.

Obviously, the condition U holds if $\Gamma$ is finite.

**Theorem 6.** Theorem 5 remains true for an infinite graph $\Gamma$ if the DLS is uniformly anti-integrable.

**Proof.** This coincides with the proof of Theorem 5. □

The condition U is very restrictive. In this subsection we present a class of discrete Lagrangian systems with $\#E = \infty$ in which the condition U requires only finitely many conditions to hold.

Suppose that a discrete group $G$ acts on $M$.

We assume that the action is discrete: any point $x \in M$ has a neighbourhood $U$ such that the sets $g(U)$, $g \in G$, do not intersect ($g'(U) \cap g''(U) = \emptyset$ for $g' \neq g''$). In this case the quotient space $\tilde{M} = M/G$ is a smooth manifold and $\pi: M \to \tilde{M}$ is a covering. The action of $G$ on $M$ generates the diagonal action of $G$ on the product $M \times M$: $g(x, y) = (g(x), g(y))$ for any pair $x, y \in M$ and any $g \in G$.

Let $L: M \times M \to \mathbb{R}$ be invariant with respect to the action of $G$:

$L(x, y) = L(g(x, y))$ \quad for all $x, y \in M$ and $g \in G$.

---

5That is, there is a homomorphism of the group $G$ to the group of diffeomorphisms of $M$. 

The corresponding DLS is said to be $G$-equivariant. The ambient system here can be defined as a map $F: P \to P$ of $P = (M \times M)/G$. In the case of the multidimensional standard map with Lagrangian (3.6) we have $M = \mathbb{R}^m$, and $G \cong \mathbb{Z}^m$ is the group of shifts preserving the potential $V$.

As another example suppose that $M$ is a Riemannian manifold and the group $G$ acts on $M$ by isometries. Let $\text{dist}(\cdot, \cdot)$ be the distance induced by the Riemannian metric. Then the Lagrangian $L(x, y) = \text{dist}^2(x, y)$ is a smooth function for any pair of sufficiently close points $x$ and $y$. It is invariant with respect to the diagonal action of $G$.

For smooth $G$-invariant functions $V: M \to \mathbb{R}$ and $u: M \times M \to \mathbb{R}$ consider the DLS determined by the discrete Lagrangian $L$ with

$$L(x, y) = u(x, y) + V(y).$$

We define the AI limit in such a system as the limit as $\|u\|_{C^2} \to 0$.

If the dynamics in a system with this Lagrangian is globally defined, then it can be assumed that an ambient system is defined.

Suppose that the configuration space $\tilde{M} = M/G$ is compact. Let $\text{dist}(\cdot, \cdot)$ be the corresponding distance on $M$. We take a finite set of non-degenerate critical points of $V$ on $\tilde{M}$, and let $I$ be the corresponding $G$-invariant set of critical points of $V$ on $M$. The latter is a finite union of the orbits

$$\{g(k): g \in G\}, \quad k \in I,$$

of the group action. We take a small $\rho$-neighbourhood of $\text{Cr}$ in $\tilde{M}$, which corresponds to a $G$-invariant neighbourhood $U = \bigcup_{k \in I} U_k$ of $I$ in $M$.

For a large constant $N$ we define a graph $\Gamma$:

$$J = \{(\kappa_-, \kappa_+) \in I^2: \text{dist}(\kappa_-, \kappa_+) < N\}, \quad E = \{\gamma = (\kappa, \kappa') \in J^2: \kappa_+ = \kappa_-\}.$$

In the notation of §4.1,

$$L_\kappa = L|_{U_{\kappa_-} \times U_{\kappa_+}}, \quad V^-_\kappa = 0, \quad V^+_\kappa = V|_{U_{\kappa_+}}, \quad u_\kappa = u|_{U_{\kappa_-} \times U_{\kappa_+}}.$$

If the group $G$ is infinite, then the graph $\Gamma$ is infinite, but due to the $G$-invariance of $L$, the condition $U$ follows from the finiteness of the set $\text{Cr}$ and the boundedness of $N$. One can see that $G$ acts also on $\Gamma$, and $\tilde{\Gamma} = \Gamma/G$ is a finite graph. Then we obtain the following.

**Proposition 2.** The map $F$ has a hyperbolic set $K$ such that $F|_K$ is conjugate to the topological Markov chain determined by the graph $\Gamma$.

In fact we need only the local existence of an ambient system near the hyperbolic set. It will exist if $\det B(\kappa_-, \kappa_+) \neq 0$ for all $(\kappa_-, \kappa_+) \in J$.

Note that if $M$ is not compact, then $M^2/G$ is also not compact, but the set $\mathcal{K}$ is compact because all its points are at a distance less than $N + \sigma$ from the compact set $\{(x, x) \in M^2\}/G$. 
5. Examples: systems with discrete time

5.1. Light particle and periodic kicks. Consider a particle with a small mass \( \varepsilon^2 \) which moves in \( \mathbb{R}^m \) in the force field with potential

\[
\mathcal{V}(x, t) = \frac{1}{2\pi} V(x) \delta(t),
\]

(5.1)

where the function \( V \) is smooth on \( \mathbb{R}^m \) and \( \delta \) is a periodic δ-function:

\[
\delta(t) = \begin{cases} \infty, & t \in 2\pi\mathbb{Z}, \\ 0, & t \in \mathbb{R} \setminus 2\pi\mathbb{Z}, \end{cases}
\]

\[
\int_{2\pi k - \sigma}^{2\pi k + \sigma} \delta(t) \, dt = 1 \quad \text{for any } k \in \mathbb{Z}, \sigma \in (0, \pi).
\]

The Hamiltonian of the system has the form

\[
H = \frac{1}{2\varepsilon^2} |p|^2 + \frac{1}{2\pi} V(x) \delta(t),
\]

where \( p = (p_1, \ldots, p_m) \) is the momentum canonically conjugate to the coordinates \( x = (x_1, \ldots, x_m) \). The Hamiltonian equations are

\[
\dot{p} = -\frac{1}{2\pi} \frac{\partial V}{\partial x}(x) \delta(t), \quad \dot{x} = \frac{p}{\varepsilon^2}.
\]

Therefore, \( p \) gets increments \( -(2\pi)^{-1} \partial V/\partial x \) at the moments of time \( 2\pi l, l \in \mathbb{Z} \). During the remaining time the particle is free.

For any integer \( l \) we put \( x(2\pi l - 0) = x_l \) and \( p(2\pi l - 0) = p_l \). Then

\[
\begin{pmatrix} x_l \\ p_l \end{pmatrix} \mapsto \begin{pmatrix} x(2\pi l + 0) \\ p(2\pi l + 0) \end{pmatrix} = \begin{pmatrix} x_l \\ p_l - \frac{1}{2\pi} \frac{\partial V}{\partial x}(x_l) \end{pmatrix}
\]

\[
\mapsto \begin{pmatrix} x_{l+1} \\ p_{l+1} \end{pmatrix} = \begin{pmatrix} x_l + 2\pi \varepsilon^{-2} p_{l+1} \\ p_l - \frac{1}{2\pi} \frac{\partial V}{\partial x}(x_l) \end{pmatrix}.
\]

The quantities \( x_{l-1}, x_l, \) and \( x_{l+1} \) satisfy the equation

\[
x_{l+1} - 2x_l + x_{l-1} = \varepsilon^{-2} \frac{\partial V}{\partial x}(x_l).
\]

We have a DLS with Lagrangian of the form (3.6), where \( B = \varepsilon^2 I \):

\[
L(x, y) = \varepsilon^2 \frac{|x - y|^2}{2} - V(y).
\]

(5.2)

The corresponding discrete dynamical system is the multidimensional standard map. By Theorem 5, if \( V \) has two non-degenerate critical points, then for small \( \varepsilon \) the system has a chaotic hyperbolic set. If \( V \) is \( \mathbb{Z} \)-periodic on \( \mathbb{R}^m \), then we have a \( \mathbb{Z} \)-equivariant DLS corresponding to a symplectic twist self-map of \( \mathbb{T}^m \times \mathbb{R}^m \).
Note that for small $\varepsilon$ the system remains close to the AI limit if the potential $V$ is a smooth periodic function close to (5.1) (as a distribution).

If a light particle travels in a potential force field, and the potential $V(x,t)$ does not satisfy (5.1), then the theory of the AI limit becomes technically more complicated. We discuss these results in §6.3.

5.2. Billiards in a wide strip. Consider a plane billiard system in a wide strip bounded by the graphs of two 1-periodic functions. In other words, we assume that a particle is moving in the domain

$$D = \{(x,y) \in \mathbb{R}^2: f_1(x) \leq y \leq f_2(x) + d\},$$

where $f_1$ and $f_2$ are 1-periodic functions and the parameter $d$ is large (Fig. 1). The motion of the particle inside the domain is assumed to be free, and reflections from the boundary are elastic.

![Figure 1. Billiards in a wide strip.](image)

This is a billiard system (see §3.1). Let $L$ be the length of the line segment between two subsequent points of impact with the boundary. We will consider motions such that the particle collides alternately with the upper and lower walls.

Let $x_1$ be the coordinate on the lower boundary and $x_2$ the coordinate on the upper boundary. The length of the corresponding line segment is

$$L(x_1, x_2) = \sqrt{(x_2 - x_1)^2 + (d + f_2(x_2) - f_1(x_1))^2}.$$ 

The Lagrangian is $\mathbb{Z}$-invariant and for large $d$ has the form (3.6):

$$L(x_1, x_2) = d + f_2(x_2) - f_1(x_1) + \frac{\varepsilon}{2}(x_2 - x_1)^2 + O(\varepsilon^2), \quad d = \varepsilon^{-1}.$$ 

The action functional has the form

$$A(x) = \sum_{i \in \mathbb{Z}} (L(x_{2i-1}, x_{2i}) + L(x_{2i+1}, x_{2i})).$$ 

In the notation of §3.2 we have a DLS $\mathcal{L} = \{L_1, L_2\}$ with the configuration space $M = \mathbb{R}_1 \cup \mathbb{R}_2$ (a union of two copies of $\mathbb{R}$) and with the Lagrangians $L_1(x_1, x_2) = L(x_1, x_2)$ and $L_2(x_2, x_1) = L(x_1, x_2)$. The action functional

$$A(x) = \sum_{i \in \mathbb{Z}} L_{\kappa_i}(x_i, x_{i+1}) = \sum_{i \in \mathbb{Z}} (d + 2(-1)^i f_{\kappa_i}(x_i) + O(\varepsilon)), \quad \kappa_i = i \mod 2,$$
has the uncoupled form (4.1). Thus, the DLS is anti-integrable if both functions $f_1$ and $f_2$ have non-degenerate critical points. Let $Cr_1 \subset \mathbb{R}_1$ and $Cr_2 \subset \mathbb{R}_2$ denote the $\mathbb{Z}$-invariant sets of non-degenerate critical points of $f_1$ and $f_2$. Suppose that the sets $Cr_1/\mathbb{Z}$ and $Cr_2/\mathbb{Z}$ are finite and non-empty.

The corresponding graph $\Gamma$ has vertices of two types: $J = J_1 \cup J_2$, where
\[ J_1 = \{ (\kappa, x, y) \in \{1\} \times Cr_1 \times Cr_2 : |x - y| < N \}, \]
\[ J_2 = \{ (\kappa, x, y) \in \{2\} \times Cr_2 \times Cr_1 : |x - y| < N \}. \]

There is an edge joining the vertices $(\kappa, x, y)$ and $(\kappa', x', y')$ if and only if $\kappa \neq \kappa'$ and $y = x'$.

As in the standard equivariant situation (§4.3), we obtain the existence of a hyperbolic set carrying a dynamics conjugate to the dynamics in a topological Markov chain. The graph $\Gamma$ is infinite, but the condition of uniform anti-integrability holds.

The multidimensional analogue of this system is straightforward: it is sufficient to say that $x \in \mathbb{R}^m$ and the functions $f_1$ and $f_2$ are periodic in all components of the vector variable $x$.

5.3. Billiard systems with small scatterers. Let $D \subset \mathbb{R}^m$ be a domain with smooth boundary $\Sigma = \partial D$. Following [30], we consider a billiard system in the domain $\Omega_\varepsilon = D \setminus (\bigcup_{j=1}^N D_j)$, where the $D_j \subset D$ are small subdomains playing the role of scatterers. They are small in the following sense. Each domain $D_j$ is associated with some point $a_j \in D$. The boundary $\partial D_j$ has the form
\[ a_j + \varepsilon \Sigma_j = \{ q \in \mathbb{R}^m : q = a_j + \varepsilon \phi_j(x), x \in S^{m-1} \}, \]
where the vector-functions $\phi_j : S^{m-1} \to \mathbb{R}^m$ are smooth embeddings which define smooth submanifolds $\Sigma_j = \phi_j(S^{m-1})$. If $\varepsilon$ is sufficiently small, then $D_j \subset D$.

As $\varepsilon \to 0$, $\Omega_\varepsilon$ degenerates into $\Omega_0 = D \setminus A$, where $A = \{a_1, \ldots, a_N\}$.

We shall explain that for small $\varepsilon$ this billiard system generates a DLS with the discrete Lagrangian $\mathcal{L} = \{L_\kappa\}$, where the index $\kappa$ corresponds to a non-degenerate billiard trajectory in $D$ starting and ending in $A$ (see Fig. 2).

![Figure 2. A billiard orbit in $\Omega_0$.](image)

Consider a billiard trajectory $\kappa$ in $D$ starting and ending at points $a_i, a_j \in A$. It can be a segment $\kappa = (a_i, a_j)$ joining $a_i$ and $a_j$ or a broken line
\[ \kappa = (a_i, p_1, \ldots, p_n, a_j), \quad p_1, \ldots, p_n \in \Sigma, \] (5.3)
joining the points $a_i, p_1, \ldots, p_n, a_j$. Then $(p_1, \ldots, p_n)$ is a critical point of the length function

$$l(a_i, p_1, \ldots, p_n, a_j) = |a_i - p_1| + |p_1 - p_2| + \cdots + |p_{n-1} - p_n| + |p_n - a_j|$$
on $\Sigma^n$. The trajectory $\kappa$ is said to be non-degenerate if $n = 0$ or the critical point $(p_1, \ldots, p_n)$ is non-degenerate. We call the broken line (5.3) a quasi-trajectory if it is non-degenerate and does not contain points in $A$ other than the endpoints $a_i$ and $a_j$.

Let $a^-_\kappa = a_i$ and $a^+_{\kappa} = a_j$ be the initial and final points of the quasi-trajectory, and let $v^-_{\kappa}$ and $v^+_{\kappa}$ be its initial and final velocity vectors issuing from the points $a_i$ and $a_j$, respectively:

$$v^-_{\kappa} = \frac{p_1 - a_i}{|p_1 - a_i|} \quad \text{and} \quad v^+_{\kappa} = \frac{a_j - p_n}{|a_j - p_n|}.$$  

In the case $n = 0$ we should take $v^-_{\kappa} = v^+_{\kappa} = (a_j - a_i)/|a_j - a_i|$.

If $D$ is convex, then there are always non-degenerate quasi-trajectories with $n = 0$ joining $a_i$ and $a_j$. Also, there always exist quasi-trajectories with $n = 1$ corresponding to the minimum of $l$ on $\Sigma$. Generically, they will be non-degenerate.

For $m = 2$, by using the same arguments as in the Birkhoff theorem [13] on periodic trajectories of a convex billiard system (see also [55]), it is easy to show that if $D$ is convex, then for any $n \geq 1$ there are at least $2n$ orbits (5.3).

Any non-degenerate trajectory (5.3) is smoothly deformed when we slightly change its starting and ending points. Hence, for small $\varepsilon \geq 0$ there exist smooth functions

$$p_1(x, y), \ldots, p_n(x, y), \quad x, y \in S^{m-1},$$  

such that

$$B_{\varepsilon}(x, y) = (a_i + \varepsilon \phi_i(x), p_1(x, y), \ldots, p_n(x, y), a_j + \varepsilon \phi_j(y))$$  

is a billiard trajectory in $D$ which coincides with $\kappa$ when $\varepsilon = 0$. This trajectory is contained in $\Omega_{\varepsilon}$ if the only common points of $B_{\varepsilon}(x, y)$ and $\partial\Omega_{\varepsilon}$ are the endpoints $a_i + \varepsilon q_i(x)$ and $a_j + \varepsilon q_j(y)$. This holds if $\varepsilon$ is sufficiently small and $x \in U^-_{\kappa}$, $y \in U^+_{\kappa}$, where $U^\pm_{\kappa} \subset S^{m-1}$ are some open sets (see Fig. 3). For example, $U^-_{\kappa}$ is the set of $x$ such that the ray starting at $\phi_i(x)$ in the direction of $v^-_{\kappa}$ does not cross $\Sigma_i$.

![Figure 3. The bold part of the boundary $\partial D_i$ is the set $a_i + \varepsilon \phi_i(U^-_{\kappa})$.](image-url)
The anti-integrable limit

The discrete Lagrangian corresponding to $\kappa$ is

$$L_\kappa(x, y) = \varepsilon^{-1} l(B_\varepsilon)$$

$$= \varepsilon^{-1} |a_i - p_1(x, y) + \varepsilon \phi_i(x)| + \varepsilon^{-1} |p_1(x, y) - p_2(x, y)| + \cdots$$

$$+ \varepsilon^{-1} |p_{n-1}(x, y) - p_n(x, y)| + \varepsilon^{-1} |p_n(x, y) - a_j - \varepsilon \phi_j(y)|. $$

Let us show that

$$L_\kappa(x, y) = \varepsilon^{-1} l(\kappa) + \langle v^+_\kappa, \phi_i(x) \rangle - \langle v^-_\kappa, \phi_j(y) \rangle + O(\varepsilon), \quad (x, y) \in S^{m-1} \times S^{m-1}. $$

Indeed, by Hamilton’s first-variation formula,

$$dl(B_\varepsilon) = v^+_\kappa dq_+ - v^-_\kappa dq_-, \quad dq_- = \phi_i(x) d\varepsilon, \quad dq_+ = \phi_j(y) d\varepsilon.$$ 

Therefore,

$$\left. \frac{\partial}{\partial \varepsilon} \right| \varepsilon=0 l(B_\varepsilon) = \langle v^+_\kappa, \phi_j(y) \rangle - \langle v^-_\kappa, \phi_i(x) \rangle.$$ 

Let the vertex set of the graph $\Gamma$ be a finite collection $J = \{\kappa\}$ of quasi-trajectories. We connect vertices $\kappa$ and $\kappa'$ by an edge if

1) the end of $\kappa$ is the start of $\kappa'$: $a^+_\kappa = a^-_{\kappa'} = a_k \in A,$

2) the direction changes: $v^+_\kappa \neq v^-_{\kappa'}.$

If $\Sigma_k$ is strictly convex, then no other conditions are needed. If not, then we need one more. Let $v \in \mathbb{R}^m$ be the unit vector

$$v = \frac{v^+_{\kappa'} - v^-_\kappa}{|v^+_{\kappa'} - v^-_\kappa|}$$

(see Fig. 4). There always exists an $s \in S^{m-1}$ such that $v$ is the outward normal to the tangent plane $T_{\phi_k(s)} \Sigma_k$ and $s \in U^+_\kappa \cap U^-_{\kappa'}.$

3) Assume that the second fundamental form $\langle v, d^2 \phi_k(s) \rangle$ of $\Sigma_k$ at $\phi_k(s)$ is non-degenerate.

The condition 3) always holds if $\Sigma_k$ is strictly convex; then there exists a unique $s$ obtained by maximizing $\langle \phi_k(s), v \rangle$. If $\Sigma_k$ is not convex, then there can be several such points $s$, in which case we have several edges $\gamma$ joining the vertices $\kappa$ and $\kappa'.$

Thus, we have defined a graph with vertex set $J = \{\kappa\}$ and edge set $E = \{\gamma\}$. The Lagrangian $L_\kappa$ has the anti-integrable form with $V^+_\kappa = \langle v^+_\kappa, \phi_j \rangle$ and $V^-_\kappa = -\langle v^-_\kappa, \phi_i \rangle$. Any path in the graph defines a code. Theorem 5 now implies the following.

---

**Figure 4.** The vectors $v^+_\kappa$, $v^-_{\kappa'}$, and $v$. 

---
Theorem 7 (cf. [30]). Suppose that \( \varepsilon \) is sufficiently small. Then for any code \( k \) there is a billiard trajectory \( x \) in \( \Omega_\varepsilon \) shadowing the chain of quasi-trajectories \( k = (\kappa_i) \). The orbit \( x \) is hyperbolic.

The billiard system in \( \Omega_\varepsilon \) is an ambient system for the DLS we have constructed. Lower estimates for the topological entropy of the ambient system can be derived with the help of the topological Markov chain obtained. A more detailed study of the problem of topological entropy for a billiard system with small scatterers is presented in [31].

We remark that if the scatterers are convex, then under certain conditions the billiard system will be hyperbolic, and then much stronger results hold; in particular, the metric entropy will be positive [28].

5.4. Birkhoff–Smale–Shilnikov theorem. As a general example of an application of the AI limit method in studying symplectic maps, we prove the existence of a chaotic hyperbolic invariant set for a discrete dynamical system with a transversal homoclinic trajectory. This theorem goes back to Birkhoff [13] and Smale [65], and in final form it was proved by Shilnikov [63] (see also [52] and [64]).

We will prove the Birkhoff–Smale–Shilnikov theorem for symplectic maps. However, a general (non-symplectic) map can be reduced to a symplectic one by doubling the dimension (see Remark 8 at the end of this subsection).

Let \( F: P \to P \) be a symplectic map of a \( 2m \)-dimensional symplectic manifold and \( O \) a hyperbolic fixed point of \( F \). Suppose that there exist transversal homoclinic trajectories to \( O \), and take a finite collection \( \{ \gamma_k \}_{k \in I} \) of them. We will show that the trajectories of \( F \) which stay in a neighbourhood of the homoclinic set \( \bigcup_{k \in I} \gamma_k \cup O \) can be described by an anti-integrable DLS. This provides a symbolic representation of the trajectories in a neighbourhood of the homoclinic set.

It is convenient to denote by \( W^+ \) and \( W^- \) the stable and unstable manifolds of \( O \):

\[
W^\pm = \{ x: F^n(x) \to O \text{ as } n \to \pm \infty \}.
\]

There exist symplectic coordinates \( q \) and \( p \) in a neighbourhood \( D \) of \( O \) such that \( O = (0, 0) \) and the local stable and unstable manifolds \( W^\pm_{\text{loc}} \subset D \) are Lagrangian graphs over a ball \( U \subset \mathbb{R}^m \):

\[
W^+_{\text{loc}} = \{ (q, p): q \in U, p = -\nabla S_+(q) \},
\quad W^-_{\text{loc}} = \{ (q, p): q \in U, p = \nabla S_-(q) \}.
\]  

Let

\[
\lambda_j, \lambda_j^{-1}, \quad 0 < |\lambda_j| < 1, \quad j = 1, \ldots, m,
\]

be the eigenvalues of the operator \( DF(O) \). We fix

\[
\alpha \in (\max\{ |\lambda_j| \}, 1).
\]

The local dynamics near \( O \) is described by the next statement, which can be deduced from a lemma of Shilnikov [63] or from the strong \( \lambda \)-lemma [39].

\[\text{\textsuperscript{6}The standard notation is } W^s \text{ and } W^u.\]
Lemma 3. Let $N$ be sufficiently large. Then for any $n \geq N$ and $q^\pm \in U$ there exist $p^\pm$ such that $F^n(q^+, p^+) = (q^-, p^-)$ and $F^i(q^+, p^+) \in D$ for $0 \leq i \leq n$. The map $(q^+, p^+) \mapsto (q^-, p^-)$ is a symplectic diffeomorphism $F^n : D^+_n \to D^-_n$ of open subsets of $D$ and is given by the generating function (defined up to a constant)

$$S_n(q^+, q^-) = S_+(q^+) + S_-(q^-) + u_n(q^+, q^-), \quad (5.5)$$

where $\|u_n\|_{C^2(U \times U)} \leq C \alpha^n$.

Thus,

$$F^n(q^+, p^+) = (q^-, p^-) \iff p^+ dq^+ - p^- dq^- = dS_n(q^+, q^-).$$

For large $n$ the generating function $S_n$ of the local map $F^n : D^+_n \to D^-_n$ has an anti-integrable form. To obtain chaotic dynamics we need to come back to a neighbourhood of $O$ by using a global return map along a homoclinic trajectory.

Let $\gamma_k$ be a transversal homoclinic trajectory to the fixed point $O$. Then there are points $z^\pm_k = (q^\pm_k, p^\pm_k) \in \gamma_k \cap W^\pm_{\text{loc}}$ and integers $m_k$ such that $F^{m_k}(z^-_k) = z^+_k$. The map $F^{m_k}$ from a neighbourhood $G^-_k$ of $z^-_k$ to a neighbourhood $G^+_k$ of $z^+_k$ is called the global map. Since this map is symplectic, there exist functions $\Phi_k$ such that

$$F^{m_k}(q^-, p^-) = (q^+, p^+) \iff p^+ dq^+ - p^- dq^- = d\Phi_k. \quad (5.6)$$

By perturbing the coordinates if necessary, we can assume that

$$\det \left( \frac{\partial q^+}{\partial p^-} \right) \neq 0.$$

Then we can express $\Phi_k$ locally as a function $\Phi_k(q^-, q^+)$ on $U^-_k \times U^+_k$, where $U^\pm_k \subset U$ is a neighbourhood of $q^\pm_k$. And then $\Phi_k$ is the generating function of the global map $F^{m_k} : G^-_k \to G^+_k$.

Equations (5.4) and (5.6) imply that the function

$$R_k(q^-, q^+) = S_-(q^-) + \Phi_k(q^-, q^+) + S_+(q^+)$$

has a critical point $(q^-_k, q^+_k)$ which corresponds to the homoclinic trajectory $\gamma_k$. Since $\gamma_k$ is transversal, the critical point is non-degenerate.

The trajectories of $F$ which stay in a neighbourhood of the homoclinic set correspond to trajectories of the DLS with Lagrangians $\{\Phi_k, S_n\}_{k \in I, n \geq N}$. However, formally this system is not anti-integrable. The application of Theorem 5 is simpler if we consider another DLS with 2$m$ degrees of freedom.

Let $J = \{k = (k, n) : k \in I, n \geq N\}$. We define the DLS $\{L_k\}_{k \in J}$ with the discrete Lagrangian $L_k$ representing the composition $F^n \circ F^{m_k}$ of the global map and the local map. Let

$$L_k(x, y) = \Phi_k(x) + S_n(x^+, y^-), \quad x = (x^-, x^+) \in U^-_k \times U^+_k, \quad y = (y^-, y^+) \in U^2.$$

Remark 7. The Lagrangian $L_k$ has 2$m$ degrees of freedom. We can replace it by the reduced Lagrangian

$$\widetilde{L}_k(x^-, y^-) = \text{Crit}_{x^+}(\Phi_k(x^-, x^+) + S_n(x^+, y^-)).$$
with \( m \) degrees of freedom. This requires an extra non-degeneracy condition which can be satisfied by perturbing the coordinates.

The trajectories of \( F \) shadowing the homoclinic chain \((\gamma_{k_i})\) correspond to critical points of the action functional

\[
A_k(x) = \sum L_{k_i}(x_i, x_{i+1}), \quad x_i = (x_i^-, x_i^+) \in U_{k_i}^- \times U_{k_i}^+, \quad k_i = (k_i, n_i).
\]

To obtain an anti-integrable DLS, we replace \( L_{k_i} \) by a gauge-equivalent Lagrangian (with the same action functional)

\[
\hat{L}_{k_i}(x, y) = L_{k_i}(x, y) + S_-(x^-) - S_-(y^-) = R_k(x) + O(\alpha^n),
\]

where \( R_k \) has a non-degenerate critical point \((q^-_k, q^+_k)\). In the notation of Theorem 5 we have an anti-integrable DLS defined by a graph \( \Gamma \) with vertex set \( J \) and any two vertices joined by an edge. If there is no upper bound on \( n_i \), then the graph will be infinite, but it is easy to see that Theorem 6 works. We obtain the Birkhoff–Smale–Shilnikov theorem.

**Theorem 8.** For any code \( \kappa_i = (k_i, n_i) \in J \) the corresponding homoclinic chain \((\gamma_{k_i})\) is shadowed by a unique hyperbolic trajectory of \( F \) which follows \( \gamma_{k_i} \), then stays in a neighbourhood of \( O \) for \( n_i \) iterations of \( F \), then shadows \( \gamma_{k_{i+1}} \), and so on. Thus, \( F \) has a chaotic hyperbolic invariant set.

The standard graph which gives a symbolic representation of the dynamics in a neighbourhood of a homoclinic set is a bit different [52]: there is one vertex for every homoclinic trajectory \( \gamma_k \) and one more for the fixed point \( O \). After \( \gamma_{k_i} \), a path in the graph may stay at \( O \) for several steps \( n_i \) before following \( \gamma_{k_{i+1}} \). Therefore, a path corresponds to a sequence \((k_i, n_i)\) as described above, so the dynamics is equivalent.

**Remark 8.** The Birkhoff–Smale–Shilnikov theorem holds for a general (non-symplectic) map \( f \). Indeed, the map \( q^+ = f(q^-) \) can be reduced to a symplectic map \( F \) with the generating function \( S(q^-, p^+) = (f(q^-), p^+) \) by introducing the conjugate momentum:

\[
p^- = \partial_q S, \quad q^+ = \partial_p S = f(q^-).
\]

If \( f \) has a hyperbolic fixed point possessing transversal homoclinic trajectories, then so does \( F \). Thus, we have proved Theorem 8 also for non-symplectic maps.

### 5.5. Shadowing a chain of invariant tori.

The following application of the AI limit is based on [15]. Let \( F: P \to P \) be a smooth symplectic diffeomorphism which has a \( d \)-dimensional hyperbolic invariant torus \( \Gamma \). Then \( \Gamma \) is the image of a smooth embedding \( h: \mathbb{T}^d \to P \), and \( F|_\Gamma \) is a translation with the rotation vector \( \rho \in \mathbb{R}^d \):

\[
F(h(x)) = F(x + \rho).
\]

If the rotation vector is Diophantine, that is,

\[
|\langle \rho, j \rangle - j_0| \geq \alpha|j|^{-\beta}, \quad \alpha, \beta > 0,
\]
for all \( j \in \mathbb{Z}^d \setminus \{0\} \) and \( j_0 \in \mathbb{Z} \), then the torus is said to be Diophantine. We assume the torus to be isotropic, that is, \( \omega|_\Gamma = 0 \). If the torus is Diophantine and the symplectic structure is exact, this is always so.

**Definition 2.** The torus \( \Gamma \) is said to be hyperbolic if there exist two smooth \((m - d)\)-dimensional subbundles \( E^\pm \) of the tangent bundle \( T_\Gamma P \) such that:

(i) the \( E^\pm \) are invariant under the linearized map \( DF \), that is,

\[
DF(x)E^\pm_x = E^\pm_F(x) \quad \text{for all } x \in \Gamma;
\]

(ii) the linearized map is contracting on \( E^+ \) and expanding on \( E^- \), that is, for some \( c > 0 \) and \( \lambda > 1 \) and for all \( x \in \Gamma \),

\[
\|DF^k(x)|_{E^+_x}\| \leq c\lambda^{-k} \quad \text{and} \quad \|DF^{-k}(x)|_{E^-_x}\| \leq c\lambda^{-k}, \quad k \in \mathbb{N}.
\]

We fix some Riemannian metric, and let \( \| \cdot \| \) be the operator norm defined by it. Since \( \Gamma \) is compact, the definition is independent of the metric.

**Definition 3.** A torus \( \Gamma \) is non-degenerate if all bounded trajectories of the linearized map are tangent to \( \Gamma \). Thus, if \( x \in \Gamma, v \in T_x M, \) and \( \|DF^k(x)v\| \leq c \) for all \( k \in \mathbb{Z} \), then \( v \in T_x \Gamma \).

We can rewrite the definition of a hyperbolic torus in coordinate form as follows.

**Definition 4.** An invariant torus \( \Gamma \) is said to be hyperbolic if in a tubular neighbourhood \( D \) of it there exist symplectic coordinates \( x \in \mathbb{T}^d, y \in \mathbb{R}^d, \) and \( z^\pm \in \mathbb{R}^{m-d} \) such that:

(a) \( \omega|_D = dy \wedge dx + dz_+ \wedge dz_-; \)

(b) \( \Gamma \) is given by the equations \( y = 0, z_- = z_+ = 0; \)

(c) the map \( F|_D \) has the form

\[
\begin{pmatrix}
x \\
y \\
z_- \\
z_+
\end{pmatrix}
\mapsto
\begin{pmatrix}
x + \rho + Ay \\
y \\
B^*(x)^{-1}z_- \\
B(x)z_+
\end{pmatrix}
+ O_2(y, z_-, z_+);
\]

(d) the Lyapunov exponents for the skew product map \((x, z) \mapsto (x + \rho, B(x)z)\)

are negative, that is, there is a norm such that \( \|B(x)\| \leq \alpha < 1; \)

(e) the symmetric matrix \( A \) is constant.

For a Diophantine isotropic torus Definitions 2 and 4 are equivalent (see [25]). If we use Definition 4, then for what follows we do not need to assume that the torus is Diophantine. However, hyperbolic invariant tori arising in applications are usually Diophantine. A hyperbolic torus is non-degenerate if \( \det A \neq 0 \). By the KAM-theory, hyperbolic non-degenerate Diophantine invariant tori persist under small smooth exact symplectic perturbations of the map.

The hyperbolic torus \( \Gamma \) has \( m \)-dimensional stable manifold \( W^+ \) and unstable

manifold \( W^- \) consisting of trajectories \( F^n(z) \) that are asymptotic to \( \Gamma \) as \( n \to +\infty \) and \( n \to -\infty \), respectively.
Let \( \{\Gamma_k\}_{k \in I} \) be a finite set of non-degenerate hyperbolic tori of dimension \( 0 < d_k < m \), and let \( W_k^± \) be their stable and unstable manifolds. A point \( z \in W_j^- \cap W_k^+ \) is called a heteroclinic point, and its orbit \( \gamma = (F^n(z))_{n \in \mathbb{Z}} \) is called a heteroclinic orbit connecting \( \Gamma_j \) with \( \Gamma_k \). Such a trajectory is said to be transversal if \( T_z W_j^- \cap T_z W_k^+ = \{0\} \).

Let \( \{\gamma_k\}_{k \in J} \) be a finite set of transversal heteroclinic orbits connecting pairs of tori in the set \( \{\Gamma_k\}_{k \in I} \). Let \( G \) be the oriented graph with vertices \( k \in I \) and edges \( \kappa \in J \) corresponding to heteroclinic orbits.

**Theorem 9.** Let \( N > 0 \) be sufficiently large and take an arbitrary sequence of integers \( (n_i)_{i \in \mathbb{Z}} \), \( n_i \geq N \). Let \( \kappa = (\kappa_i)_{i \in \mathbb{Z}} \) be a path on the graph \( G \) corresponding to the sequence \( (\gamma_{\kappa_i})_{i \in \mathbb{Z}} \) of heteroclinic trajectories connecting \( \Gamma_{k_{i-1}} \) with \( \Gamma_{k_i} \). Then there exists a (non-unique) trajectory of \( F \) shadowing the heteroclinic chain \( (\gamma_{\kappa_i})_{i \in \mathbb{Z}} \), and staying in a neighbourhood of the torus \( \Gamma_{k_i} \) for \( n_i \) iterations of \( F \) after shadowing \( \gamma_{\kappa_i} \).

Such results can be proved by Easton’s window method [46], but the AI limit proof seems simpler and more constructive. An alternative to the AI limit approach is given by variational methods which do not require transversality of the heteroclinic trajectories (see, for example, [58], [49], [11], [32], [33], [50]), but these methods are limited to positive-definite Lagrangian systems.

**Theorem 9** is analogous to the shadowing lemma in the general theory of hyperbolic sets of dynamical systems [52]. The difference is that the set \( \bigcup_{k \in I} \Gamma_k \bigcup_{\kappa \in J} \gamma_\kappa \) is not hyperbolic, so the standard results of hyperbolic theory do not apply. Theorem 9 can be used to construct diffusion orbits in the problem of Arnold diffusion for an \( \varepsilon \)-small perturbation of an a priori unstable integrable system (see §7.2). However, this is possible only in the absence of strong resonances, and moreover, the diffusion speed given by Theorem 9 will be of order \( O(\varepsilon^2) \), which is much weaker than the diffusion speed estimate \( O(\varepsilon/|\log \varepsilon|) \) obtained by using the AI limit approach for the separatrix map (see §7.2).

When the hyperbolic tori are hyperbolic fixed points, we obtain a weaker version of Theorem 8. In general the proof is similar. We will replace Lemma 3 with Lemma 4 below.

Let \( \Gamma \) be a hyperbolic torus. Choose symplectic coordinates \( (q, p) \) in a small tubular neighbourhood \( D \) of \( \Gamma \) in such a way that the Lagrangian local stable and unstable manifolds are given by generating functions:

\[
W_{\text{loc}}^+ = \{(q, p) \in D: q \in U, p = -\nabla S^+(q)\},
\]
\[
W_{\text{loc}}^- = \{(q, p) \in D: q \in U, p = \nabla S^-(q)\}.
\]

The projection of \( \Gamma = W_{\text{loc}}^+ \cap W_{\text{loc}}^- \) on \( U \) is the torus \( T = \{q: \nabla(S^+ + S^-)(q) = 0\} \).

Since \( U \) is a tubular neighbourhood of \( T \), we have a (non-canonical) projection \( U \to \mathbb{T}^d \). Let \( \tilde{U} \) be the universal covering of \( U \). Then we have a map \( \phi: \tilde{U} \to \mathbb{R}^d \). The map \( F \) lifts to a map \( \tilde{F}: \tilde{D} \to \tilde{D} \).

**Lemma 4.** Let \( r > 0 \). Then there exists an \( N > 0 \) such that:

\(^7\)Below, ‘heteroclinic’ always means heteroclinic or homoclinic.
(a) for all \( n \geq N \) and all \((q_+, q_-)\) in the set

\[
Y^n = \{(q_+, q_-) \in \widetilde{U} \times \widetilde{U} : |\phi(q_-) - \phi(q_+) - np| \leq r\}
\]

(5.8)

there exist unique \( z_+ = (q_+, p_+) \in \widetilde{D} \) and \( z_- = (q_- , p_- ) \in \widetilde{D} \) such that \( F^n(z_+) = z_- \) and \( F^j(z_+) \in \widetilde{D} \) for \( j = 0, 1, \ldots, n \);

(b) the map \( F^n : (q_+, p_+) \to (q_-, p_-) \) is given by the generating function \( Q^n \):

\[
F^n(q_+, p_+) = (q_-, p_-) \iff p_- = \partial_{q_-}Q^n(q_+, q_-), \ p_+ = \partial_{q_+}Q^n(q_+, q_-);
\]

(c) \( Q^n \) is a smooth function on \( Y^n \) and has the form

\[
Q^n(q_+, q_-) = S^+(q_+) + S^-(q_-) + n^{-1}v_n(q_+, q_-),
\]

where \( \|v_n\|_{C^2(Y^n)} \leq Cr^2 \), with a constant \( C > 0 \) independent of \( n \).

We do not need the rotation vector \( \rho \) to be Diophantine or non-resonant, but it is important that \( A \) is non-degenerate. Lemma 4 is proved in [15] for \( d = m - 1 \). For any \( 0 < d < m \) the proof is similar. For \( d = 0 \), Lemma 3 gives a sharper estimate.

For every torus \( \{ \Gamma_k \}_{k \in I} \) we define the tubular neighbourhoods \( U_k \) and \( D_k \) and generating functions \( S^\pm_k \) and \( Q^n_k \) as above. Lemma 4 provides the discrete Lagrangian describing the local map near the hyperbolic tori \( \Gamma_k \). Next we define the discrete Lagrangian describing the global map along a transversal heteroclinic trajectory \( \gamma_k \).

Take two tori \( \Gamma_j \) and \( \Gamma_k \) joined by a heteroclinic (homoclinic if \( k = j \)) trajectory \( \gamma_k \subset W^- \cap W^+ \). There exist points

\[
z^-_k = (q^-_k, p^-_k) \in \gamma_k \cap D_j \quad \text{and} \quad z^+_k = (q^+_k, p^+_k) \in \gamma_k \cap D_k
\]

and a number \( m_k \in \mathbb{N} \) such that \( F^{m_k}(z^-_k) = z^+_k \).

The symplectic coordinates can be chosen in such a way that the global map \( F^{m_k} \) from a neighbourhood of \( z^-_k \) to a neighbourhood of \( z^+_k \) is given by a generating function \( \Phi_k \) defined on \( X^-_k \times X^+_k \), where \( X^\pm_k \) is a small neighbourhood of \( q^\pm_k \):

\[
F^{m_k}(q_-, p_-) = (q_+, p_+) \iff p_+ = \partial_{q_+} \Phi_k(q_-, q_+), \ p_- = \partial_{q_-} \Phi_k(q_-, q_+).
\]

(5.10)

Since the heteroclinic trajectory \( \gamma_k \) is transversal, we conclude exactly as in §5.4 that \((q^-_k, q^+_k)\) is a non-degenerate critical point of the function

\[
R_k(q_-, q_+) = S^-_j(q_-) + \Phi_k(q_-, q_+) + S^+_k(q_+), \quad (q_-, q_+) \in X^-_k \times X^+_k.
\]

Let \( \pi : \widetilde{U}_k \to U_k \) be the universal covering with the transformation group

\[
\tau_v : \widetilde{U}_k \to \widetilde{U}_k, \quad v \in \mathbb{Z}^{d_k}, \quad d_k = \dim \Gamma_k.
\]

Fix connected components \( Z^\pm_k \) of the set \( \widetilde{X}^\pm = \pi^{-1}(X^\pm_k) \). Then \( \widetilde{X}^\pm = \bigcup_{v \in \mathbb{Z}^{d_k}} \tau_v Z^\pm_k \).

To reduce the non-uniqueness, we assume that the sets \( Z^\pm_k \) intersect the same fundamental domain \( K \) of the group action. Then

\[
\max\{d(x, y) : x \in Y^+_k, y \in Y^-_k\} \leq 2 \dim K.
\]

(5.11)
Let $\rho_k \in \mathbb{R}^{d_k}$ be the rotation vector of the torus $\Gamma_k$. Take $R > \sqrt{d_k}$. For any $n \in \mathbb{N}$, there exists a vector $v \in \mathbb{Z}^{d_k}$ such that

$$|v - n \rho_k| \leq R.$$  \hfill (5.12)

Then

$$|\phi(y) - \phi(x) - n \rho_k| \leq r = R + 2 \text{ diam } K$$

for all $x \in Z^+_k$ and $y \in \tau_v Z^-_k$. Let $N > 0$ be sufficiently large and let $n \geq N$. Then

$$Z^+_k \times \tau_v Z^-_k \subset Y^n_k,$$

where $Y^n_k$ is the set (5.8) corresponding to the torus $\Gamma_k$.

Let $\sigma = (\kappa, n, v)$, where $n \geq N$ and $v$ is chosen as in (5.12) above. We define the discrete Lagrangian $L_\sigma$ by

$$L_\sigma(x, y) = \Phi_\kappa(x) + Q^n_k(x_+, \tau_v y_-), \quad x = (x_-, x_+), \quad y = (y_-, y_+),$$

where $Q^n_k$ is the function (5.9) corresponding to the torus $\Gamma_k$. The Lagrangian $L_\sigma$ represents trajectories which shadow the heteroclinic trajectory $\gamma_\kappa$ and then stay in a neighbourhood of $\Gamma_k$ for $n$ iterations of $F$. As in §5.4, $L_\sigma$ has $2m$ degrees of freedom, but if desired, we can replace it by a Lagrangian with $m$ degrees of freedom.

To obtain an anti-integrable Lagrangian, we make a gauge transformation:

$$\hat{L}_\sigma(x, y) = L_\sigma(x, y) + S_j^-(x_-) - S_k^-(y_-).$$

By (5.9), $\hat{L}_\sigma = R_\kappa(x) + O(n^{-1})$, where $R_\kappa$ has a non-degenerate critical point $(q_\kappa^-, q_\kappa^+)$. Suppose we are given a chain $(\gamma_{\kappa_i})_{i \in \mathbb{Z}}$ of heteroclinic trajectories, where $\gamma_{\kappa_i}$ joins $\gamma_{\kappa_{i-1}}$ and $\gamma_{\kappa_i}$. Consider the infinite product

$$Z = \prod_{i \in \mathbb{Z}} Z^-_{\kappa_i} \times Z^+_{\kappa_i},$$

which is the set of sequences

$$x = (x_i), \quad x_i = (x_i^-, x_i^+) \in Z^-_{\kappa_i} \times Z^+_{\kappa_i}.$$  

For a sequence $(n_i, v_i)$ such that $n_i \geq N$ and $|v_i - n_i \rho_k| \leq R$, let $\sigma_i = (\kappa_i, n_i, v_i)$ and define a formal functional on $Z$ by

$$A_\sigma(x) = \sum_{i \in \mathbb{Z}} L_{\sigma_i}(x_i, x_{i+1}) = \sum_{i \in \mathbb{Z}} \hat{L}_{\sigma_i}(x_i, x_{i+1}).$$

Its critical points correspond to trajectories shadowing the chain $(\gamma_{\kappa_i})$. Since $\hat{L}_\sigma$ is anti-integrable, Theorem 9 follows from Theorem 5.

6. The AI limit in continuous Lagrangian systems

In the previous section we gave examples of applications of the AI limit in DLS. Now we discuss applications of the AI limit method to continuous Lagrangian systems, two examples for autonomous systems and one for time-dependent systems.
6.1. Turaev–Shilnikov theorem for Hamiltonian systems. As an example of the anti-integrable limit in autonomous Lagrangian or Hamiltonian systems, we prove the Turaev–Shilnikov theorem [72] for a Hamiltonian system with a hyperbolic equilibrium. This situation is more delicate than for hyperbolic fixed points of a symplectic map: without additional conditions, systems with transversal homoclinic trajectories may be integrable [40]. The proof we give follows the approach in [24].

Consider a Hamiltonian system with Hamiltonian $H$ on a symplectic manifold $P$. Let $\phi^t$ be the flow and $O$ a hyperbolic equilibrium point. We can assume that $H(O) = 0$. Let

$$\pm \lambda_j, \quad 0 < \text{Re} \lambda_1 \leq \cdots \leq \text{Re} \lambda_m,$$

be the eigenvalues. Suppose that the eigenvalue $\lambda_1$ with smallest real part is real and

$$0 < \lambda_1 < \text{Re} \lambda_2.$$

The case of complex $\lambda_1$ is somewhat simpler; it is studied in [40], [56], and [27].

Let $v_+$ be an eigenvector corresponding to $-\lambda_1$ and $v_-$ an eigenvector corresponding to $\lambda_1$. We fix a metric and assume that $|v_+| = 1$. To reduce the non-uniqueness of eigenvectors we can assume that $\omega(v_-, v_+) > 0$, where $\omega$ is the symplectic 2-form. Then the eigenvectors $v_\pm$ are uniquely defined up to the change $(v_+, v_-) \mapsto (-v_+, -v_-)$.

Let $W^\pm$ be the stable and unstable manifolds of the equilibrium $O$. They contain the strong stable and unstable manifolds $W^{++} \subset W^+$ and $W^{--} \subset W^-$ corresponding to the eigenvalues $\pm \lambda_j$ with $j > 1$. Any trajectory in $W^+ \setminus W^{++}$ is tangent to $v_+$ as $t \to +\infty$, and any trajectory in $W^- \setminus W^{--}$ is tangent to $v_-$ as $t \to -\infty$.

Let $\gamma(t) = \phi^t(z)$ be a homoclinic trajectory: $\lim_{t \to \pm \infty} \gamma(t) = O$. It is said to be transversal if $T_{\gamma(0)} W^+ \cap T_{\gamma(0)} W^- = \mathbb{R} \dot{\gamma}(0)$. Then the intersection of $W^+$ and $W^-$ along $\gamma$ is transversal on the energy level $H = 0$.

Suppose that there exist several transversal homoclinic trajectories $\{\gamma_\kappa\}_{\kappa \in J}$ which do not belong to the strong stable and unstable manifolds of $O$. Then $\gamma_\kappa$ will be tangent to $v_\pm$ as $t \to \pm \infty$, respectively:

$$\lim_{t \to \pm \infty} \frac{\dot{\gamma}_\kappa(t)}{|\dot{\gamma}_\kappa(t)|} = s_\kappa^\pm v_\pm, \quad s_\kappa^\pm = + \text{ or } -.$$

Let us define a graph $\Gamma_-$ as follows. The vertices $\kappa \in J$ correspond to transversal homoclinic trajectories $\gamma_\kappa$. We join the vertices $\kappa$ and $\kappa'$ by an edge if $s_\kappa^+ = s_{\kappa'}^-$. The graph $\Gamma_+$ is defined in the same way, but we join $\kappa$ and $\kappa'$ by an edge if $s_\kappa^+ = -s_{\kappa'}^-$. 

Theorem 10. There exists an $\varepsilon_0 > 0$ such that for any $\varepsilon \in (0, \varepsilon_0)$ and any path $k = (\kappa_i)_{i \in \mathbb{Z}}$ in the graph $\Gamma_+$ there exists a hyperbolic trajectory with energy $H = \varepsilon$ shadowing the homoclinic chain $(\gamma_{\kappa_i})$. Similarly, for any path $k = (\kappa_i)_{i \in \mathbb{Z}}$ in the graph $\Gamma_-$ the chain $(\gamma_{\kappa_i})$ is shadowed by a hyperbolic trajectory with energy $H = -\varepsilon$.

To prove this result we construct an anti-integrable DLS which describes the shadowing trajectories. Then it remains to use Theorem 5.

There exist symplectic coordinates $q, p$ in a neighbourhood $D$ of $O$ such that $O = (0, 0)$ and the local stable and unstable manifolds $W^\pm_{\text{loc}}$ are Lagrangian graphs.
over a small ball $U$ in $\mathbb{R}^m$:

$$\begin{align*}
W_{loc}^+ &= \{(q,p) : q \in U, p = -\nabla S_+(q)\}, \\
W_{loc}^- &= \{(q,p) : q \in U, p = \nabla S_-(q)\}. \\
\end{align*}$$

Let $S_\pm(0) = 0$. We can assume that the coordinates are chosen in such a way that the strong local stable and unstable manifolds are given by

$$\begin{align*}
W^+_{loc} &= \{(q,p) \in W^+_{loc} : q_1 = 0\}, \\
W^-_{loc} &= \{(q,p) \in W^-_{loc} : q_1 = 0\}.
\end{align*}$$

Fix a small $\delta > 0$ and for $s = +$ or $s = -$ let

$$U_s = \{q \in U : sq_1 > \delta|q|\} \quad \text{and} \quad W_s^\pm = \{(q,p) \in W_s^\pm_{loc} : q \in U_s\}.$$ 

We choose the signs in such a way that the vector $v_\pm$ points towards $W_{s_s}^\pm$. Then the homoclinic trajectory $\gamma_\kappa$ satisfies the condition $\gamma_\kappa(t) \in W_{s_\kappa}^\pm$ for $t \to \pm\infty$.

![Figure 5. Homoclinic trajectories $\gamma_\kappa$ and $\gamma_{\kappa'}$ with $s_\kappa^- = s_\kappa^+ = s_{\kappa'}^+ = +1$ and $s_{\kappa'}^- = -1$.](image)

Take any $\alpha \in (0,1)$.

**Lemma 5.** Let $\varepsilon_0 > 0$ be sufficiently small, let $\varepsilon \in (0,\varepsilon_0)$, and let $s = +$ or $s = -$. For any $q_\pm \in U_s$ there exist $p_\pm$ and $\tau > 0$ such that $H(q_\pm,p_\pm) = -\varepsilon$ and $\phi^\tau(q_+,p_+) = (q_-,p_-)$. The trajectory $\gamma(t) = \phi^t(q_+,p_+)$ with $0 \leq t \leq \tau$ stays in $D$. The Maupertuis action of the trajectory $\gamma$ has the form

$$S(q_+,q_-,\varepsilon) = \int_\gamma pdq = S_+(q_+) + S_-(q_-) + w(q_+,q_-,\varepsilon),$$

where $\|w\|_{C^2} \leq C\varepsilon^\alpha$. If $q_+ \in U_s$ and $q_- \in U_{-s}$, then the same is true, but the trajectory $\gamma$ will have energy $H = \varepsilon$.

Therefore, for small $\varepsilon > 0$ the local flow near $O$ is described by an anti-integrable discrete Lagrangian. Lemma 5 is proved in [24].

**Proof of Theorem 10.** For definiteness we will consider shadowing trajectories with negative energy $H = -\varepsilon$. 

Let $\Sigma = \partial U$. The homoclinic orbit $\gamma$ exits $W_{\text{loc}}^+$ at the point $(q_+^-, p_-^+)$ and enters $W_{\text{loc}}^+$ at the point $(q_+^+, p_-^+)$, where $q_\pm^\pm \in \Sigma$. Then there are moments of time $t_\pm > 0$ such that $\phi(t_\pm)(q_\pm^-, p_-^+) = (q_+^+, p_-^+)$. We can assume that $\det(\partial q_+^\pm / \partial p_-^\pm) \neq 0$. Then there are neighbourhoods $X^\pm_{\kappa^\pm}$ of the points $q_\pm^\pm$ in $\Sigma$ such that for any $(q_-, q_+) \in X^-_{\kappa^-} \times X^+_{\kappa^+}$ and any $\varepsilon \in (0, \varepsilon_0)$ there exist $p_\pm$ such that $H(q_\pm, p_\pm) = -\varepsilon$ and the trajectory $\phi(t)(q_-, p_-)$, $0 \leq t \leq \tau = t_+ + O(\varepsilon)$, with energy $-\varepsilon$ satisfies $\phi(t_\pm)(q_-^ \pm, p_-^ \pm) = (q_+, p_+^\pm)$. The map $(q_-, p_-) \mapsto (q_+, p_+)$ has a generating function $\Phi_\kappa(q_-, q_+, \varepsilon)$ such that $d\Phi_\kappa = p_+ dq_+ - p_- dq_-$. 

For $\varepsilon = 0$, (6.1) implies that the function 

$$R_\kappa(q_-, q_+) = S_-(q_-) + \Phi_\kappa(q_-, q_+, 0) + S_+(q_+), \quad (q_-, q_+) \in X^-_{\kappa^-} \times X^+_{\kappa^+},$$

has a critical point $(q^-_\kappa, q^+_\kappa)$ corresponding to the homoclinic trajectory $\gamma$. Since $\gamma$ is transversal, the critical point is non-degenerate.

For small $\varepsilon > 0$, trajectories with $H = -\varepsilon$ that shadow heteroclinic orbits $\{\gamma\}$ correspond to trajectories of a DLS $\{\Phi, S\}$ with Lagrangians $\Phi$ and $S$ defined on open sets in the sphere $\Sigma$. This system has $m - 1$ degrees of freedom. To apply Theorem 5 it is convenient to introduce a DLS with $2m - 2$ degrees of freedom.

We can assume that $X^\pm_{\kappa^\pm} \subset U_{s^\pm}$. Define the discrete Lagrangian $L_\kappa$ on $X^-_{\kappa^-} \times X^+_{\kappa^+} \times (U_{s^\pm} \cap \Sigma)$ by 

$$L_\kappa(x, y, \varepsilon) = \Phi_\kappa(x, \varepsilon) + S(x, y, \varepsilon), \quad x = (x_-, x_+), \quad y = (y_-, y_+).$$

This Lagrangian has $2m - 2$ degrees of freedom.

For a given path $k = (\kappa_i)$ in the graph $\Gamma_-$, we obtain the discrete action functional 

$$A_k(x) = \sum L_{\kappa_i}(x_i, x_{i+1}, \varepsilon), \quad x_i = (x_i^-, x_i^+).$$

As explained in §3.4, trajectories with $H = -\varepsilon$ that shadow the homoclinic chain $(\gamma_k)$ correspond to critical points of $A_k$, that is, trajectories of the DLS $\mathcal{L} = \{L_k\} \in \mathcal{J}$.

By a gauge transformation we can replace $L_\kappa$ with an anti-integrable Lagrangian 

$$\hat{L}_\kappa(x, y, \varepsilon) = L_\kappa(x, y) + S_-(x) - S_-(y) = R_\kappa(x) + O(\varepsilon^\alpha),$$

where the function $R_\kappa$ has a non-degenerate critical point $(q^-_\kappa, q^+_\kappa)$. Now Theorem 10 follows from Theorem 5. □

Theorem 10 was formulated in [72] (in a different form) and proved in [24] (for positive-definite Lagrangian systems) and in [71]. In [24] variational methods were used, so the transversality of the homoclinic trajectories was not assumed. Similar results hold for systems with several hyperbolic equilibria on the same energy level.

For example, consider a Lagrangian system on a compact manifold $M$ with $L(q, \dot{q}) = ||\dot{q}||^2/2 - V(q)$, where $V$ attains its maximum on the set $A = \{a_1, \ldots, a_n\}$ and each maximum point is non-degenerate. There exist many minimal heteroclinic trajectories joining the points in $A$, for example any pair of points can be joined by a chain of heteroclinic trajectories. Suppose that the minimal heteroclinic trajectories $\gamma_\kappa$ do not belong to the strong stable or unstable manifolds, and define the $s^\pm$
as above. Define the graph $\Gamma_-$ by joining $\kappa$ and $\kappa'$ by an edge if $s_{\kappa}^+ = s_{\kappa'}^-$. Then for small $\varepsilon > 0$ and any path in $\Gamma_-$ there exists a trajectory with energy $-\varepsilon$ that shadows the corresponding chain of heteroclinic trajectories. See [8] for interesting concrete examples.

6.2. $n$-centre problem with small masses. The problem we consider in this subsection is somewhat similar to the billiards problem discussed in §5.3: instead of small scatterers we have small singularities of the potential.

Let $M$ be a smooth manifold and let $A = \{a_1, \ldots, a_n\}$ be a finite set in $M$. Consider a Lagrangian system with the configuration space $M \setminus A$ and the Lagrangian

$$L(q, \dot{q}, \varepsilon) = L_0(q, \dot{q}) - \varepsilon V(q),$$

where $\varepsilon$ is a small positive parameter. We assume that $L_0$ is smooth on $TM$ and quadratic in the velocity:

$$L_0(q, \dot{q}) = \frac{1}{2} \|\dot{q}\|^2 + \langle w(q), \dot{q} \rangle - W(q),$$

where $\| \cdot \|$ is a Riemannian metric on $M$.

The potential $V$ is a smooth function on $M \setminus A$ with Newtonian singularities: in a small ball $U_k$ about $a_k$,

$$V(q) = -\frac{\phi_k(q)}{\text{dist}(q, a_k)},$$

where $\phi_k$ is a smooth positive function on $U_k$. The distance is defined by means of the Riemannian metric $\| \cdot \|$. We call the system with the Lagrangian (6.2) the $n$-centre problem. For $\varepsilon = 0$ the limit system with the Lagrangian (6.3) has no singularities.

Let

$$H(q, \dot{q}, \varepsilon) = H_0 + \varepsilon V, \quad H_0(q, \dot{q}) = \frac{1}{2} \|\dot{q}\|^2 + W(q),$$

be the energy integral. We fix $E$ such that the domain $D = \{W < E\}$ contains the set $A$ and study the system on the energy level $\{H = E\}$.

We call a trajectory $\gamma: [a, b] \to D$ of the limit system with the Lagrangian (6.3) a non-degenerate collision orbit if $\gamma(a), \gamma(b) \in A$, $\gamma(t) \notin A$ for $a < t < b$, and the endpoints $\gamma(a)$ and $\gamma(b)$ are non-conjugate along $\gamma$ on the energy level $\{H = E\}$, that is, for the Maupertuis action functional.

Suppose that there are several non-degenerate collision orbits $\{\gamma_{\kappa}\}_{\kappa \in J}$ joining points $a^-_{\kappa}, a^+_{\kappa} \in A$. We denote by $v^-_{\kappa}$ and $v^+_{\kappa}$ the initial and final velocity of $\gamma_{\kappa}$. Consider a graph $\Gamma$ with vertex set $J$ and with vertices $\kappa$ and $\kappa'$ joined by an edge if $a^+_{\kappa} = a^-_{\kappa'}$ and $v^+_{\kappa} \neq \pm v^-_{\kappa'}$.

The next result was proved in [19].

**Theorem 11.** There exists an $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0]$ and any path $\kappa = (\kappa_i)_{i \in \mathbb{Z}}$ in the graph $\Gamma$ there exists a unique (up to a time shift) hyperbolic trajectory of energy $E$ shadowing the chain $(\gamma_{\kappa_i})_{i \in \mathbb{Z}}$ of collision orbits.
Hence, there is an invariant subset in \( \{ H = E \} \) on which the system is conjugate to a suspension over a topological Markov chain. The topological entropy is positive if the graph \( \Gamma \) has a connected branched subgraph.

The conditions of Theorem 11 are stated in terms of the Lagrangian \( L_0 \) and the set \( A \) only, not involving the potential \( V \) if it has Newtonian singularities on \( A \). We can also add a smooth \( O(\varepsilon) \)-small perturbation to \( L \).

**Corollary 2.** Suppose that \( M \) is a closed manifold and 
\[
E > \min_{q \in M} \left( \frac{1}{2} \| w(q) \|^2 + W(q) \right).
\]
Then for any \( n \geq 2 \), almost all the points \( a_1, \ldots, a_n \in M \), and sufficiently small \( \varepsilon > 0 \) there exists a chaotic hyperbolic invariant set of trajectories of energy \( E \) that are close to chains of collision orbits.

Indeed, the assumption implies that the Jacobi metric 
\[
ds_E = \sqrt{2(E - W(q))} \| dq \| + (w(q), dq)
\]
is a positive-definite Finsler metric on \( M \), and by Morse theory any two generic points can be connected by an infinite number of non-degenerate geodesics, that is, trajectories of energy \( E \). Therefore, any two points in a generic finite set \( A \subset M \) can be connected by an infinite number of non-degenerate collision trajectories of energy \( E \), and no other points in \( A \) lie on these trajectories.

If \( n \) is large enough, then chaotic trajectories exist for purely topological reasons [21], so that the smallness of \( \varepsilon \) is not needed.

**Remark 9.** There is another corollary of Theorem 11 for systems containing no small parameter. Namely, let 
\[
L(q, \dot{q}) = \frac{1}{2} \| \dot{q} \|^2 - V(q),
\]
but with large energy \( E = \varepsilon^{-1} \). After the time change \( s = \varepsilon^{-1/2} t \) we obtain the Lagrangian (6.2) with \( L_0 = \| \dot{q} \|^2 / 2 \). Therefore, the conclusion of Theorem 11 holds. See [53] for the classical n-centre problem.

Next we prove Theorem 11. First consider the limit system with \( \varepsilon = 0 \). Let \( \Sigma_k = \partial U_k \). For any \( x \in \Sigma_k \) there exist a unique trajectory \( \gamma^+_x \) of energy \( E \) connecting \( x \) with \( a_k \) in \( U_k \) and a unique trajectory \( \gamma^-_x \) of energy \( E \) connecting \( a_k \) with \( x \). Let
\[
S^\pm_k(x) = \int_{\gamma^\pm_x} p \, dq = \int_{\gamma^\pm_x} ds_E.
\]
Then the \( S^\pm_k \) are smooth functions on \( \Sigma_k \).

Denote the velocity vectors of \( \gamma^\pm_x \) at the point \( a_k \) by \( u^+(x) \) and \( u^-(x) \). We fix an arbitrarily small \( \delta > 0 \) and let
\[
X_k = \{(x, y) \in \Sigma_k^2 : \| u^+(x) - u^-(y) \| \geq \delta \}.
\]
Lemma 6. Suppose that $\varepsilon_0 > 0$ is sufficiently small and let $\varepsilon \in (0, \varepsilon_0]$.  
1) For any $(x, y) \in X_k$ there exists a unique trajectory $\gamma = \gamma_{x,y}^\varepsilon$ of energy $E$ connecting $x$ with $y$ in $U_k$.  
2) $\gamma$ depends smoothly on $x$ and $y$.  
3) The Maupertuis action  

$$S_k(x, y, \varepsilon) = \int_{\gamma} p \, dq$$  

(6.8)  

is a smooth function on $X_k \times (0, \varepsilon_0]$ and, modulo a quantity of order $\varepsilon \log \varepsilon$,  

$$Q_k(x, y, \varepsilon) = S_k^-(x) + S_k^+(y) + \varepsilon u_k(x, y, \varepsilon),$$  

where $u_k$ is uniformly $C^2$-bounded on $X_k$ as $\varepsilon \to 0$.

The proof is based on regularization of singularities and on a modification of Lemma 3 required because all the positive eigenvalues for the regularized system will be equal (see [19]).

For any $\kappa \in J$, let $x_\kappa \in \Sigma^-_\kappa$ and $y_\kappa \in \Sigma^+_\kappa$ be the points of intersection of the collision orbit $\gamma_\kappa$ with $\Sigma^-_\kappa$ and $\Sigma^+_\kappa$, respectively. If the spheres $\Sigma_k$ are small enough, then the points $x_\kappa$ and $y_\kappa$ are non-conjugate along $\gamma_\kappa$.

Let $U^-_\kappa \subset \Sigma^-_\kappa$ be a small neighbourhood of $x_\kappa$ and $U^+_\kappa \subset \Sigma^+_\kappa$ a small neighbourhood of $y_\kappa$. Taking $\delta > 0$ small enough, we can assume that for any edge $(\kappa, \kappa')$ we have $U^-_\kappa \times U^-_{\kappa'} \subset X_k$, where $a_k = a_\kappa^+ = a_{\kappa'}^-$. If the neighbourhoods $U^\pm_\kappa$ are small enough and $\varepsilon \in (0, \varepsilon_0)$, then any points $x \in U^-_\kappa$ and $y \in U^+_\kappa$ are joined by a unique trajectory $\gamma_\kappa$ with $H = E$ which is close to $\gamma_\kappa$. Let  

$$\Phi_\kappa(x, y, \varepsilon) = \int_{\gamma_\kappa} p \, dq$$

be the Maupertuis action of $\gamma_\kappa$. Then $\Phi_\kappa$ is a smooth function on $U^-_\kappa \times U^+_\kappa$.

Lemma 7. The function $R_k(x, y) = \Phi_\kappa(x, y, 0) + S^-_\kappa(x) + S^+_\kappa(y)$ on $U^-_\kappa \times U^+_\kappa$ has a non-degenerate critical point $(x_\kappa, y_\kappa)$.

Lemma 7 follows from the assumption that $\gamma_\kappa$ is a non-degenerate critical point of the action functional. Indeed, $R_k(x, y)$ is the Maupertuis action of the piecewise smooth trajectory of the limit system obtained by gluing together the trajectories $\gamma^-_x$, $\beta_0$, and $\gamma^+_y$. Hence, $R_k$ is the restriction of the action functional to a finite-dimensional submanifold consisting of broken trajectories (with break points $x$ and $y$) connecting $a_{\kappa^-}$ with $a_{\kappa^+}$.

We define the discrete Lagrangian of a system with $2m - 2$ degrees of freedom by  

$$L_\kappa(z_-, z_+, \varepsilon) = \Phi_\kappa(z_-) + Q_\kappa^+(y_-, x_+, \varepsilon), \quad z_- = (x_-, y_-), \quad z_+ = (x_+, y_+).$$

For any path $k = (\kappa_i)$ in the graph $\Gamma$, a critical point of the functional  

$$A_k(z) = \sum L_{\kappa_i}(z_i, z_{i+1}, \varepsilon), \quad z_i \in U^-_{\kappa_i} \times U^+_{\kappa_i},$$

corresponds to a trajectory with $H = E$ that shadows the collision chain $(\gamma_{\kappa_i})$. 
As in §6.1, we replace $L_\kappa$ by a gauge-equivalent anti-integrable Lagrangian

$$\tilde{L}_\kappa(z_-, z_+, \varepsilon) = L(z_-, z_+, \varepsilon) + S^-_\kappa(x_-) - S^+\kappa(x_+) = R_\kappa(z_-) + O(\varepsilon).$$

By Lemma 7, $R_\kappa$ has a non-degenerate critical point. Now Theorem 11 follows from Theorem 5.

For a concrete example we consider the spatial circular restricted 3-body problem (Sun, Jupiter, and an asteroid) and suppose that the mass $\varepsilon$ of Jupiter is much smaller than the mass $1-\varepsilon$ of the Sun. The centre of mass is stationary and the first two masses move in circular orbits about it with angular velocity $1$.

Consider the motion of the asteroid in the frame $Oxyz$ rotating about the $z$-axis passing through the Sun at $O = (0,0,0)$. Jupiter can be chosen at $P = (1,0,0)$. The motion $q = (x,y,z)$ of the asteroid is described by a Lagrangian system of the form (6.2), where

$$L_0(q, \dot{q}) = \frac{1}{2}|\dot{q}|^2 + x\dot{y} - y\dot{x} + \frac{1}{2}|q|^2 + \frac{1}{|q|} \quad \text{and} \quad V(q) = \frac{1}{|q|} - \frac{1}{|q-P|} + x.$$

We have $M = \mathbb{R}^3 \setminus \{O\}$ and the singular set consists of one point $P$. The energy integral

$$H = \frac{1}{2}|\dot{q}|^2 - \frac{1}{2}|q|^2 - \frac{1-\varepsilon}{|q|} - \frac{\varepsilon}{|q-P|} + \varepsilon x$$

in the rotating coordinate frame is called the Jacobi integral, and $C = -2H$ is called the Jacobi constant.

For $\varepsilon = 0$, the limit system is the Kepler Sun-asteroid problem. Its bounded orbits are transformations of ellipses with parameters $a$, $e$, and $\iota$ to the rotating frame, where $a$ is the semimajor axis, $e$ is the eccentricity, and $\iota$ the inclination of the orbit to the plane of the orbit of the Sun and Jupiter. They have angular frequency $\Omega = a^{-3/2}$ and Jacobi constant $C = a^{-1} + 2\sqrt{a(1-e^2)} \cos \iota$. The collision orbits are rotating Kepler arcs starting and ending at $P$.

For $C \in (-2, +3)$ we define the set $A_C$ of allowed frequencies of Kepler ellipses to be

- $(0, 1)$ if $C \in [-1, +2],$
- $(0, (2+C)^{3/2})$ if $C \in (-2, -1),$
- $((3-C)^{3/2}, 1)$ if $C \in (2, 3).$

The next result was proved in [20]

**Theorem 12.** For any $C \in (-2, +3)$ there exists a dense subset $S \subset A_C$ such that for any $\Omega \in S$ there is a non-degenerate collision orbit $\gamma_\Omega$ with frequency $\Omega$ and inclination $\iota = \cos^{-1}(C/2 - \Omega^{2/3})$.

Then Theorem 11 implies the following.

**Corollary 3.** For any finite set $\Lambda \subset S$ there exists an $\varepsilon_0 > 0$ such that for any sequence $(\Omega_n)_{n \in \mathbb{Z}}$ in $\Lambda$ and any $\varepsilon \in (0, \varepsilon_0)$ there is a trajectory of the spatial circular restricted 3-body problem with Jacobi constant $C$ which $O(\varepsilon)$-shadows a concatenation of collision orbits formed from the Kepler arcs $\gamma_{\Omega_n}$. 
Trajectories of the 3-body problem which shadow chains of collision orbits of the Kepler problem were called second species solutions by Poincaré.

In [17] and [22] similar results were obtained for the elliptic restricted and non-restricted plane 3-body problem with 2 small masses. The proof is also based on a reduction to a DLS, but the Lagrangian is only partly anti-integrable. The corresponding generalizations of Theorem 5 were proved in [16] and [23].

6.3. Lagrangian systems with slow time dependence. A straightforward analogue of an anti-integrable DLS is a continuous Lagrangian system with slow time dependence:

\[ L = L(q, \varepsilon \dot{q}, t, \varepsilon), \quad q \in M, \]  

where \( \varepsilon > 0 \) is a small parameter. For example,

\[ L = \frac{\varepsilon^2}{2} \|\dot{q}\|^2 + U(q, t), \]  

where \( \| \cdot \| \) is a Riemannian metric on \( M \), possibly depending on \( t \). This looks similar to the anti-integrable discrete Lagrangian (5.2). Thus, we may expect that the limit as \( \varepsilon \to 0 \) is similar to the anti-integrable limit in the DLS.

Introducing the fast time \( s = t/\varepsilon \), we obtain a system depending slowly on the time \( s \):

\[ L = L(q, q', t, \varepsilon), \quad q' = \frac{dq}{ds}. \]  

If \( L \) satisfies the Legendre condition, then the Lagrangian system can be represented as a Hamiltonian system:

\[ q' = \partial_p H, \quad p' = -\partial_q H, \quad t' = \varepsilon, \]  

where

\[ p = \partial_{q'} L(q, q', t, \varepsilon), \quad H(q, p, t, \varepsilon) = \langle p, q' \rangle - L. \]  

Letting \( z = (q, p) \), we obtain a differential equation of the form

\[ z' = v(z, t, \varepsilon), \quad t' = \varepsilon. \]  

This is the standard form of a singularly perturbed differential equation [73].

Next we present a simplified version of some results in [18]. References to earlier classical results due to Cherry and Palmer can be found there. A similar anti-integrable limit approach was used in [61] for the Mather acceleration problem. Interesting results on chaotic energy growth for systems of the type (6.9) were obtained recently in [45] by a very different approach.

For \( \varepsilon = 0 \) the frozen Lagrangian (6.11) takes the form

\[ L_t(q, q') = L(q, q', t, 0), \]  

where the time \( t \) is now frozen. Hence, the energy \( H_t(q, p) = H(q, p, t, 0) \) is a first integral of the frozen system.

Suppose that for each \( t \) the frozen system has a hyperbolic equilibrium point \( O_t \), and there exists a homoclinic trajectory \( \gamma_t : \mathbb{R} \to M \) with \( \gamma_t(\pm \infty) = O_t \). Without loss of generality we can assume that \( H_t = 0 \) at the equilibrium \( O_t \).
Remark 10. For a natural system (6.10) on a compact manifold a homoclinic trajectory always exists if $O_t$ is a point of strict non-degenerate minimum of $U_t = U(q, t)$ (see [14]). If $M$ is not simply connected, then there are at least as many such trajectories as there are generators of the semigroup $\pi_1(M)$.

Let $I$ be the open set of $t$ such that the homoclinic trajectory $\gamma_t$ is transversal: the intersection of the stable and unstable manifolds $W^+_t$ along $\gamma_t$ is transversal on the energy level $\{H_t = 0\}$. The Maupertuis action

$$f(t) = J(\gamma_t) = \int_{\gamma_t} p \, dq$$

is a smooth function on $I$. Hamilton’s principle implies that

$$f'(t) = -\int_{-\infty}^{\infty} (\partial_t H_t)|_{\gamma_t(s)} \, ds,$$

where the homoclinic orbit $\gamma_t$ is parametrized by the fast time $s$. Therefore, $f$ is a version of the Poincaré–Melnikov function.

For simplicity assume that $L$ is 1-periodic in the time. Take a finite set of non-degenerate critical points of $f$ in $T = \mathbb{R}/\mathbb{Z}$. The corresponding points in $\mathbb{R}$ form a discrete periodic set $K \subset \mathbb{R}$. If $\rho > 0$ is small enough, then the intervals $I_k = (k - \rho, k + \rho)$, $k \in K$, are non-intersecting.

**Theorem 13.** Fix a small $\delta > 0$ and arbitrary constants $0 < c_1 < c_2$, and let $\varepsilon > 0$ be sufficiently small. Then for any increasing sequence $k_i \in K$ there exist a unique trajectory $q(t)$ and sequences $t_i \in I_{k_i}$ and $c_1 < T_i < c_2$ such that:

(a) $d(q(t), \gamma_{k_i}(\mathbb{R})) \leq \delta$ for $t \in [t_i, t_i + \varepsilon T_i]$;

(b) $d(q(t), O_t) \leq \delta$ for $t \in [t_i + \varepsilon T_i, t_i + 1]$;

and, moreover,

(c) $|t_i - k_i| \leq C \varepsilon$ and $d(q(t), \gamma_{k_i}(\mathbb{R})) \leq C \varepsilon$ for $t \in [t_i, t_i + \varepsilon T_i]$.

We call $q(t)$ a multibump trajectory (see Figure 6). Theorem 13 holds also for non-periodic $L$ if certain uniformity assumptions hold.

In [18] a similar result was proved without the transversality assumption. Consider the system (6.10) on a non-simply connected compact manifold, and let $f(t)$ be the minimum of the actions of non-contractible homoclinic trajectories for the frozen system. Then it is enough to assume that $f$ is non-constant.

**Proof of Theorem 13.** First let $\varepsilon = 0$. Since $O_t$ is a hyperbolic equilibrium of the limit system, it has local stable and unstable Lagrangian manifolds $W^+_t$ and $W^-_t$ in the phase space $T^*M$. We assume for simplicity that the projection $W^\pm_t \to M$ is non-degenerate at $O_t$. This holds if the limit system is natural as in (6.10); in general we can achieve this by changing symplectic coordinates near $O_t$. Hence,

$$W^+_t = \{(q, p) : q \in D_t, p = -\nabla S^+_t(q)\} \quad \text{and} \quad W^-_t = \{(q, p) : q \in D_t, p = \nabla S^-_t(q)\},$$

where $D_t$ is a small $\delta$-neighbourhood of $O_t$. Since $S^\pm_t$ is defined up to an additive function of the time, we can assume without loss of generality that $S^+_t(O_t) = 0$. Then $-S^+_t(q)$ is the action of the trajectory of the frozen system starting at $q$ and
Figure 6. Multibump trajectory shadowing the homoclinic $\gamma_k$. asymptotic to $O_t$ as $s \to +\infty$, while $S_t^-$ is the action of the trajectory asymptotic to $O_t$ as $s \to -\infty$ and ending at $q$.

In the next lemma the slow time $t$ is used.

**Lemma 8.** Fix $0 < c_1 < c_2$. For sufficiently small $\varepsilon > 0$, any $a < b$ with $b - a \in (c_1, c_2)$, and any $x \in D_a$ and $y \in D_b$, there exists a unique trajectory $q(t) \in D_t$ with $a \leq t \leq b$ such that $q(a) = x$ and $q(b) = y$. Moreover:

1) the function $q(t) = q(t, a, b, x, y, \varepsilon)$ is smooth for $\varepsilon > 0$;

2) the trajectory $q(t)$ is $C^0$-close to the concatenation of the asymptotic trajectories of the frozen system;

3) if

$$S(a, b, x, y, \varepsilon) = \varepsilon^{-1} \int_a^b L(q(t), \varepsilon \dot{q}(t), t, \varepsilon) \, dt$$

denotes the action of the trajectory $q(t)$ normalized to the fast time, then

$$S = S^+(x) + S^-(y) + \varepsilon v(a, b, x, y, \varepsilon), \quad (6.16)$$

where $v$ is a smooth function for $\varepsilon > 0$ and $\|v\|_{C^2} \leq C$ for some constant $C$ independent of $\varepsilon$.

We note that $q(t)$ behaves badly as $\varepsilon \to 0$: its time derivative is unbounded of order $\varepsilon^{-1}$, while the generating function $S$ is regular as $\varepsilon \to 0$. If the Lagrangian (6.9) is independent of $t$ and $\varepsilon$, then Lemma 8 follows from Shilnikov’s lemma [63] or from the strong $\lambda$-lemma [39].

In the general case, the existence of a solution of the given boundary value problem can be obtained using results in the theory of singularly perturbed differential equations [73].

Let $p(t) = \partial_q L(q(t), \varepsilon \dot{q}(t), t, \varepsilon)$ be the momentum (6.13) of the trajectory $q(t)$. By the first-variation formula,

$$\partial_x S = -p(a), \quad \partial_y S = p(b), \quad \partial_a S = H(x, p(a), a, \varepsilon), \quad \partial_b S = -H(y, p(b), b, \varepsilon). \quad (6.17)$$
Lemma 8 describes trajectories which stay near the equilibrium $O_t$ during a finite interval of the slow time $t$. Next we describe trajectories which travel near the homoclinic orbit $\gamma_k$ in a short interval of slow time of order $\varepsilon$. Hence, the fast time $s = \varepsilon^{-1} t$ will be used.

To simplify the notation, we can assume that $D_t = D_k$ is independent of $t$ for $t \in I_k$. Let $\Sigma_k = \partial D_k$. Let $\gamma_k(0)$ and $\gamma_k(\tau_k)$ be the points of intersection of the transversal homoclinic $\gamma_k : \mathbb{R} \to M$ with $\Sigma_k$. By changing $\delta$ if needed, we can assume that these points are non-conjugate along $\gamma_k$. Then for $z = (x, y, t, T)$ close to $z^0_k = (\gamma_k(0), \gamma_k(\tau_k), k, \tau_k)$ and for sufficiently small $\varepsilon > 0$ there exists for the system with Lagrangian $L(q, q', t + \varepsilon s, \varepsilon)$ a trajectory $\beta_\varepsilon(s)$ with $0 \leq s \leq T$ satisfying the boundary conditions $\beta_\varepsilon(0) = x$ and $\beta_\varepsilon(T) = y$ and staying close to the homoclinic orbit $\gamma_k$. Let

$$
\Phi_k(z, \varepsilon) = \int_0^T L(\beta_\varepsilon(s), \beta_\varepsilon'(s), t + \varepsilon s, \varepsilon) \, ds
$$

be its action. The derivatives of $\Phi_k$ satisfy equations analogous to (6.17):

$$
\begin{align*}
\partial_x \Phi_k &= -p(0), & \partial_y \Phi_k &= p(T), \\
\partial_t \Phi_k &= H(x, p(0), t, \varepsilon), & \partial_T \Phi_k &= -H(y, p(T), t + \varepsilon T, \varepsilon),
\end{align*}
$$

(6.18)

where $p(s)$ is the momentum of $\beta_\varepsilon(s)$.

**Lemma 9.** The function

$$
R_k(z) = S^+_k(x) + \Phi_k(x, y, t, T, 0) + S^-_k(y), \quad x, y \in \Sigma_k,
$$

has a non-degenerate critical point $z^0_k$.

Indeed, $R_k$ is the action of the concatenation of an asymptotic trajectory of the frozen system starting at $O_t$ and ending at $x$, a trajectory $\beta_0$ joining $x$ and $y$ in a neighbourhood of $\gamma_k$, and a trajectory starting at $y$ and asymptotic to $O_t$. The equation $\partial_T R_k = 0$ implies that $H_1 = 0$ along $q(s)$. Then the equations $\partial_x R_k = 0$ and $\partial_y R_k = 0$ imply that the concatenation is a smooth trajectory of the frozen system which is homoclinic to $O_t$. Thus, this trajectory coincides with $\gamma_t$. Finally, $\partial_t R_k = 0$ means that $t$ is a critical point of $f(t)$.

Next we define the DLS describing multibump trajectories. Let $J = \{\kappa = (\kappa_-, \kappa_+) \in K^2 : \kappa_- < \kappa_+\}$ be the vertex set of the graph, and let $\kappa, \kappa' \in J$ be joined with an edge if $\kappa_+ = \kappa'_-$. We define the discrete Lagrangian of the system with $2m$ degrees of freedom by

$$
L_k(z_-, z_+, \varepsilon) = \Phi_k(z_-, \varepsilon) + S(t_+ + \varepsilon T_-, t_-, y_-, x_+, \varepsilon), \quad z_\pm = (x_\pm, y_\pm, t_\pm, T_\pm).
$$

Thus, $L_k(z_-, z_+, \varepsilon)$ is the action of the broken trajectory which starts at $x_- \in \Sigma_{\kappa_-}$ at time $t_-$, travels near the homoclinic trajectory $\gamma_{\kappa_-}$ until it reaches $\Sigma_{\kappa_+}$ at the point $y_-$ at the time $t + \varepsilon T_-$, then stays close to the hyperbolic equilibrium $O_t$ for $t_- + \varepsilon T_- \leq t \leq t_+$, and ends at $x_+ \in \Sigma_{\kappa_+}$ at the time $t_+$.

By (6.17) and (6.18), critical points of the functional

$$
A_k(z) = \sum_{i} L_{\kappa_i}(z_i, z_{i+1}, \varepsilon), \quad \kappa_i = (k_i, k_{i+1}),
$$
Figure 7. The separatrix map.

where

\[ z_i = (x_i, y_i, t_i, T_i), \quad x_i, y_i \in \Sigma_{k_i}, \quad t_i \in I_{k_i}, \quad T_i > 0, \]

 correspond to trajectories shadowing the chain of homoclinic trajectories \((\gamma_{k_i})\).

We make a gauge transformation replacing \(L_{\kappa}\) by the Lagrangian

\[
\tilde{L}_{\kappa}(z_-, z_+, \varepsilon) = L_{\kappa}(z_-, z_+, \varepsilon) + S_{\kappa-}(x_-) - S_{\kappa+}(x_+) = R_{\kappa-}(z_-) + O(\varepsilon),
\]

where \(R_{\kappa-}\) has a non-degenerate critical point \(z^0_{\kappa-}\). Then \(\tilde{L}_{\kappa}\) has an anti-integrable form.

Note that the graph \(\Gamma\) describing the anti-integrable system is infinite. However, the Lagrangian is invariant with respect to the \(Z\)-action on \(\mathbb{R}\), so that the uniform anti-integrability required in Theorem 6 holds. \(\square\)

7. Separatrix map

7.1. The AI limit in the Zaslavsky separatrix map. We consider an integrable area-preserving map \(F_0\) having a hyperbolic fixed point with two homoclinic separatrix loops. Let \(F_{\varepsilon}\) be a perturbed map also assumed to be area-preserving. In an \(\varepsilon\)-neighbourhood of the unperturbed separatrix loops the dynamics of \(F_{\varepsilon}\) is determined by the separatrix map.

The construction is presented in Fig. 7. The left-hand side of the figure presents the phase space of the map \(F_{\varepsilon}\). We see the hyperbolic fixed point \(p_\varepsilon\) and its asymptotic curves (separatrices), which are split for \(\varepsilon \neq 0\). We also see two grey domains \(\Delta_{\varepsilon}^\pm\) on which the separatrix map will be defined. The boundaries of these domains are curvilinear quadrilaterals. Their ‘horizontal’ sides can be regarded as lying on invariant KAM curves (this is convenient, but not necessary) while the ‘vertical’ sides for each quadrangle are the images of each other under
the maps $F_\varepsilon$ and $F_\varepsilon^{-1}$. For any point $z \in \Delta_\varepsilon = \Delta_\varepsilon^+ \cup \Delta_\varepsilon^-$ its image $F_\varepsilon(z)$ lies outside $\Delta_\varepsilon$. By definition, the image of $z$ under the separatrix map is $F_\varepsilon^n(z)$, where $n = n(\varepsilon)$ is the minimal natural number such that $F_\varepsilon^n(z) \in \Delta_\varepsilon$. In Fig. 7 two such points $z$ are shown together with their images.

In some convenient coordinates the separatrix map can be computed in the form of an explicit part plus small error terms,

$$
\begin{pmatrix}
y \\
x \\
\sigma
\end{pmatrix}
\rightarrow
\begin{pmatrix}
y_+ \\
x_+ \\
\sigma_+
\end{pmatrix},
$$

$$
\begin{align*}
y_+ &= y + \lambda \frac{\partial V_\sigma}{\partial x} + O(\varepsilon), \\
x_+ &= x + \frac{1 + O(\varepsilon)}{\lambda}(\omega_\sigma + \log |y_+|), \\
\sigma_+ &= \sigma \text{ sgn } y_+,
\end{align*}
$$

(7.1)

Here the variable $x \mod 1$ changes along the separatrix, the variable $y$ changes across the separatrix, and the discrete variable $\sigma = \pm 1$ indicates which of the loops the orbit is near at a given moment of time. The functions $V_\sigma(x)$ (Poincaré–Melnikov potentials) are periodic with period $1$. (Details can be found, for example, in [62] and [70].)

The map (7.1) has the form

$$
y = \frac{\partial W}{\partial x}, \quad x_+ = \frac{\partial W}{\partial y_+}, \quad \sigma_+ = \sigma \text{ sgn } y_+,
$$

(7.2)

with generating function

$$W = W(y_+, x, \sigma) = xy_+ - \lambda V_\sigma(x) + \frac{1 + O(\varepsilon)}{\lambda}(\omega_\sigma + \log |y_+| - 1)y_+.
$$

The first two equalities in (7.2) can be written as

$$x_+ dy_+ + y dx = dW(y_+, x, \sigma).
$$

To represent (7.1) in Lagrangian form, consider another generating function (the Legendre transform of $W$)

$$\lambda L = x_+ y_+ - W, \quad y_+ dx_+ - y dx = \lambda dL.
$$

(7.3)

Here we use the fact that a Lagrangian is defined up to a non-zero constant multiplier. It is easy to obtain an explicit formula for $L$:

$$L(x, x_+, \sigma, \theta_+) = (1 + O(\varepsilon))\theta_+ e^{\lambda(x_+-x-\tilde{\omega}_\sigma)} + V_\sigma(x),$$

$$\theta_+ = \text{ sgn } y_+, \quad \tilde{\omega}_\sigma = \omega_\sigma + \lambda^{-1} \log \lambda^2.
$$

(7.4)

Thus, the separatrix map is a $\mathbb{Z}$-equivariant DLS, where $\mathbb{Z}$ acts on $M = \mathbb{R} \times \mathbb{Z}_2 \times \mathbb{Z}_2$ by the shifts

$$M \ni (x, \sigma, \theta) \mapsto k(x, \sigma, \theta) = (x + k, \sigma, \theta).
$$

The Lagrangian form of the separatrix map is

$$
\begin{pmatrix}
x_-
\\
x
\\
\sigma_-
\\
\theta
\end{pmatrix}
\rightarrow
\begin{pmatrix}
x
\\
x_+
\\
\sigma
\\
\theta_+
\end{pmatrix},

\sigma = \sigma_- \theta, \quad \theta_+ = \sigma \sigma_+,

\frac{\partial}{\partial x}(L_- + L) = 0,
$$

(7.5)
where
\[ L_- = L(x, x, \sigma, \vartheta) \quad \text{and} \quad L = L(x, x, \sigma, \vartheta). \]
The first two equalities in (7.5) arise from the definition of \( \vartheta \) (\( \vartheta = \text{sgn}\, I \)) and from the last equality in (7.1). The third equality in (7.5) follows from (7.3) since, according to (7.3) and the analogous equality \( y \, dx - y_- \, dx_- = dL_- \), we have
\[ y = -\frac{\partial L}{\partial x} = \frac{\partial L_-}{\partial x}. \]

It is easy to check that the quantities \( x_+, \sigma, \) and \( \vartheta_+ \) are computed uniquely from (7.5) in terms of \( x_-, x, \sigma_- \), and \( \vartheta \).

We obtain a DLS with extra discrete variables \( \sigma \) and \( \vartheta \). Any sequence
\[ z = \{z_j\}, \quad z_j = \begin{pmatrix} x_j \\ \sigma_j \\ \vartheta_j \end{pmatrix}, \quad \sigma_{j+1} = \sigma_j \vartheta_{j+1}, \]
is called a path. Let \( \Sigma \) be the set of all paths.

In general the index \( j \) takes all integer values. However, it is possible to consider also semi-infinite and finite paths. Paths finite from the left begin with a triple \( z_j \), where \( x_j = +\infty \). Paths finite from the right end with \( z_j \), where \( x_j = -\infty \). Paths finite from the left and from the right are said to be finite.

The action \( A \) is defined as the formal sum
\[ A = A(z) = \sum_j L(x_j, x_{j+1}, \sigma_j, \vartheta_{j+1}). \]
The path \( z^0 \) is called an extremal (or a trajectory) if \( \partial A/\partial x_j|_{z=z^0} = 0 \) for any \( j \).

We note that semifinite trajectories belong to separatrices. Finite ones belong to both stable and unstable separatrices, and therefore they are homoclinic trajectories.

We define the distance \( \rho \) on \( \Sigma \) as follows. Let \( z' \) and \( z'' \) be paths, where
\[ z'_j = \begin{pmatrix} x'_j \\ \sigma'_j \\ \vartheta'_j \end{pmatrix} \quad \text{and} \quad z''_j = \begin{pmatrix} x''_j \\ \sigma''_j \\ \vartheta''_j \end{pmatrix}. \]
We put \( \rho(z', z'') = \infty \) if the sequences \( \sigma'_j \) and \( \vartheta'_j \) do not coincide with \( \sigma''_j \) and \( \vartheta''_j \) or if for some \( j \) only one of the triples is defined. Otherwise we put
\[ \rho(z', z'') = \sup_j |x'_j - x''_j|. \]
Here we take \(|-\infty - (\infty)| = |+\infty - (+\infty)| = 0\).

Let \( \text{Cr}(\sigma) \) denote a finite set of non-degenerate critical points of the function \( V_\sigma \).

The set \( \Sigma \) contains the subset \( \Pi \) of simple paths (codes). By definition, a path \( z \) is simple if \( x_j \in \text{Cr}(\sigma_j) \) for any \( j \).

**Theorem 14.** Suppose that the constants \( c_1 \) and \( c_2 = c_2(c_1) \) are sufficiently large. Then, for any simple path \( z^* \) such that \( x^*_j - x^*_{j+1} > c_1 \) for all \( j \) there exists a unique trajectory \( z \) in the \( c_2^{-1} \)-neighbourhood of \( z^* \), and \( z \) is hyperbolic.
Theorem 14 establishes a symbolic dynamics in a neighbourhood of separatrices of an area-preserving map. It can be deduced from Theorem 5 by introducing the DLS with Lagrangian

\[ L_\kappa(x_-, x_+) = L(x_- + k, x_+, \sigma, \theta), \quad \kappa = (k, \sigma, \theta), \quad x_\pm \in (0, 2\pi), \]

where \( k > c_1 \) is sufficiently large. The corresponding graph has vertices \( \kappa \), and two vertices \( \kappa \) and \( \kappa' \) are joined by an edge if \( \sigma' = \sigma \theta \). There is an edge \( \gamma \) for every non-degenerate critical point of \( V_\sigma \) in \((0, 2\pi)\). For large \( k \) this DLS is anti-integrable. The graph \( \Gamma \) is infinite, but the uniform anti-integrability condition \( U \) obviously holds.

The traditional approach to symbolic dynamics near separatrices was presented in [13], [2], and [60]. Another version of the separatrix map was constructed by Shilnikov and Afraimovich ([1], and also [64]). This construction is also given in §6.1.

7.2. Separatrix map and Arnold diffusion. Ideas of the AI limit can be applied in the problem of Arnold diffusion. Here we discuss only the \textit{a priori} unstable case where, unlike the original (\textit{a priori} stable) situation [3], the unperturbed integrable system contains a normally hyperbolic manifold \( N \). In the perturbed system chaos is mainly concentrated near \( N \) and its asymptotic manifolds.

We consider a non-autonomous near-integrable Hamiltonian system on the phase space \( T^n_x \times \mathcal{D} \times D \times \mathbb{T}_t \), where \( \mathcal{D} \subset \mathbb{R}^n_y \), is an open domain with compact closure \( \mathcal{D} \) and \( D \subset \mathbb{R}^2_{(v,u)} \) is an open domain. The Hamiltonian function and the symplectic structure are as follows:

\[ H(y, x, v, u, t, \varepsilon) = H_0(y, v, u) + \varepsilon H_1(y, x, v, u, t) + \varepsilon^2 H_2(y, x, v, u, t, \varepsilon), \]
\[ \omega = dy \wedge dx + dv \wedge du. \]

As usual, \( \varepsilon \geq 0 \) is a small parameter. The Hamiltonian equations have the form

\[ \dot{y} = -\frac{\partial H}{\partial x}, \quad \dot{x} = \frac{\partial H}{\partial y}, \quad \dot{v} = -\frac{\partial H}{\partial u}, \quad \dot{u} = \frac{\partial H}{\partial v}. \tag{7.6} \]

Assume that in the unperturbed Hamiltonian the variables \( y \) are separated from \( u \) and \( v \), that is, \( H_0(y, v, u) = F(y, f(v, u)) \). The function \( f \) has a non-degenerate saddle point \( (v, u) = (0, 0) \) that is a unique critical point on a compact connected component of the set

\[ \gamma = \{(v, u) \in D: f(v, u) = f(0, 0)\}. \]

In dynamical terminology \((0, 0)\) is a hyperbolic equilibrium of the Hamiltonian system \((D, dv \wedge du, f)\) with one degree of freedom, and the corresponding separatrices \( \gamma \) are doubled. Topologically, these separatrices form a figure-eight: two loops, \( \gamma^\pm \), issuing from a single point, \( \gamma = \gamma^+ \cup \gamma^- \).

Hence, the unperturbed normally hyperbolic manifold is

\[ N = T^n_x \times \mathcal{D} \times (0, 0) \times \mathbb{T}_t. \]
It is foliated by the tori
\[ N_y = \{ x, y, u, v, t \}: u = v = 0, \ y = \text{const} \}, \]
which carry quasi-periodic dynamics with frequencies
\[ \left( \nu(y) \right) \frac{1}{1}, \ \nu(y) = \frac{\partial H_0}{\partial y} (y, 0, 0). \]

We are interested in the perturbed dynamics near manifolds asymptotic to \( N \).

The problem of Arnold diffusion in an \textit{a priori} unstable case has three aspects, contained in the following conjecture.

\textbf{Conjecture 1} ([5], [69]).

A \textit{(genericity)}. Diffusion exists for an open dense set of \( C^r \)-perturbations, where \( r \in \mathbb{N} \cup \{ \infty, \omega \} \) is sufficiently large.

B \textit{(freedom)}. The projection of a diffusion trajectory on the \( y \)-space can move in a small neighbourhood of any smooth curve \( \chi \subset \mathcal{D} \).

C \textit{(velocity)}. There are ‘fast’ diffusion trajectories whose average velocity along \( \chi \) is of order \( \varepsilon / |\log \varepsilon| \).

There are several different approaches to the problem. The traditional approach is based on the construction of transition chains of hyperbolic tori ([3], [34], [35], [43], [36]–[38]), and later it was effectively supplemented with the idea of the scattering map ([36]–[38]) and symbolic dynamics in polysystems [26]. A variational approach was developed in [9]–[11], [32], [33], [46], [49], [50].

Conjecture 1 has been proved only in the case \( n = 1 \) (two-and-a-half degrees of freedom) [68]. Here we explain briefly the ideas and methods in [68]. The system (7.6)\( |_{\varepsilon = 0} \) has the \( n \)-parametric family of (partially) hyperbolic \((n+1)\)-tori \( N_y \) which foliate the normally hyperbolic manifold \( N \). The manifolds asymptotic to \( N \) have two components \( \{ y \} \times \mathbb{T}_x^n \times \gamma^\pm \times \mathbb{T}_t \). Therefore, after a perturbation the situation reminds us of the one discussed in §7.1. In the phase space of the time-1 map a picture analogous to that in Fig. 7 appears. However, unlike the case considered in §7.1, we now have to deal with a family of hyperbolic tori and their asymptotic manifolds. The separatrix map can still be defined [66], and explicit formulae for it can be regarded as a multidimensional generalization of (7.1).

The separatrix map can be presented in the form
\[
(\zeta, \rho, \tau, \sigma, \theta) \mapsto (\zeta_+, \rho_+, \tau, \sigma_+, \theta_+),
\]
\[
\rho = \frac{\partial \mathcal{R}}{\partial \zeta}, \ \ z_+ = \frac{\partial \mathcal{R}}{\partial \rho_+}, \ \ \frac{\partial}{\partial \tau} (\mathcal{R}_- + \mathcal{R}) = 0, \ \ \sigma = \sigma_\tau, \ \ \vartheta_+ = \sigma \sigma_+, \ \ \mathcal{R}_+ = R(\zeta, \rho, \tau, \sigma, \theta, t, \varepsilon), \ \ \mathcal{R}_- = R(\zeta, \rho, \tau, \sigma, \theta, t, \varepsilon)
\]
(cf. (7.5)). Here up to small error terms the equality \( \varepsilon \rho = y \) holds, and \( \zeta - x \) is a function of \( y, u, \) and \( v \). The integer variable \( t \) has the meaning of the time during which the trajectory of the time-1 map travels outside the analogues of the domains \( \Delta_\varepsilon^\pm \) (see Fig. 7). The variables \( \sigma \) and \( \vartheta \) are analogous to the corresponding ones in
§ 7.1, and
\[
\mathcal{R} = (\rho_+ + \zeta + \nu t_+) - (\tau_+ - \tau - t_+) H(\varepsilon \rho_+, \zeta)
\]
\[
- \vartheta_+ e^{\lambda(\tau_+ - \tau - \omega_+^\sigma)} + \tilde{\Theta}^\sigma(\varepsilon \rho_+, \zeta, \tau),
\]
\[
\omega_+^\sigma = t_+ + \lambda^{-1} \log \varepsilon + f(\varepsilon \rho_+), \quad \nu = \nu(\varepsilon \rho_+), \quad \lambda = \lambda(\varepsilon \rho_+).
\]

If \( n = 0 \), then we do not have the variables \( \rho \) and \( \zeta \), and only the last two terms in \( \mathcal{R} \) remain. In this case \( \mathcal{R} \) becomes \( L \) up to a constant multiplier (see (7.4)).

We are interested in the case when the quantities \( t_+ + \lambda^{-1} \log \varepsilon \) are greater than a large positive constant \( K_0 \), so that \( e^{K_0} \) plays the role of a large parameter in the AI limit. The function \( H \) is essential only in small neighbourhoods of strong (low-order) resonances
\[
\{ \rho: \langle \nu(\rho), k \rangle + k_0 = 0 \}, \quad k \in \mathbb{Z}^n, \quad k_0 \in \mathbb{Z}, \quad |k| + |k_0| < C.
\]

We see that the discrete dynamical system (7.7) is partially Hamiltonian (with respect to \( \rho \) and \( \zeta \)) and partially Lagrangian (with respect to \( \tau \)). Because of the presence of the new Hamiltonian variables, we have to use a certain generalization of the AI limit method. Unfortunately, in this generalization the symbolic dynamics is not as transparent and standard as in § 4.1 or § 7.1. The construction is as follows, [67]. Having a finite piece of a trajectory of the separatrix map and the corresponding quasi-trajectory, a piece of the same length (the code), we present a rule according to which the code can be extended by adding a new point. Then according to the main result in [67] the trajectory can be slightly deformed and extended while staying close to the longer code. So we have again a pair: a piece of the trajectory with a code. By using a certain freedom in the rule for extending a code, one can hope to push the trajectory in the desired direction in the \( y \)-space.

Here another difficulty appears. This possibility of pushing the orbit in the proper direction is relatively simple in the non-resonant zone, that is, where the frequency vector \( (\nu(y), 1) \in \mathbb{R}^{n+1} \) (1 is the time frequency) does not allow low-order resonances. In a near-resonant domain the construction of a reasonable extension of a code is a separate delicate problem, which has been solved completely only in the case of two-and-a-half degrees of freedom [68]. Diffusion in domains free of low-order resonances in the case of arbitrary dimension was established in [69].

Finally, we mention another application of the separatrix map. According to [29], [48], [51], in (so far, special cases of) \textit{a priori} unstable systems with two-and-a-half degrees of freedom a large set of trajectories has been constructed whose projections on the \( y \)-axis for small \( \varepsilon > 0 \) behave like trajectories of Brownian motion. This shows that the term ‘diffusion’ proposed by Chirikov is quite adequate for the phenomenon discussed above.

Bibliography

[1] В. С. Афраймович, Л. П. Шильников, “О малых периодических возмущениях автономных систем”, \textit{Докл. АН СССР} 214:4 (1974), 739–742; English transl., V. S. Afraimovich and L. P. Shil’nikov, “On small periodic perturbations of autonomous systems”, \textit{Soviet Math. Dokl.} 15 (1974), 206–211.
В. М. Алексеев, “Квазислучайные динамические системы. I. Квазислучайные
dиффеоморфизмы”, Матем. сб. 76(118):1 (1968), 72–134; “II. Одномерные
нелинейные колебания в периодически возмущаемом поле” 77(119):4 (1968),
545–601; “III. Квазислучайные колебания одномерных осцилляторов”
78(120):1 (1969), 3–50; English transl., V. M. Alekseev, “Quasirandom
dynamical systems. I. Quasirandom diffeomorphisms”, Math. USSR-Sb. 5:1
(1968), 73–128; “II. One-dimensional nonlinear oscillations in a field with periodic
perturbation”, Math. USSR-Sb. 6:4 (1968), 505–560; “III. Quasirandom oscillations
of one-dimensional oscillators”, Math. USSR-Sb. 7:1 (1969), 1–43.

В. И. Арнольд, “О неустойчивости динамических систем со многими степенями
свободы”, Докл. АН СССР 156 (1964), 9–12; English transl., V. I. Arnol’d,
“Instability of dynamical systems with several degrees of freedom”, Soviet Math.
Dokl. 5 (1964), 581–585.

В. И. Арнольд, Математические методы классической механики, Наука,
M. 1974, 431 с.; English transl., V. I. Arnold, Mathematical methods of classical
mechanics, Grad. Texts in Math., vol. 60, Springer-Verlag, New York 1978,
xvi+462 pp.

В. И. Арнольд, В. В. Козлов, А. И. Нейштадт, Динамические системы – 3,
Итоги науки и техн. Сер. Соврем. пробл. матем. Фундам. направления, 3,
ВИНИТИ, М. 1985, с. 5–290; English transl., V. I. Arnol’d, V. V. Kozlov,
and A. I. Neishtadt, “Mathematical aspects of classical and celestial mechanics”,
Dynamical systems, III, Encyclopaedia Math. Sci., vol. 3, Springer-Verlag, Berlin
1988, pp. 1–291.

S. Aubry and G. Abramovici, “Chaotic trajectories in the standard map. The
concept of anti-integrability”, Phys. D 43:2-3 (1990), 199–219.

S. Aubry, R. S. MacKay, and C. Baesens, “Equivalence of uniform hyperbolicity for
symplectic twist maps and phonon gap for Frenkel–Kontorova models”, Phys. D
56:2-3 (1992), 123–134.

C. Baesens, Y.-C. Chen, and R. S. MacKay, “Abrupt bifurcations in chaotic
scattering: view from the anti-integrable limit”, Nonlinearity 26:9 (2013),
2703–2730.

M. Berti and P. Bolle, “A functional analysis approach to Arnold diffusion”, Ann.
Inst. H. Poincaré Anal. Non Linéaire 19:4 (2002), 395–450.

M. Berti, L. Biasco, and P. Bolle, “Drift in phase space: a new variational
mechanism with optimal diffusion time”, J. Math. Pures Appl. (9) 82:6 (2003),
613–664.

U. Bessi, “An approach to Arnold’s diffusion through the calculus of variations”,
Nonlinear Anal. 26:6 (1996), 1115–1135.

M. Bialy, “Maximizing orbits for higher-dimensional convex billiards”, J. Mod. Dyn.
3:1 (2009), 51–59.

G. D. Birkhoff, Dynamical systems, Amer. Math. Soc. Colloq. Publ., vol. 9, Amer.
Math. Soc., New York 1927, viii+295 pp.

С. В. Болотин, “Либрационные движения натуральных динамических систем”,
Вестн. Моск. ун-та. Сер. 1 Матем. Мех., 1978, № 6, 72–77; English transl.,
S. V. Bolotin, “Librations of natural dynamic systems”, Mosc. Univ. Mech. Bull.
33:5-6 (1978), 49–53.

S. V. Bolotin, “Symbolic dynamics near minimal hyperbolic invariant tori of
Lagrangian systems”, Nonlinearity 14:5 (2001), 1123–1140.

S. V. Bolotin, “Shadowing chains of collision orbits”, Discrete Contin. Dyn. Syst.
14:2 (2006), 235–260.
The anti-integrable limit

[17] S. Bolotin, “Symbolic dynamics of almost collision orbits and skew products of symplectic maps”, *Nonlinearity* 19:9 (2006), 2041–2063.

[18] S. Bolotin and R. MacKay, “Multibump orbits near the anti-integrable limit for Lagrangian systems”, *Nonlinearity* 10:5 (1997), 1015–1029.

[19] S.V. Bolotin and R.S. MacKay, “Periodic and chaotic trajectories of the second species for the n-centre problem”, *Celestial Mech. Dynam. Astronom.* 77:1 (2000), 49–75.

[20] S. Bolotin and R.S. MacKay, “Nonplanar second species periodic and chaotic trajectories for the circular restricted three-body problem”, *Celestial Mech. Dynam. Astronom.* 94:4 (2006), 433–449.

[21] S.V. Bolotin and P. Negrini, “Global regularization for the n-center problem on a manifold”, *Discrete Contin. Dyn. Syst.* 33:3 (2013), 1009–1032.

[22] S. Bolotin and P. Negrini, “Shilnikov lemma for a nondegenerate critical manifold of a Hamiltonian system”, *Regul. Chaotic Dyn.* 18:6 (2013), 774–800.

[23] S.V. Bolotin and P.H. Rabinowitz, “A variational construction of chaotic trajectories for a reversible Hamiltonian system”, *J. Differential Equations* 148:2 (1998), 364–387.

[24] S.V. Bolotin and D.V. Treschev, “Remarks on the definition of hyperbolic tori of Hamiltonian systems”, *Regul. Chaotic Dyn.* 5:4 (2000), 401–412.

[25] A. Bounemoura and E. Pennamen, “Instability for a priori unstable Hamiltonian systems: a dynamical approach”, *Discrete Contin. Dyn. Syst.* 32:3 (2012), 753–793.

[26] B. Buffoni and E. Séré, “A global condition for quasi-random behavior in a class of conservative systems”, *Comm. Pure Appl. Math.* 49:3 (1996), 285–305.

[27] L. A. Bunimovich, Ya. G. Sinai, N. I. Chernov, “Статистические свойства двумерных гиперболических биллиардов”, *УМН* 46:4(280) (1991), 43–92; English transl., L. A. Bunimovich, Ya. G. Sinai, and N. I. Chernov, “Statistical properties of two-dimensional hyperbolic billiards”, *Russian Math. Surveys* 46:4 (1991), 47–106.

[28] O. Castejón and V. Kaloshin, *Random iteration of maps on a cylinder and diffusive behavior*, 2015, 75 pp. http://www.math.umd.edu/~vkaloshi/papers/random-iteration.pdf

[29] Y.-C. Chen, “Anti-integrability in scattering billiards”, *Dyn. Syst.* 19:2 (2004), 145–159.

[30] Y.-C. Chen, “On topological entropy of billiard tables with small inner scatterers”, *Adv. Math.* 224:2 (2010), 432–460.

[31] C.-Q. Cheng and J. Yan, “Existence of diffusion orbits in a priori unstable Hamiltonian systems”, *J. Differential Geom.* 67:3 (2004), 457–517.

[32] C.-Q. Cheng and J. Yan, “Arnold diffusion in Hamiltonian systems: a priori unstable case”, *J. Differential Geom.* 82:2 (2009), 229–277.

[33] L. Chierchia and G. Gallavotti, “Drift and diffusion in phase space”, *Ann. Inst. H. Poincaré Phys. Théor.* 60:1 (1994), 1–144.

[34] A. Delshams, R. de la Llave, and T.M. Seara, *A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model*, Mem. Amer. Math. Soc., vol. 179, № 844, Amer. Math. Soc., Providence, RI 2006, viii+141 pp.

[35] A. Delshams, R. de la Llave, and T.M. Seara, “Geometric properties of the scattering map of a normally hyperbolic invariant manifold”, *Adv. Math.* 217:3 (2008), 1096–1153.
[37] A. Delshams and G. Huguet, “Geography of resonances and Arnold diffusion in a priori unstable Hamiltonian systems”, Nonlinearity 22:8 (2009), 1997–2077.

[38] A. Delshams and G. Huguet, “A geometric mechanism of diffusion: rigorous verification in a priori unstable Hamiltonian systems”, J. Differential Equations 250:5 (2011), 2601–2623.

[39] Bo Deng, “The Šil’nikov problem, exponential expansion, strong λ-lemma, C¹-linearization, and homoclinic bifurcation”, J. Differential Equations 79:2 (1989), 189–231.

[40] R. L. Devaney, “Homoclinic orbits in Hamiltonian systems”, J. Differential Equations 21:2 (1976), 431–438.

[41] P. Duarte, “Plenty of elliptic islands for the standard family of area preserving maps”, Ann. Inst. H. Poincaré Anal. Non Linéaire 11:4 (1994), 359–409.

[42] R. W. Easton, J. D. Meiss, and G. Roberts, “Drift by coupling to an anti-integrable limit”, Phys. D 156:3-4 (2001), 201–218.

[43] E. Fontich and P. Martín, “Arnold diffusion in perturbations of analytic integrable Hamiltonian systems”, Discrete Contin. Dyn. Syst. 7:1 (2001), 61–84.

[44] G. Gallavotti, G. Gentile, and V. Mastroianni, “Hamilton–Jacobi equation, heteroclinic chains and Arnol’d diffusion in three time scale systems”, Nonlinearity 13:2 (2000), 323–340.

[45] V. Gelfreich and D. Turaev, “Unbounded energy growth in Hamiltonian systems with a slowly varying parameter”, Comm. Math. Phys. 283:3 (2008), 769–794.

[46] M. Gidea and C. Robinson, “Diffusion along transition chains of invariant tori and Aubry–Mather sets”, Ergodic Theory Dynam. Systems 33:5 (2013), 1401–1449.

[47] C. Golé, Symplectic twist maps. Global variational techniques, Adv. Ser. Nonlinear Dynam., vol. 18, World Scientific Publishing Co., Inc., River Edge, NJ 2001, xviii+305 pp.

[48] M. Guardia, V. Kaloshin, and J. Zhang, A second order expansion of the separatrix map for trigonometric perturbations of a priori unstable systems, 2015, 50 pp., arXiv:1503.08301v2.

[49] V. Kaloshin and M. Levi, “Geometry of Arnold diffusion”, SIAM Rev. 50:4 (2008), 702–720.

[50] V. Kaloshin and Ke Zhang, Dynamics of the dominant Hamiltonian, with applications to Arnold diffusion, 2014 (v2 – 2015), 84 pp., arXiv:1410.1844.

[51] V. Kaloshin, Jianlu Zhang, and Ke Zhang, Normally hyperbolic invariant laminations and diffusive behaviour for the generalized Arnold example away from resonances, 2015, 85 pp., arXiv:1511.04835.

[52] A. Katok and B. Hasselblatt, Introduction to the modern theory of dynamical systems, Encyclopedia Math. Appl., vol. 54, Cambridge Univ. Press, Cambridge 1995, xviii+802 pp.

[53] M. Klein and A. Knauf, Classical planar scattering by Coulombic potentials, Lecture Notes in Phys. New Ser. Monogr., vol. M13, Springer-Verlag, Berlin 1992, v+142 pp.

[54] O. Knill, “Topological entropy of standard type monotone twist maps”, Trans. Amer. Math. Soc. 348:8 (1996), 2999–3013.

[55] В. В. Козлов, Д. В. Трецьёв, Биллиарды. Генетическое введение в динамику систем с ударами, Изд-во Моск. ун-та, М. 1991, 168 с.; English transl., V. V. Kozlov and D. V. Treschev, Billiards. A genetic introduction to the dynamics of systems with impacts, Transl. Math. Monogr., vol. 89, Amer. Math. Soc., Providence, RI 1991, viii+171 pp.
The anti-integrable limit

[56] L. M. Lerman, “Complex dynamics and bifurcations in a Hamiltonian system having a transversal homoclinic orbit to a saddle focus”, *Chaos* 1:2 (1991), 174–180.

[57] R. S. MacKay and J. D. Meiss, “Cantori for symplectic maps near the anti-integrable limit”, *Nonlinearity* 5:1 (1992), 149–160.

[58] J. N. Mather, “Existence of quasi-periodic orbits for twist homeomorphisms of the annulus”, *Topology* 21:4 (1982), 457–467.

[59] D. McDuff and D. Salamon, *Introduction to symplectic topology*, 2nd ed., Oxford Math. Monogr., The Clarendon Press, Oxford Univ. Press, New York 1998, x+486 pp.

[60] Z. Nitecki, *Differentiable dynamics. An introduction to the orbit structure of diffeomorphisms*, The M.I.T. Press, Cambridge, MA–London 1971, xv+282 pp.

[61] G. N. Piftankin, “Diffusion speed in the Mather problem”, *Nonlinearity* 19:11 (2006), 2617–2644.

[62] Г.Н. Пифтанкин, Д. В. Трецёв, “Сепаратрисное отображение в гамильтоновых системах”, *УМН* 62:2(374) (2007), 3–108; English transl., G.N. Piftankin and D.V. Treschev, “Separatrix map in Hamiltonian systems”, *Russian Math. Surveys* 62:2 (2007), 219–322.

[63] Л. П. Шильников, “Об одной задаче Пуанкаре–Биркгофа”, *Матем. сб.* 74(116):3 (1967), 378–397; English transl., L.P. Shilnikov, “On a Poincaré–Birkhoff problem”, *Math. USSR-Sb.* 3:3 (1967), 353–371.

[64] L. P. Shilnikov, A. L. Shilnikov, D. V. Turaev, and L. O. Chua, *Methods of qualitative theory in nonlinear dynamics. Part II*, World Sci. Ser. Nonlinear Sci. Ser. A Monogr. Treatises, vol. 5, World Scientific Publishing Co., Inc., River Edge, NJ 2001, xxiii+565 pp.

[65] S. Smale, “Diffeomorphisms with many periodic points”, *Differential and combinatorial topology*, A symposium in honor of M. Morse, Princeton Univ. Press, Princeton, NJ 1965, pp. 63–80.

[66] D. Treschev, “Multidimensional symplectic separatrix maps”, *J. Nonlinear Sci.* 12:1 (2002), 27–58.

[67] D. Treschev, “Trajectories in a neighbourhood of asymptotic surfaces of a priori unstable Hamiltonian systems”, *Nonlinearity* 15:6 (2002), 2033–2052.

[68] D. Treschev, “Evolution of slow variables in a priori unstable Hamiltonian systems”, *Nonlinearity* 17:5 (2004), 1803–1841.

[69] D. Treschev, “Arnold diffusion far from strong resonances in multidimensional a priori unstable Hamiltonian systems”, *Nonlinearity* 25:9 (2012), 2717–2757.

[70] D. Treschev and O. Zubelevich, *Introduction to the perturbation theory of Hamiltonian systems*, Springer Monogr. Math., Springer-Verlag, Berlin 2010, x+211 pp.

[71] D. Turaev, “Hyperbolic sets near homoclinic loops to a saddle for systems with a first integral”, *Regul. Chaotic Dyn.* 19:6 (2014), 681–693.

[72] Д. В. Тураев, Л. П. Шильников, “О гамильтоновых системах с гомоклиническими кривыми седла”, *Докл. AH CCCP* 304:4 (1989), 811–814; English transl., D. V. Turaev and L. P. Shilnikov, “On Hamiltonian systems with homoclinic saddle curves”, *Soviet Math. Dokl.* 39:1 (1989), 165–168.

[73] А. Б. Васильева, В. Ф. Бутузов, *Асимптотические разложения решений сингулярно-возмущенных уравнений*, Наука, М., 1973, 272 с. [A. B. Vasil’eva and V. F. Butuzov, *Asymptotic expansions of the solutions of singularly perturbed equations*, Nauka, Moscow 1973, 272 pp.]

[74] А. П. Веселов, “Интегрируемые отображения”, *УМН* 46:5(281) (1991), 3–45; English transl., A. P. Veselov, “Integrable maps”, *Russian Math. Surveys* 46:5 (1991), 1–51.
[75] P. Walters, An introduction to ergodic theory, Grad. Texts in Math., vol. 79, Springer-Verlag, New York–Berlin 1982, ix+250 pp.

Sergey V. Bolotin  
Steklov Mathematical Institute  
of Russian Academy of Sciences  
E-mail: bolotin@mi.ras.ru

Dmitrii V. Treschev  
Steklov Mathematical Institute  
of Russian Academy of Sciences  
E-mail: treschev@mi.ras.ru

Received 17/OCT/15  
Translated by THE AUTHORs