THE DUCK AND THE DEVIL: CANARDS ON THE STAIRCASE

J. GUCKENHEIMER AND YU. ILYASHENKO

Abstract. Slow-fast systems on the two-torus $T^2$ provide new effects not observed for systems on the plane. Namely, there exist families without auxiliary parameters that have attracting canard cycles for arbitrary small values of the time scaling parameter $\varepsilon$. In order to demonstrate the new effect, we have chosen a particularly simple family, namely $\dot{x} = a - \cos x - \cos y, \dot{y} = \varepsilon$, $a \in (1, 2)$ being fixed. There is no doubt that a similar effect may be observed in generic slow-fast systems on $T^2$. The proposed paper is the first step in the proof of this conjecture.


Key words and phrases. Slow-fast systems on the torus, canard solution, devil’s staircase, Poincaré map.

1. Introduction

Generic slow-fast systems in the plane

$$\dot{x} = f(x, y, \varepsilon), \quad \dot{y} = \varepsilon g(x, y, \varepsilon), \quad (x, y) \in \mathbb{R}^2$$

(1.1)

exhibit a particularly simple description of the shape of orbits as $\varepsilon \to 0$. Namely, the orbits of (1.1) tend to the orbits of a degenerate system constituted by alternating segments of phase curves $y = \text{const}$ of the fast system and stable portions of the slow curve $M: f(x, y, 0) = 0$.

In the presence of an auxiliary parameter, a new kind of limit behaviour of solutions was discovered in [D] (also see [DR] and the references therein). Namely, for the two-parameter family,

$$\dot{x} = y - x^2 - x^3, \quad \dot{y} = \varepsilon(a - y)$$

(1.2)

there is an exponentially narrow horn $a \in (a_1(\varepsilon), a_2(\varepsilon)), \quad \varepsilon_0 > \varepsilon > 0, \quad |a_2(\varepsilon) - a_1(\varepsilon)| < e^{-c/\varepsilon}$

such that for $(a, \varepsilon)$ inside this horn solutions are close to an attracting cycle constituted by arcs passing not only near phase curves of the fast system and stable...
portions of the slow curve, but also near unstable portions of the slow curve, see Fig. 2(a).

Solution of this type are called ducks or canards (ducks in French).

The existence of canard solutions in two-dimensional slow-fast systems was discovered in different situations, see [AAIS] and references there. But all the known examples contained an auxiliary parameter like $a$ in (1.2).

The system considered in the present paper

\begin{align}
\dot{x} &= a - \cos x - \cos y, \\
\dot{y} &= \varepsilon,
\end{align}

has no auxiliary parameter: $a$ in (1.2) is fixed, and one may replace $a$ by, say, 1.1. There is a sequence of intervals $C_n \subset \{ \varepsilon > 0 \}$, $C_n \to 0$ as $n \to \infty$, such for any $\varepsilon \in C_n$ (1.3) has an attracting canard cycle, see Fig. 1.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figures}
\caption{(a) Regular attracting cycle of system (1.3). (b) Canard cycle of the slow-fast system (1.3) on a torus.}
\end{figure}

A great majority of solutions are attracted to this cycle. Therefore, a random solution in the slice containing the slow curve looks like the one shown in Fig. 1. We conjecture that there is an open set of slow-fast system in $T^2$ with the same property. However, we have chosen a particular example in order to exhibit the new effect and elaborate the tools for the study of the generic case.

Canard solutions of slow-fast systems may be easily seen in the plane if we do not require that these solutions be attracting periodic. Indeed, consider the classical picture of a jump near a fold point of the slow curve, Fig. 2(b). A “majority” of solutions initiating to the left of the fold point approach the stable part of the slow curve, then gather in the exponentially narrow flow box of width $e^{-c/\varepsilon}$ and jump down near the fold point. This picture already contains the canard solutions; they may be seen if we pronounce the exorcism: “reverse the time”, see Fig. 2(c).
The initial conditions of positive semi-orbits with canard behaviour form an exponentially narrow tube. It is so narrow that in computer simulations it is shown by a single curve, see Fig. 2 (c).

System (1.3) is chosen to be reversible: the symmetry $\sigma_r: (x, y) \mapsto -(x, y)$, together with time reversal, brings system (1.3) to itself. The reason is twofold: reversibility facilitates the proofs; it is intimately related with the canard solutions. Yet we believe that reversibility is not necessary for the existence of canards without auxiliary parameters on the two-torus.
System (1.3) has a global cross-section \( \Gamma \) defined by \( y = \pm \pi \), and hence a Poincaré map \( P \), which is a diffeomorphism of \( \Gamma \) with rotation number \( \rho(\varepsilon) \). The rotation number varies continuously with \( \varepsilon \). In generic families of diffeomorphisms of the circle, the function \( \rho(\varepsilon) \) is a Cantor function whose graph is often called a “devil’s staircase” [M].

The steps of the devil’s staircase occur at rational values of \( \rho \) for generic families of circle maps. The lengths of these steps and the intervening gaps satisfy universal scaling properties that are intimately related to the continued fraction expansions of the rotation numbers. One of our main concerns is to examine this geometry for the Poincaré map of (1.3). The limit \( \varepsilon \to 0 \) of (1.3) is singular. As \( \varepsilon \to 0 \), the rotation number \( \rho(\varepsilon) \) increases without bound. At \( \varepsilon = 0 \), the zeros of the function \( a - \cos(x) - \cos(y) \) are equilibrium points of the system (1.3) and the flow no longer has a global cross-section. The singular nature of the limit \( \varepsilon \to 0 \) affects the scaling of the steps of the devil’s staircase. In particular, there is a set of dominant steps for which \( \rho(\varepsilon) \) is an integer and the Poincaré map has fixed points. We give an analysis here demonstrating that the gap lengths between these dominant steps in the devil’s staircase are very short: there is a constant \( c > 0 \) so that the gap lengths tend to 0 faster than \( e^{-c/\varepsilon} \).

The vector field (1.3) has a time reversing symmetry \( \sigma_r \) given by \( \sigma_r(x, y) = (-x, -y) \). Since \( \sigma_r \) reverses time, it maps stable periodic orbits of (1.3) into unstable periodic orbits of (1.3) and vice-versa. Periodic orbits that contain one of the fixed points \((0, 0), (\pi, 0), (\pi, 0) \) or \((\pi, \pi) \) of \( \sigma_r \) must be neutrally stable. (If a periodic orbit contains one fixed point of \( \sigma_r \), then it also contains a second one.) Such orbits occur for parameters at the endpoints of \( C_n \). For the Poincaré map \( P \), we have \( P_{-1}(-x) = -P(x) \), implying that the graph of \( P \) is symmetric with respect to the reflection

\[ \sigma_p: \mathbb{T}^2 \to \mathbb{T}^2, \quad (x, y) \mapsto (-y, -x) \]

in the circle \( x + y = 0 \). Thus \( P'(x) = (P_{-1})'(-x) \). If \( x + P(x) = 0 \), then \( P(x) = -x \) and \( P'(x) = 1 \). The neutral fixed points of \( P \) (if they exist) are 0 and \( \pi \).

The graph of the Poincaré map of (1.3) has very small slope at most points. The slope grows rapidly inside a small interval and decays just as quickly. The graph is approximated by vertical and horizontal circles that intersect on the fixed point circle \( x + y = 0 \) of \( \sigma_p \). The Poincaré map has exactly two fixed points, one with small slope and one with large slope, unless this intersection point is close to 0 or \( \pm \pi \). Characterizing the Poincaré map where its derivative changes from very small to very large is a critical part of our analysis. We decompose the Poincaré map in this region into a composition of maps produced by the flow across different vertical slices of the torus.

In any closed cylinder formed by the orbits of the fast system that do not intersect the slow curve, the slow-fast system is orbitally smoothly equivalent to a system for which the Poincaré map is a mere shift. The normalizing coordinates tend to a smooth limit as \( \varepsilon \to 0 \). (Theorem 2).

Near any closed arc in the stable (or unstable) part of the slow curve, the slow-fast system is orbitally smoothly equivalent to a linear one with respect to the fast variable. (Theorem 3).
The Poincaré map near these parts of the slow curves is a pure linear contraction (and expansion) with the coefficient exponentially tending to zero (respectively, to infinity) as $\epsilon \to 0$. Transitions from the unstable slow curve to the stable slow curve, or across the critical points of $y$ on the slow manifold are short enough to play an inconsequential role in determining the derivative of the Poincaré map. The dominant contribution to the derivative of the Poincaré map comes from the relative length of time that a trajectory spends close to the slow manifold. If the trajectory spends more time close to the stable slow curve than to the unstable slow curve, the slope of the Poincaré map is very small. If the trajectory spends more time close to the unstable slow curve $s^+$ than to the stable slow curve $s^-$, the slope of the Poincaré map is very large. The Poincaré map has a slope that is $O(1)$ only if the trajectory crosses from $s^+$ to $s^-$ near $y = 0$. Thus the bifurcations of the periodic orbits adjoin the region where the stable periodic orbits contain canards.

The main results of the paper are stated in Theorem 1 below. The outline of the proof is given in §2, using three lemmas that are proved in §4. Section §3 contains the proofs of Theorems 2 and 3. There are five additional supporting results that we call propositions. The propositions are proved where stated and serve to structure a portion of the proof of a theorem or lemma.

2. Canard cycles on the torus and the description of the Poincaré map

2.1. Statement of the theorem. Fix an arbitrary cross-section $J$ of the unstable part of the slow curve independent of $\epsilon$. We will call a solution of (1.3) that crosses $J$ a canard solution. The motivation is that in order to reach $J$ the solution must contain a segment close to an unstable arc of the slow curve whose length is independent of $\epsilon$. Later on we will specify the choice of $J$.

**Theorem 1.** There exists a sequence of intervals $R_n = (\alpha_n, \beta_n)$ and two sequences of disjoint intervals $C^\pm_n \subset R_n$ with the properties:

1. $|R_n| = O(e^{-cn})$.
2. $\alpha_n = O(1/n)$

and such that:

3. For small $\epsilon \not\in \bigcup R_n$, $\rho(\epsilon) = 0 \pmod{\mathbb{Z}}$. There are exactly two hyperbolic periodic orbits of (1.3), the unstable solution containing canards.
4. For any $\epsilon \in C^\pm_n$ system (1.3) has exactly two periodic orbits, both hyperbolic canard cycles. One cycle is stable and the other is unstable.

2.2. The Poincaré map. Let, as before, $\Gamma = \{y = -\pi\}$, $P_\epsilon : \Gamma \to \Gamma$ be the Poincaré map along the orbits of (1.3). Let $\gamma_\epsilon \subset S^1 \times S^1$ be the graph of $P_\epsilon$. The shape of the graph $\gamma_\epsilon$ tends to the union of a horizontal and a vertical circles as $\epsilon \to 0$. This is formalized in the following lemma:

**Lemma 1** (Shape Lemma). For any small $\epsilon > 0$ there exists an interval $D_\epsilon$ with the following properties:

1. $|D_\epsilon| = O(e^{-c/\epsilon})$. 

2. $|P'_\epsilon|_{S^1 \setminus D_\epsilon} = O(e^{-c/\epsilon})$.
3. In the notation $D'_\epsilon = -D_\epsilon$, $\Pi_\epsilon = D_\epsilon \times S^1$, $\Pi'_\epsilon = S^1 \times D'_\epsilon$, we have $\gamma_\epsilon \subset \Pi_\epsilon \cup \Pi'_\epsilon$.

The same statements hold for the inverse Poincaré map $P_{\epsilon}^{-1}$ with $D_\epsilon$ and $D'_\epsilon$ interchanged.

Lemma 1 is proved in 4.1. The next lemma formalizes the statement that the graph $\gamma_\epsilon$ of the Poincaré map moves monotonically to the upper left as $\epsilon \to 0$.

To formulate it, we need to specify a cross-section $J$ from the beginning of 2.1. Let $\tau = \cos^{-1}(a - 1)$ be the maximal value of $y$ on the slow curve

$$M = \{a - \cos(x) - \cos(y) = 0\}.$$ 

Pick $0 < \delta < \tau$; let $\alpha = \tau - \delta$; in what follows, $\delta$ will be chosen small. Consider the following cross-sections:

$$\Gamma^\pm = S^1 \times \{\mp \alpha\},$$

$$J^+ = \{(x, y) \in \Gamma^+ \mid x \in [0, \pi]\}, \quad J^- = \{(x, y) \in \Gamma^- \mid x \in [-\pi, 0]\}.$$ 

Solutions that cross $J^+$ will be called canard solutions; these solutions are mentioned in Theorem 1.

For two points $a, b$ on the oriented circle that split it in two nonequal arcs, let $a < b$ if $a$ is the left end (in the sense of the orientation of $S^1$) of the shortest arc with ends $a$ and $b$. Denote by $(a, b)$ the shortest arc that connects $a$ and $b$ oriented from $a$ to $b$; the orientation may be opposite to that of $S^1$.

For any $a, b$ as above, denote by $P^{a,b}_\epsilon$ the Poincaré map of the cross-section $y = a$ to $y = b$ along the orbits of (1.3) passing over the arc $(a, b)$.

In Lemma 1, choose $D_\epsilon = P^{\alpha,-\pi}_\epsilon(J^+)$, $D_\epsilon = [d^-_\epsilon, d^+_\epsilon]$, $d^-_\epsilon < d^+_\epsilon$. (2.1)

By definition, all the orbits of (1.3) that cross $D_\epsilon$ are canard solutions.

Denote by $A^\pm_\epsilon$ the points on the graph $\gamma_\epsilon$ lying above $d^\pm_\epsilon$:

$$A^\pm_\epsilon = (d^\pm_\epsilon, P_\epsilon(d^\pm_\epsilon)).$$

Let $\Delta' = \{x + y = 0\}$, $\{B^\pm_\epsilon\} = \gamma_\epsilon \cap \Delta'$, $B^\pm_\epsilon = (b^\pm_\epsilon, -b^\pm_\epsilon)$, $b^-(\epsilon) < b^+(\epsilon)$.

The circle $\Delta'$ is the mirror of the symmetry $\sigma_p$, so $P_\epsilon(b^\pm_\epsilon) = 1$. Hence, $b^\pm_\epsilon \in D_\epsilon$.

**Lemma 2 (Monotonicity Lemma).** Let $d^\pm_\epsilon$, $A^\pm_\epsilon$, $B^\pm_\epsilon$ be the liftings to the universal covers of $d^\pm_\epsilon$, $A^\pm_\epsilon$, $B^\pm_\epsilon$, which continuously depend on $\epsilon$. Then

1. $\frac{d}{d\epsilon} d^\pm_\epsilon \to \infty$, $\frac{d}{d\epsilon} (x - y)(C(\epsilon)) \to \infty$ as $\epsilon \to 0^+$ for any of the choices: $C(\epsilon) = A^+(\epsilon)$; $A^-(\epsilon)$; $B^+(\epsilon)$; $B^-(\epsilon)$.

2. Equation $d^-_\epsilon = -\pi n$ has the solution $\epsilon = \epsilon_n = O(1/n)$.

Now we can prove Theorem 1.
Figure 3. (a) Annuli in the Shape Lemma. (b–f) Application of the Monotonicity Lemma. Location of ε: (b) between $R_{n-1}$ and $R_n$; (c) onto $C_n^-$; (d) between $C_n^+$ and $C_n^-$; (e) onto $C_n^+$; (f) between $R_n$ and $R_{n+1}$.
2.3. Existence and lifespan of canard cycles. The proof of Theorem 1 is illustrated by Figure 3.

Remark. Figures 3 (b), (c), (e), (f) correspond to the case $\rho(\varepsilon) = 0 \pmod{1}$. All
the rest of the devil’s staircase is passed as $\varepsilon$ ranges between $C_n^+$ and $C_n^-$, see Fig. 3 (d).

Proof of Theorem 1. Define the segments $R_n$ in the following way:

$$R_n = \{ \varepsilon \mid d_x^- \leq -\pi n \leq d_x^+ \} := [\alpha_n, \beta_n], \quad R = \bigcup R_n. \quad (2.2)$$

By statement 1 of Lemma 2, $R_n$ is indeed a segment for large $n$. By statement 2 of Lemma 2, $\alpha_n = O(1/n)$. By statement 1 of Lemma 1, $|d_x^+ - d_x^-| = O(\varepsilon^{c/\varepsilon})$. By statement 1 of Lemma 2, this implies: $|R_n| = O(\varepsilon^{c/\varepsilon})$. This proves statements 1 and 2 of Theorem 1 for $R_n$ determined by (2.2).

To prove statement 3, denote by $K_\varepsilon$ the square $D_{\varepsilon} \times D_{\varepsilon}$, $D_\varepsilon = [d_x^-, d_x^+]$. Let $\Delta$ be the diagonal $x = y \pmod{2\pi}$ in $T^2$. The following proposition implies the third part of the theorem.

Proposition 1. Let $K_\varepsilon \cap \Delta = \emptyset$. Then $\rho(\varepsilon) = 0 \pmod{\mathbb{Z}}$ and $P_\varepsilon$ has exactly two hyperbolic fixed points.

Proof. On $S^1 \setminus D_{\varepsilon}$, the graph $\gamma_\varepsilon$ of $P_\varepsilon$ has slope smaller than one. The endpoints of this graph lie on the left and right boundaries of $K_\varepsilon$ at the points $A^- \varepsilon)$ and $A^+ \varepsilon)$. If we connect these points by a segment inside $K_\varepsilon$, we obtain a closed curve $\gamma_\varepsilon$ on $T^2$ that has homotopy type $(1,0)$ because outside $K_\varepsilon$ its slope is everywhere smaller than one, and it intersects each vertical circle in exactly one point. Consequently, $\gamma_\varepsilon$ intersects $\Delta$ at a unique point $p_\varepsilon$. The point $p_\varepsilon$ is not in $K_\varepsilon$ because $K_\varepsilon \cap \Delta = \emptyset$. Therefore, it is a fixed point of $P_\varepsilon$. It is stable because $P_\varepsilon'(p_\varepsilon) < 1$. Applying the symmetry $\sigma_\varepsilon$, we get a fixed point $p_\varepsilon = \sigma_\varepsilon(p_\varepsilon)$, which is unstable. This proves the proposition. \qed

By (2.2), for $\varepsilon \notin R$, $0, \pi \notin D_{\varepsilon}$; hence, $K_\varepsilon \cap \Delta = \emptyset$. An application of Proposition 1 proves statement 3 of Theorem 1. \qed

We can now prove the first part of statement 4, namely, the existence of canard cycles.

Let

$$C_n^- = \{ \varepsilon \in R_n \mid (y-x)(\bar{B}^- (\varepsilon) < 2\pi n < (y-x)(\bar{A}^- (\varepsilon)) \},$$

$$C_n^+ = \{ \varepsilon \in R_n \mid (y-x)(\bar{B}^+ (\varepsilon) < 2\pi n < (y-x)(\bar{A}^+ (\varepsilon)) \},$$

The Monotonicity Lemma implies that $C_n^-$ and $C_n^+$ are intervals for $n$ large. Obviously,

$$C_n^- \cup C_n^+ \subset R_n.$$

Let $\varepsilon \in C_n^- \setminus C_n^+$; case $\varepsilon \in C_n^+$ is treated in the same way. The points $A^- \varepsilon)$ and $B^- \varepsilon)$ lie on an arc $\lambda_\varepsilon$ of the graph $\gamma_\varepsilon$ and $\lambda_\varepsilon$ is contained in $K_\varepsilon$. By the definition of $C_n^-$, these points lie on the different sides of $\Delta$ in $K_\varepsilon$. Hence, $\lambda_\varepsilon$ crosses $\Delta$; denote the intersection point by $S_\varepsilon$. Below we prove that it is the unique stable point of $P_\varepsilon$. The point $S_\varepsilon = (s_\varepsilon, s_\varepsilon)$ corresponds to a canard cycle, because $s_\varepsilon \in D_\varepsilon$; hence, the cycle crosses $J^+$. The symmetric point $U_\varepsilon = -S_\varepsilon \in \gamma_\varepsilon \cap \Delta$ corresponds to another periodic solution, which is, in fact, unstable hyperbolic.
2.4. Uniqueness and stability of canard cycles. For any degree one smooth map \( P: S^1 \to S^1 \), between any two fixed points there exists a point at which \( P' = 1 \). This follows from Rolle’s theorem applied to the periodic function \( P(x) - x \). Consequently, the number of fixed points of \( P \) is bounded from above by the number of solutions of the equation \( P'_\varepsilon(x) = 1 \).

Lemma 3 gives us information about this equation.

Lemma 3. The set \( \{ x \in S^1 \mid P'_\varepsilon(x) \in [1/2, 2] \} \) consists of two arcs contained in \( D_\varepsilon \). On the left arc \( P'_\varepsilon \) increases, on the right one it decreases.

Lemma 3 is proved in 4.2.

Lemma 3 implies that there are always exactly two solutions of the equation \( P'_\varepsilon(x) = 1 \) on \( S^1 \). It also implies that \( P'_\varepsilon(S_\varepsilon) < 1 \) at the fixed point \( S_\varepsilon \) constructed in 2.3. Therefore, the canard cycle constructed in 2.3 is stable. This finishes the proof of the fourth and last statement of the theorem. Note that if there is a fixed point \( x \) of \( P_\varepsilon \) with \( P'_\varepsilon(x) = 1 \), then this is the only fixed point of the Poincaré map.

3. Normalization

3.1. Nonlinear translations.

Theorem 2. Consider a vector field on the cylinder \( x \in S^1 = \mathbb{R}/\mathbb{Z}, y \in \mathbb{R} \) defined by
\[
\dot{x} = f(x, y, \varepsilon), \quad \dot{y} = \varepsilon g(x, y, \varepsilon), \quad \text{where } f > 0, g > 0.
\]

If \( a, b \in \mathbb{R}, a < b, \) and \( \varepsilon > 0 \), then the Poincaré map \( P^{ab}_\varepsilon \) of a cross-section \( \Gamma^a: y = a \) to \( \Gamma^b: y = b \) has the form
\[
P^{a,b}_\varepsilon: x \mapsto G^1_\varepsilon \circ (G^2_\varepsilon(x) + T(\varepsilon)), \quad (3.1)
\]
where \( T(\varepsilon) \to +\infty \) and \( G^1_\varepsilon, G^2_\varepsilon \) are diffeomorphisms of the circle. As \( \varepsilon \to 0 \), we have \( G^1_\varepsilon, G^2_\varepsilon \to G_{1,2} \); the maps \( G_1, G_2 \) are still diffeomorphisms.

Proof. Let us divide the vector field by \( f \). We obtain a system with the same phase curves and the same Poincaré map:
\[
\dot{x} = 1, \quad \dot{y} = \varepsilon w(x, y, \varepsilon), \quad w > 0. \quad (3.2)
\]

Proposition 2. System (3.2) is smoothly equivalent to the system
\[
\dot{x} = 1, \quad \dot{y} = \varepsilon v(y, \varepsilon). \quad (3.3)
\]

Proof. Consider the time-one Poincaré map \( F_\varepsilon: \{ x = 0 \} \to \{ x = 0 \} \). It is smooth in \( y, \varepsilon \) and has the form
\[
F_\varepsilon: y \mapsto y + \varepsilon F(y, \varepsilon). \quad (3.4)
\]

Lemma 4. The smooth family (3.4) on any \( y \)-segment on which \( F \) is positive is smoothly equivalent to the family
\[
y \mapsto y + \varepsilon \quad \text{embeddable in the flow } \dot{y} = \varepsilon.
\]
This lemma is a global version of Theorem 7 from [1]. For the sake of completeness we reproduce here the proof of the above theorem with the minor changes needed for passing from germs to maps of segments.

Proof. Let us first extend (3.4) to the whole line $\mathbb{R}$, replacing $F$ by a smooth function equal to $F$ on the given segment, and to one outside some larger segment. Consider the vector field $F(y, 0) \frac{\partial}{\partial y}$ on the line, $F \neq 0$. It is smoothly equivalent to $\frac{\partial}{\partial y}$. In the rectifying chart, still denoted by $y$, the globalized map (3.4) has the form:

$$F : y \mapsto y + \varepsilon + \varepsilon^2 f(y, \varepsilon)$$

with $f \equiv 0$ outside some segment. We want to conjugate this map with $S : z \mapsto z + \varepsilon$, using the coordinate change $H : y \mapsto y + \varepsilon h(y, \varepsilon) = z$ smooth in $y$ and $\varepsilon \geq 0$. The functional equation $H \circ F = S \circ H$ takes the form:

$$h(y + \varepsilon + \varepsilon^2 f, \varepsilon) - h(y, \varepsilon) = -\varepsilon f(y, \varepsilon).$$

We will solve this equation by successive approximations, the Borel theorem and the homotopy method. Let $f = \varepsilon^n f_n(y) + \ldots$, here and below the dots replace higher order terms in $\varepsilon$. Let $h = \varepsilon^n h_n(y) + \ldots$. The functional equation implies $h_0'(\varepsilon) = -f_0$. Let

$$\tilde{f}(y, \varepsilon) = \sum f_n(u) \varepsilon^{n-1}$$

be a semiformal series (formal in $\varepsilon$ with functional coefficients in $x$) for the discrepancy. The successive approximation method produces a semiformal series

$$y + \tilde{h}, \quad \tilde{h} = \sum h_n(y) \varepsilon^n$$

that conjugates $y + \varepsilon + \varepsilon^2 \tilde{f}$ with $y + \varepsilon$ as semiformal series.

The Borel theorem provides a smooth map $y + \tilde{h}$ that conjugates $F$ with

$$F = y + \varepsilon + g(y, \varepsilon),$$

where $g$ is flat on the line $\{\varepsilon = 0\}$.

The last step is to prove that the map $F$ is $C^\infty$ embeddable. To this end consider the map $G : (y, \varepsilon, t) \mapsto (G, \varepsilon, t), \quad t \in [0, 1], \quad G = y + \varepsilon + tg$.

It is sufficient to find a vector field $(X, 1)$ in $\mathbb{R} \times (\mathbb{R}, 0) \times [0, 1]$ that is $G$-invariant. The corresponding equation is

$$G_y X - X \circ G = -g.$$

The solution of this equation is given by the formula

$$X = \sum_{0}^{\infty} (G_n^* g) \circ G^{-1};$$

here $G$ is regarded as a one-dimensional map in $y$ depending on $\varepsilon, t$ as parameters. Note that $g$ has a compact support because $f$ has the same property.

For a flat $g$ on the plane $\{\varepsilon = 0\}$ with compact support this formula gives for $\varepsilon \geq 0$ a vector field $X$ that admits a smooth extension by 0 through the line $\varepsilon = 0$. 


We omit the detailed estimates. Note only that the number of nonzero terms in the
sum above is of order $1/\varepsilon$ for $\varepsilon$ fixed, while the derivatives of $G^n$ for $n$ of order
$1/\varepsilon$ tend to zero faster than any power of $\varepsilon$. Lemma 4 is proved.

We have constructed system (3.3) which has the same Poincaré map as (3.2).
Let $V^t$ and $W^t$ be the time $t$ flow transformations for systems (3.3) and (3.2),
respectively. For $t = 1$ they bring the line $x = 0$ into itself, and coincide with the
Poincaré map on it. The map $H: (x, y) \mapsto V^x \circ W^{-x}(x, y)$ transforms the flow $W^t$
to $V^t$. Since the Poincaré maps $F_\varepsilon$ for (3.2) and (3.3) coincide,

$$H(1, y) = (1, F_\varepsilon \circ F_\varepsilon^{-1}(y)) = (1, y).$$

Hence, $H$ is well defined on the whole cylinder. This proves the proposition. The
map $H$ depends on $\varepsilon$ and tends to the identity as $\varepsilon \to 0$.  

The Poincaré map $\tilde{P}_{\varepsilon}^{a,b}: \{y = a\} \to \{y = b\}$ for system (3.3) is a rotation:

$$\tilde{P}_{\varepsilon}^{a,b}: x \mapsto x + T(\varepsilon), \quad T(\varepsilon) = \frac{1}{\varepsilon} \int_a^b \frac{dy}{v(y, \varepsilon)}.$$  \hspace{1cm} (3.5)

We complete the proof of Theorem 2 by calculating the map $P_\varepsilon^{a,b}$ in the normalizing chart that brings (3.2) to (3.3). Let $(x, a)$ be on the section $\Gamma^a = \{y = a\}$. Denote

$$w_\varepsilon(x) = \int_0^x w(\xi, a, 0) d\xi, \quad \omega(a) = w_\varepsilon(1).$$

Note that in (3.3), $v(a, 0) = \omega(a)$. Let $G_a: S^1 \to S^1$ be a diffeomorphism such
that $x' = G_a(x)$ implies $V^{x'} \circ W^{-x'}(x, a) \in \{y = a\}$.

The diffeomorphism $G_a$ depends on $\varepsilon$, but we do not indicate this by abuse of
notation. The point $p = (x, a) \in \Gamma^a$ is mapped by $H$ to $q = (x, y(x, a))$. Let $\varphi$ be the orbit of (3.2) passing through $p$, and $\psi$ be the orbit of (3.3) passing through $q$. Then $H(\varphi) = \psi$. The orbit $\psi$ crosses $\Gamma^a$ at the point $(G_a(x), a)$. The Poincaré map $P_\varepsilon^{a,b}$ of (3.3) takes this point to $(G_a(x) + T(\varepsilon), b) \in \psi$. The corresponding point on $\Gamma^b$ lying on $\varphi$ is obtained by the application of $G_b^{-1}$. Hence,

$$P_\varepsilon^{a,b}(x) = G_b^{-1} \circ (G_a(x) + T(\varepsilon)).$$

The map $G_a$ for $\varepsilon = 0$ may be easily found by using the variation equation with
respect to the parameter. For $\varepsilon = 0$, (3.2), (3.3) imply that $\frac{dw}{dx} = 0$. From Proposition 2 it follows that in (3.3) $v$ is at least $C^1$-smooth in $y, \varepsilon$. Hence,

$$V^{x'}(0, a) = (x', a + \varepsilon x' \omega(a) + o(\varepsilon)).$$

On the other hand

$$W^x(0, a) = (x, a + \varepsilon w_\varepsilon(x) + o(\varepsilon)).$$

The definition of $x'$ implies

$$x' \omega(a) = w_\varepsilon(x) + o(1).$$

Hence,

$$G_a(x) \to \frac{w_\varepsilon(x)}{\omega(a)} \quad \text{as} \quad \varepsilon \to 0$$

and the limit is a diffeomorphism of $S^1$. This proves Theorem 2.  \hspace{1cm} □
3.2. Normalization along slow curves.

Theorem 3. Consider a system
\[ \dot{x} = f(x, y, \epsilon), \quad \dot{y} = \epsilon g(x, y, \epsilon) \]
with a curve of hyperbolic fixed points of the fast system, \( \epsilon > 0 \) small and \( g > 0 \).
In the neighborhood of the curve of fixed points, the system is orbitally smoothly equivalent to a family linear in a fast variable. The normal form is:
\[ \dot{x} = a(y, \epsilon)x, \quad \dot{y} = \epsilon. \]  
(3.6)

Remark. For system (1.3), the normal form near a segment \( x = s(y, \epsilon) \) of the true slow curve is
\[ \dot{x} = (\sin s(y, \epsilon))x, \quad \dot{y} = \epsilon \]  
(3.7)

Proof. When \( \epsilon > 0 \) is small, near a curve \( \gamma_0 \) of hyperbolic fixed points of the fast system there exists an invariant curve \( \gamma_\epsilon \), “the true slow surface,” close to \( \gamma_0 \) by the Fenichel theorem [F]. Therefore the system may be written in the form
\[ \dot{x} = x f_1(x, y, \epsilon), \quad \dot{y} = \epsilon g_1(x, y, \epsilon), \quad g_1 > 0. \]
Dividing by \( g_1 \), we obtain a vector field with the same trajectories and
\[ \dot{x} = x a(y, \epsilon) + O(x^2), \quad \dot{y} = \epsilon. \]  
(3.8)

We assert that system (3.8) may be transformed by a finitely smooth coordinate change
\[ x = X + \epsilon h(X, Y, \epsilon), \quad y = Y \]  
(3.9)
to (3.6) in which \( x \) is replaced by \( X \). To do so, we extend the family (3.8) to a single system by adding the equation \( \dot{\epsilon} = 0 \). This equation has a 2-dimensional invariant manifold \( x = 0 \). We will prove that many lower nonlinear terms in \( x \) may be killed by the coordinate change (3.9). Then (3.8) will be equivalent to
\[ \dot{x} = x a(y, \epsilon) + \epsilon x^N b_1(x, y, \epsilon), \quad \dot{y} = \epsilon, \quad N \gg 1. \]  
(3.10)
The jets of order \( N - 1 \) of systems (3.6) and (3.10) coincide on \( x = 0 \). Hence, the systems are finitely smoothly equivalent by the Belitski—Samovol theorem [II].

It remains to prove the equivalence of systems (3.8) and (3.10). For this, we use the method of successive approximations and conjugate system (3.10) with a similar system in which \( N \) is replaced by \( N + 1 \). The composition of these conjugations gives the desired transformation. So it remains to prove the following:

Proposition 3. The system
\[ \dot{x} = xa + \epsilon x^N b(y, \epsilon) + O(x^{N+1}), \quad \dot{y} = \epsilon \]
may be transformed by the coordinate change
\[ x = X + \epsilon X^N h(Y, \epsilon), \quad y = Y \]
The orbits of system \( (Y, \varepsilon) \) has the form \( a \gamma \), where \( \gamma \) is the unstable portion of the slow curve. The function \( h \) is uniformly in \( x \) initial coordinate \( s \).

\[ \varepsilon h_{x^s} = a(1 - N)h + b. \]  

Equation (3.11) is equivalent to the system

\[ \hat{h} = a(1 - N)h + b, \quad \hat{Y} = \varepsilon. \]  

Proof. An elementary calculation gives the following equation for \( h \):

\[ \varepsilon h_{y^s} = a(1 - N)h + b. \]  

For \( \varepsilon = 0 \), (3.11) has a smooth solution \( h = b/a \). This is the slow curve for (3.12) with \( \varepsilon = 0 \). The “true slow curve” for (3.12) with \( \varepsilon > 0 \) is the desired solution \( h(Y, \varepsilon) \) of (3.11). Hence, (3.11) has a smooth solution. This proves Proposition 3 and completes the proof of Theorem 3.

4. The Poincaré Map

In this section Lemmas 1, 2 and 3 are proved.

4.1. Distortion: Proof of Lemma 1. Let \( \alpha = \tau - \delta \), \( \sigma = [\alpha, \alpha] \). Let \( s^+ \) and \( s^- \) be the unstable and stable portions of the slow curve over \( \sigma \). Let, as before,

\[ J^+ = \Gamma^+ \cap \{ x \in [0, \pi] \}, \quad J^- = \Gamma^- \cap \{ x \in [-\pi, 0] \}. \]

Denote, for brevity, \( Q_\varepsilon = P_{\varepsilon}^{-\tau, -\alpha} : \Gamma \to \Gamma^+ \), and let \( D_\varepsilon = (Q_\varepsilon)^{-1}(J^+) \).

Proposition 4. 1. \( |D_\varepsilon| = O(e^{c_1(\delta)/\varepsilon}) \), \( c_1(\delta) \to 0 \) as \( \delta \to 0 \).

2. For any \( y_0 \in [-\pi, -\alpha] \) the derivative of the map along the orbits

\[ P_{\varepsilon}^{-\tau, y_0} : \Gamma \to \Gamma_{y_0} = \{ y = y_0 \} \]

is no greater than \( O(e^{c_1(\delta)/\varepsilon}) \), \( c_1(\delta) = O(\delta^{3/2}) \).

Proof. Consider a decomposition of \( Q_\varepsilon^{-1} \) into three factors. Take

\[ \Gamma_1 = \{ y = -\tau + \delta^2 \}; \quad \Gamma_2 = \{ y = -\tau - \delta^2 \} \]

and let \( Q_1 : \Gamma^+ \to \Gamma_1; \quad Q_2 : \Gamma_1 \to \Gamma_2 \) and \( Q_3 : \Gamma_2 \to \Gamma \) be the mappings along the orbits of system (1.3) with time reversed. Then

\[ Q_\varepsilon^{-1} = Q_3 \circ Q_2 \circ Q_1. \]

To study \( Q_1 \), apply the normal form (3.7). Let \( x = x(y, \varepsilon) \) be the true slow curve for system (1.3) regarded over the segment \( [-\tau + \delta^2, \tau - \delta^2] \) where \( x = x(y, 0) > 0 \) is the graph of the unstable portion of the slow curve. The function \( a(y, \varepsilon) = \sin x(y, \varepsilon) \) has the form

\[ a(y, \varepsilon) = (C(a) + o(1))\sqrt{y + \tau} + O_\delta(\varepsilon), \]

where \( o(1) \to 0 \) as \( \delta \to 0 \); \( O_\delta(\varepsilon) \) is of order \( \varepsilon \) for any fixed \( \delta \). The constant \( C(a) \) may be calculated explicitly, but its value is of no importance. The mapping \( Q_1 \) is linear in its normal form coordinate \( x_1 \):

\[ Q_1 : \quad x_1 \mapsto \Lambda(\varepsilon)x_1, \quad \Lambda(\varepsilon) = e^{-\frac{3\delta^2}{2}(C(a) + o(1) + O_\delta(\varepsilon))} < e^{-\frac{3\delta^2}{4}C(a)} \]

for small \( \delta \) and \( \varepsilon < \varepsilon_0 \), where \( \varepsilon_0 \) depends on \( \delta \). Let us now estimate \( Q_3^2 \) in terms of its initial coordinate \( x \). (The transition function from \( x \) to \( x_1 \) has bounded derivative uniformly in \( \varepsilon \), so it may be neglected in future calculations.) The variational
equation for the derivative of the solutions of (1.3) (with time reversed) with respect to its initial conditions has the form

\[ X' = -\frac{1}{\varepsilon} \sin x(y)X, \tag{4.1} \]

where \( x = x(y) \) is the orbit of (1.3). The time of motion from \( \Gamma_1 \) to \( \Gamma_2 \) is equal to \( 2\delta^2 \).

Hence, \( Q_2 \leq e^{2\delta^2} \ll \Lambda^{-1}(\varepsilon) \).

In particular, \( Q_1^2 Q_2' \leq e^{-\delta^2 / 2} c \), \( c = O(\delta^3 / 2) \) for small \( \varepsilon \). The map \( Q_3 \) is very simple. By Theorem 2, it is a rotation in a specific chart combined with transition functions that have bounded derivatives. Therefore, \( (Q_2^{-1})' \setminus J^{-} \) is contracting, by (4.2). This proves statement 1 of Proposition 4 and, consequently, statement 1 of Lemma 1.

The same argument shows that for any \( y_0 \in [-\pi, -\alpha] \)

\[ |(P_+^{-\pi,y_0})'| \leq e^{c_1 / \varepsilon}, \quad c_1 = O(\delta^3 / 2) \tag{4.2} \]

with another constant \( c_1 \). The proposition is proved.

Statement 2 of Lemma 1 follows from Theorem 3, (4.2) and reversibility. Indeed,

\[ P_\varepsilon = \tilde{Q}_\varepsilon \circ P_+^\varepsilon \circ Q_\varepsilon, \]

where \( Q_\varepsilon: \Gamma \to \Gamma^+, P_+^\varepsilon: \Gamma^+ \to \Gamma^- \) and \( \tilde{Q}_\varepsilon: \Gamma^- \to \Gamma \) are maps along the orbits of (1.3). Let \( \sigma_r \) be the time reversing symmetry \( \sigma_r(x, y) = (-x, -y) \). Note that \( \sigma_r(\Gamma^+) = \Gamma^- \). By reversibility, \( \tilde{Q}_\varepsilon = Q_\varepsilon^{-1} \circ \sigma_r \). Hence, \( \tilde{Q}_\varepsilon | J^- \) is contracting, by (4.2). We have

\[ Q_\varepsilon(\Gamma - D_\varepsilon) = \Gamma^+ - J^+ \]

by the definition of \( D_\varepsilon \). As is shown below, \( P_+^\varepsilon(\Gamma^+ - J^+) \subset J^- \), and \( \tilde{Q}_\varepsilon \) is contracting on \( J^- \).

Orbits starting on \( \Gamma^+ - J^+ \) enter a fixed neighborhood of the stable portion of the slow curve over the segment \( \sigma \) in a time \( O(1) \). By Theorem 3, the derivative of the map along the orbits defined on a segment close to \( \Gamma^+ \) in this neighborhood is contracting with the coefficient

\[ \Lambda \ll e^{-C / \varepsilon}. \]

Hence, \( P_+^\varepsilon(\Gamma^+ - J^+) \subset J^- \).

Now, in the decomposition for \( P_\varepsilon \), only the first factor is expanding, but this expansion is strongly dominated by the contraction of \( P_+^\varepsilon \). (We study \( P_+^\varepsilon \) more thoroughly in 4.2.) This proves statement 2 of Lemma 1.

Statement 3 of Lemma 1 follows from statement 2 and reversibility. Reversibility implies that the graph \( \gamma_\varepsilon \) of \( P_\varepsilon \) is symmetric with respect to the diagonal \( x + y = 0 \). The symmetry \( \sigma_p \) interchanges \( \Pi_1 \) and \( \Pi_2 \). If there is a point \( p = (x, y) \in \gamma \) outside \( \Pi_1 \cup \Pi_2 \), then the symmetric point \( q = (-y, -x) \in \gamma \) also lies outside \( \Pi_1 \cup \Pi_2 \). By
parts of the true slow curve of \((y)\) intersects the normalizing coordinates. By Lemma 1, the graph of the solution \(x\) Let \(U\) and \(s\) below that these orbits stay near the unstable part \(y\) stable part \(s\) direction along the \(x\)

We will prove that the second term dominates the two others. \(P\) We decompose the Poincaré map \(P\) as above:

\[ P_\varepsilon = Q_\varepsilon \circ P^- \circ Q_\varepsilon. \]

By the chain rule

\[ \log(P_\varepsilon) = \log \tilde{Q}_\varepsilon' \circ P_+ \circ Q_\varepsilon + \log(P_+)' \circ Q_\varepsilon + \log Q_\varepsilon'. \]

Hence,

\[ \frac{P''_\varepsilon(u)}{P'_\varepsilon(u)} = \frac{d}{du} \log \tilde{Q}_\varepsilon' \circ P_+ \circ Q_\varepsilon(u) + \frac{d}{du} \log(P_+)' \circ Q_\varepsilon(u) + \frac{d}{du} \log Q_\varepsilon'(u). \]

We will prove that the second term dominates the two others.

First of all let us describe the orbits \(x = x(y, u)\) with initial condition \(u\) such that \(P_\varepsilon(u) \in [1/2, 2]\). The dependence on \(\varepsilon\) is omitted in the notation. We will prove below that these orbits stay near the unstable part \(s^+\) of the slow curve over the arc \([-\alpha, y_+]\) with \(y_+\) close to 0. Then some orbits jump “downward” (in negative direction along the \(x\)-circle), the others jump “upward” to the neighborhood of the stable part \(s^-\) of the slow curve, and remain there when \(y\) ranges in \([y_-, \alpha]\), where \(y_-\) is \(O(\varepsilon)\)-close to \(y_+\). For the orbits under consideration that jump downward, we have \(P''_\varepsilon > 0\); for those jumping upward, \(P''_\varepsilon < 0\).

Note that the variational equation implies:

\[ \log(P_{\varepsilon}^a, b)'(u) = \frac{1}{\varepsilon} \int_a^b \sin x(y; u, a) dy, \]

\[ \frac{d}{du} \log(P_{\varepsilon}^a, b)'(u) = \frac{1}{\varepsilon} \int_a^b \cos x(y; u, a) X(y; u, a) dy, \]

where \(x(y; u, a)\) is a solution with

\[ x(a; u, a) = u; \quad X(y; u, a) = \frac{\partial x(y; u, a)}{\partial u}. \]

Let \(x = x^\pm(y, \varepsilon)\) be the equation of the unstable (for +) and stable (for −) parts of the true slow curve of (1.3) over the segment \(\sigma\). Let \(s^\pm(y) = x^\pm(y, 0)\). Let \((x_1, y), (x_2, y)\) be normalizing charts from Theorem 3 for equation (1.3) near \(s^+\) and \(s^-\), respectively. From now on we suppose that \(u \in \Gamma\) satisfies

\[ P'_\varepsilon(u) \in [1/2, 2], \]

Let \(U, S\) be neighborhoods of \(s^+\), \(s^-\) defined by \(|x_1| < b, |x_2| < b\) respectively in the normalizing coordinates. By Lemma 1, the graph of the solution \(x(y; u, −\pi)\) intersects \(J^+\) and \(J^-\). Let this solution leave \(U\) with \(y = y_+\) and enter \(S\) with \(y = y_-\). Let us prove first that

\[ |y_+| \leq \frac{3\delta}{C} \quad \text{for } \varepsilon \text{ small}. \]
Indeed, by \((4.5)\),
\[
\log(P^+)' = \frac{1}{\varepsilon} \left( \int_{-\alpha}^{\alpha} \sin s^+(y)dy + \int_{\gamma}^{\alpha} \sin s^-(y)dy \right) + O(1).
\]
The symmetry of the vector field implies that \(s^+(y) = -s^-(y)\), yielding
\[
\log(P^-)' = \frac{1}{\varepsilon} \int_{-\gamma}^{\gamma} \sin s^+(y)dy + O(1).
\]
On the other hand, the \(x\)-component of the vector field of \((1.3)\) is bounded away from zero outside any neighborhood of the slow curve. Hence, \(y^+ = y^- + O(\varepsilon)\). Therefore,
\[
|\log(P^-)'| > c\left|\frac{y^+}{\varepsilon}\right|
\]
with some \(c\) depending on \(a\) in \((1.3)\). We have proved in 4.1 that
\[
|\log Q'_\varepsilon| < \frac{\delta}{\varepsilon}, \quad |\log \tilde{Q}'_\varepsilon| < \frac{\delta}{\varepsilon}.
\]
Together with \((4.4)\) and \((4.7)\) this yields \((4.8)\).

**Proposition 5.** Let \(u \in \Gamma\) satisfy \((4.7)\). Let
\[
I = \int_{-\alpha}^{\alpha} \sin s^+(y)dy.
\]
Then for \(\varepsilon\) small,
\[
\left| \frac{d}{du} \log (P^-)' \circ Q_\varepsilon(u) \right| > \exp \frac{I}{5\varepsilon}.
\]

**Proof.** Let us use Theorem 3 applied to equation \((1.3)\). In the normalizing coordinates \((x_1, y), (x_2, y)\) near the unstable and stable parts of the slow curve, respectively, \((1.3)\) has the form
\[
\frac{dx_1}{dy} = \frac{1}{\varepsilon} \varphi(y, \varepsilon)x_1, \quad \frac{dx_2}{dy} = \frac{1}{\varepsilon} \psi(y, \varepsilon)x_2,
\]
where \(\varphi(y, \varepsilon) = \sin x^+(y, \varepsilon), \psi(y, \varepsilon) = \sin x^-(y, \varepsilon), y \in \sigma = [-\alpha, \alpha]\). By the reversibility of \((1.3)\),
\[
\psi(y, \varepsilon) = -\varphi(-y, \varepsilon).
\]
The function \(a - \cos x - \cos y\) is even in \(y\). Hence, \(\varphi|_{\varepsilon=0}\) is even as well. Therefore,
\[
\psi(y, \varepsilon) = -\varphi(y, \varepsilon) + \varepsilon \varphi_1(y, \varepsilon).
\]
Consider the primitive functions of \(\varphi\) and \(\psi\) with respect to \(y\):
\[
\Phi' = \varphi, \quad \Phi(-\alpha, \varepsilon) = 0; \quad \Psi' = \psi, \quad \Psi(\alpha, \varepsilon) = 0.
\]
Then, by \((4.12)\),
\[
\Phi(-y, \varepsilon) = \Psi(y, \varepsilon).
\]
Let \(I\) be the same as in \((4.9)\). Then
\[
\Phi(-y, 0) = I - \Phi(y, 0), \quad \Phi(-y, \varepsilon) = I - \Phi(y, \varepsilon) + \varepsilon k_1(y, \varepsilon),
\]
where \(k_1\) is a smooth function for all \(\varepsilon\) near 0, \(y \in \sigma\). The functions \(k_j\) that appear below are also smooth for all \(\varepsilon\) near 0 and the same \(y\).
The orbit of (1.3) passing from \((y, x_1) = (-\alpha, \xi) \in J^+\) and jumping downward to \((y, x_2) = (\alpha, \eta(\xi)) \in J^-\) first intersects the line \(x_1 = -b\), then \(x_2 = b\). The first intersection point is \((y_+, -b)\), the second one is \((y_-, b)\). \(y_\pm\) depend on \(\xi\). Note that the function \(\xi \mapsto \eta(\xi)\) gives the map \(P_\pm^r\) in the normalizing charts.

Obviously,

\[
y_+ = y_- + \varepsilon k_2(y_+, \varepsilon), \tag{4.16}
\]

Solutions of equations (4.11) passing through \(p\),

\[
(y, x_1)(p) = (-\alpha, \xi), \quad q = P_-(p), \quad (y, x_2)(q) = (\alpha, \eta(\xi))
\]

have the form

\[
x_1(y, \varepsilon) = \xi \exp \frac{\varepsilon}{\xi} \Phi(y, \varepsilon), \quad x_2(y, \varepsilon) = \eta(\xi) \exp \frac{1}{\varepsilon} \Psi(y, \varepsilon).
\]

Then

\[
-b = \xi \exp \frac{1}{\varepsilon} \Phi(y_+(\xi), \varepsilon), \quad b = \eta(\xi) \exp \frac{1}{\varepsilon} \Psi(y_-(\xi), \varepsilon),
\]

\[
\eta(\xi) = -\xi \exp \frac{\Phi(y_+, \varepsilon) - \Psi(y_-, \varepsilon)}{\varepsilon}. \tag{4.17}
\]

By (4.14), (4.15), (4.16):

\[
\Phi(y_+, \varepsilon) - \Psi(y_-, \varepsilon) = \Phi(y_+, \varepsilon) - \Phi(-y_-, \varepsilon) = \Phi(y_+, \varepsilon) - \Phi(-y_+, \varepsilon) + \varepsilon k_3(y_+, \varepsilon)
\]

\[
= 2\Phi(y_+, \varepsilon) - I + \varepsilon k_4(y_+, \varepsilon).
\]

On the other hand,

\[
\exp \frac{\Phi(y_+, \varepsilon)}{\varepsilon} = -\frac{b}{\xi}.
\]

Hence,

\[
\eta(\xi) = -\xi \exp \left( \frac{2\Phi(y_+, \varepsilon)}{\xi} - \frac{I}{\varepsilon} + k_4(y_+, \varepsilon) \right) = -\frac{b^2}{\xi} \exp \left( -\frac{I}{\varepsilon} + k_4(y_+, \varepsilon) \right).
\]

We will show that this function behaves like \(C\xi^{-1}\) for \(\xi < 0\), hence, is concave. In order to study \(\eta\) as a function of \(\xi\), let us express \(y_+\) in terms of \(\xi\) by making use of (4.17). We have:

\[
\Phi(y_+, \varepsilon) = \varepsilon (\log b - \log(-\xi)).
\]

The function \(\Phi\) is monotonic in \(y_+\); let \(z_\varepsilon\) be the inverse function: \(z_\varepsilon(\Phi(y_+, \varepsilon)) \equiv y_+\). Then \(y_+ = z_\varepsilon(\varepsilon (\log b - \log(-\xi)))\). Hence,

\[
b^2 \exp k_4(y_+, \varepsilon) = k_3(\varepsilon (\log b - \log(-\xi)), \xi); \]

\(k_3(u, \varepsilon)\) is smooth in \(u, \varepsilon\). Let for brevity \(k_3 = k\). Then

\[
\eta'(\xi) = \frac{k(\varepsilon (\log b - \log(-\xi))) + \varepsilon k_u(\varepsilon (\log b - \log(-\xi)), \xi)}{\xi^2} \exp \left( -\frac{I}{\varepsilon} \right),
\]

\[
\log \eta'(\xi) = -2 \log(-\xi) + \log(k + \varepsilon k_u) - \frac{I}{\varepsilon},
\]

\[
\frac{d}{d\xi} \log \eta'(\xi) = -\frac{2}{\xi} + \frac{k_u + \varepsilon k_{uu} \varepsilon}{k + \varepsilon k_u} \xi = -\frac{2 + O(\varepsilon)}{\xi} > -\frac{1}{\xi}, \quad \xi < 0.
\]
Hence, the function $\eta(\xi)$ is concave on the $\xi$-segment that we study. Let us estimate this concavity. For this let us estimate $1/|\xi|$ from below. By (4.17)

$$-\xi = b \exp -\frac{\Phi(y_+, \varepsilon)}{\varepsilon}.$$ 

By (4.8),

$$\Phi(y_+, \varepsilon) > \frac{I}{2} - |y_+| \max_\sigma \varphi > \frac{I}{3}$$

for $\delta$ small. Hence,

$$\frac{d}{d\xi} \log \eta'(\xi) > -\frac{1}{\xi} > b^{-1} \exp \frac{I}{3\varepsilon}.$$  (4.18)

When we pass from the normalizing coordinates $\xi, \eta$ to the initial coordinate $x$ bounded terms only will be added; so the following estimate holds for $P_\varepsilon^{-}$:

$$\frac{d}{dx} \log (P_\varepsilon^{-})'(u) > b^{-1} \exp \frac{I}{4\varepsilon},$$

provided that $u$ satisfies (4.7), and the corresponding orbit jumps downwards from $U$ to $S$. By (4.2),

$$\frac{d}{du} \log (P_\varepsilon^{-})'(u) \leq \frac{1}{\varepsilon} \int_{-\alpha}^{-\pi} X(y; u, -\pi) dy.$$ 

By Statement 2 of Proposition 4,

$$X(y; u, -\pi) = (P_\varepsilon^{-})'(u) \leq O(e^{O(\delta^{3/2})}).$$

Thus proves Proposition 5 for orbits jumping downward.

For the orbit jumping upward the same evaluations go with the only difference that (4.17) should be replaced by

$$b = \xi \exp \frac{\Phi(y_+, \varepsilon)}{\varepsilon}, \quad -b = \eta(\xi) \exp \frac{\Psi(y_-, \varepsilon)}{\varepsilon},$$

and (4.18) becomes

$$\frac{d}{d\xi} \log \eta'(\xi) < -\frac{2 + O(\varepsilon)}{\xi} < -\frac{1}{\xi} < -b^{-1} \exp \frac{I}{3\varepsilon}; \quad \xi > 0.$$ 

The derivative in $u$ is estimated as above. This concludes the estimate of the second term in (4.4) and proves Proposition 5. \hfill \Box

Now let us estimate the two other terms in (4.4). By (4.6),

$$\left| \frac{d}{du} \log Q'_\varepsilon(u) \right| \leq \frac{1}{\varepsilon} \int_{-\alpha}^{-\pi} X(y; u, -\pi) dy.$$ 

By Statement 2 of Proposition 4,

$$X(y; u, -\pi) = (P_\varepsilon^{-})'(u) \leq O(e^{O(\delta^{3/2})}).$$

Hence,

$$\left| \frac{d}{du} \log Q'_\varepsilon(u) \right| \leq O(e^{O(\delta^{3/2})}).$$

On the other hand, for $y \in [\alpha, \pi],$

$$X(y; u, \pi) = \frac{d}{du} P_\varepsilon^{-y} \circ P_\varepsilon = (P_\varepsilon^{-y})' \circ P_\varepsilon(u) \cdot P'_\varepsilon(u).$$
The last factor is in $[1/2, 2]$ by assumption of the lemma. The first one is $O(\varepsilon^{1/4})$ by reversibility and Proposition 2. Hence, the third term in (4.4) admits the same estimate as the first one. Therefore, $\frac{d}{d\mu} \log P_{\varepsilon}'$ has the same sign as the second term in (4.4) for $\delta$ small. This proves Lemma 3.

4.3. Parameter variations: proof of Lemma 2.

Lemma 2 (Monotonicity Lemma). Let $d_{\varepsilon}^{\pm}, \tilde{A}^{\pm}(\varepsilon), \tilde{B}^{\pm}(\varepsilon)$ be liftings to the universal covers of $d_{\varepsilon}^{\pm}, A^{\pm}(\varepsilon), B^{\pm}(\varepsilon)$ which continuously depend on $\varepsilon$. Then

1. $\frac{d}{d\varepsilon} d_{\varepsilon}^{\pm} \to \infty, \quad \frac{d}{d\varepsilon} (x - y)(C(\varepsilon)) \to \infty \quad \text{as} \quad \varepsilon \to 0^+$ \quad (4.19)

   for any of the choices: $C(\varepsilon) = \tilde{A}^{+}(\varepsilon); \tilde{A}^{-}(\varepsilon); \tilde{B}^{+}(\varepsilon); \tilde{B}^{-}(\varepsilon)$.

2. Equation $d_{\varepsilon}^{-} = -\pi n$ has a solution $\varepsilon = \varepsilon_n = O(\frac{1}{n})$.

Proof. Let us first prove (4.19) for $d_{\varepsilon}^{+}$ and $\tilde{B}^{+}(\varepsilon)$. The proof for $d_{\varepsilon}^{-}, \tilde{A}^{\pm}(\varepsilon), \tilde{B}^{-}(\varepsilon)$ requires a few additional arguments provided below.

By definition, $D_{\varepsilon} = Q_{\varepsilon}^{-1}(J^{+})$, where $Q_{\varepsilon}$ is the map $\Gamma \to \Gamma^{+}$ along the phase curves of (1.3). Here $\Gamma = \{y = -\pi\}, \Gamma^{+} = \{y = -\alpha\}, \alpha = \tau - \delta$. We have:

$$d_{\varepsilon}^{-} = Q_{\varepsilon}^{-1}(\pi).$$

Let $Q_{3} \circ Q_{2} \circ Q_{1}$ be the decomposition of $Q_{\varepsilon}^{-1}$ defined in 4.1. Denote $R_{\varepsilon} = Q_{2} \circ Q_{1}$, now stressing its dependence on $\varepsilon$. By (3.1),

$$Q_{3}(x) = x + T(\varepsilon) + \varphi_{\varepsilon}(x). \quad (4.20)$$

Here $\varphi_{\varepsilon}$ is an $\varepsilon$-dependent $x$ periodic function uniformly bounded in the $C^{1}$-norm as $\varepsilon \to 0$; $T(\varepsilon)$ is provided by (3.5):

$$T(\varepsilon) = \frac{1}{\varepsilon} \int_{-\tau - \delta^{2}}^{-\pi} \frac{dy}{v(y, \varepsilon)}.$$

Now we have

$$v(y, \varepsilon) = v(y, 0) + \varepsilon v_{1}(y, \varepsilon), \quad v(y, 0) = \int_{0}^{2\pi} \frac{dx}{a - \cos x - \cos y},$$

Hence,

$$T(\varepsilon) = \frac{C}{\varepsilon} + T_{1}(\varepsilon), \quad C = \int_{-\tau - \delta^{2}}^{-\pi} \frac{dy}{v(y, 0)} < 0 \quad (4.21)$$

with $T_{1}$ smooth in $\varepsilon$. Note that $C < 0$ since $-\pi < -\tau - \delta^{2}$. Differentiating (4.21), we find

$$T'(\varepsilon) \to +\infty \quad \text{as} \quad \varepsilon \to 0^+. \quad (4.22)$$

The derivative of $R_{\varepsilon}(\pi)$ in $\varepsilon$ is obviously positive. Indeed, outside the domain $M^{-}; a - \cos x - \cos y \leq 0$ bounded by the slow curve $M$, the orbit of system (1.3) is a curve that is transverse to the orbits of the same system with $\varepsilon$ replaced by a larger $\mu; \varepsilon < \mu$. Denote the first system by (1.3)$_{\varepsilon}$, and the second one by (1.3)$_{\mu}$. The map $R_{\varepsilon}$ is the map along the orbits of (1.3)$_{\varepsilon}$ with time reversed. The vectors of
The derivative $s\epsilon$ coordinates is an $Q$ proved for $\mu$, hence, for $\epsilon < \mu$,

$$R_\epsilon(\pi) < R_\mu(\pi) \implies \frac{d}{d\epsilon} R_\epsilon(\pi) \geq 0.$$ 

Together with (4.20) and (4.22), this proves (4.19).

For $B^+(\epsilon)$ the same argument works. Namely, $(x - y)(B^+(\epsilon)) = 2b^+(\epsilon) = 2P_{\epsilon}^{0,-\pi}(\pi)$. The corresponding orbit lies outside $M^-$, because no orbit quits $M^-$ in the domain $\{y \in [-\pi, 0]\}$ and no orbit enters $M^-$ in $\{y \in [0, \pi]\}$.

Decompose

$$P_{\epsilon}^{0,-\pi}(\pi) = Q_{3\epsilon} \circ \tilde{R}_\epsilon(\pi), \quad \tilde{R}_\epsilon = P_{\epsilon}^{0,-\tau-\delta^2}.$$ 

We have:

$$\frac{d}{d\epsilon} Q_{3\epsilon} \to +\infty \quad \text{as } \epsilon \to 0^+,$$

as proved previously. On the other hand,

$$\frac{d}{d\epsilon} \tilde{R}_\epsilon(\pi) > 0.$$

The proof is the same as for $R_\epsilon$.

To prove (4.19) for $d_\epsilon^-, \tilde{A}^+(\epsilon), \tilde{B}^-(\epsilon)$, this argument should be extended. It is well known [MR] that the “true slow curve” $x_1 = 0$ ($x_2 = 0$) in the initial coordinates is an $\epsilon$-depending curve

$$x = s_\epsilon^+(y) \quad (x = s_\epsilon^-(y)), \quad y \in [-\beta, \beta], \quad \beta = \tau - \delta^2,$$ 

with

$$s_\epsilon^\pm(y) = s^\pm(y) + \epsilon s_1^\pm(y) + \ldots$$

The derivative $s_1^\pm(y)$ may be easily found from the assumption that (4.23) is an invariant curve of the system (1.3). Namely,

$$s_1^+(y) = \frac{s^+(y)}{\varphi(y, 0)}, \quad s_1^-(y) = \frac{s^-(y)}{\psi(y, 0)},$$

where $s^\pm, \varphi, \psi$ are the same as in 4.2. Hence, the difference $s_\epsilon^+(-\beta) - s_-^+(-\beta)$ is positive of order $\epsilon$. On the other hand, the maps $P_{\epsilon}^{0,-\beta}$ and $Q_{1\epsilon} = P_{\epsilon}^{0,\alpha-\beta}$ are contracting on $[0, \pi]$ with coefficient that tends to zero as $\exp(-C/\epsilon)$. This is proved for $Q_{1\epsilon} = Q_1$ in 4.1; the proof for $P_{\epsilon}^{0,-\beta}$ is the same. Hence, the points $Q_{1\epsilon}(0), P_{\epsilon}^{0,-\beta}(0)$ lie outside $M^-$, and $(P_{\epsilon}^{0,-\beta})'(0) \to 0$ as $\epsilon \to 0$. The monotonicity arguments above show that $\frac{d}{d\epsilon} Q_{2\epsilon} > 0$ outside $\Gamma_1 \cap M^-$. Hence,

$$\frac{d}{d\epsilon} R_\epsilon(0) = \left(\frac{d}{d\epsilon} Q_{2\epsilon}\right) \circ Q_{1\epsilon}(0) + Q_{2\epsilon}' \circ Q_{1\epsilon} \frac{d}{d\epsilon} Q_{1\epsilon}(0).$$

The first term is positive by the monotonicity arguments above; the second one tends to zero. Hence, $\frac{d}{d\epsilon} R_\epsilon(0)$ has a nonnegative lower limit as $\epsilon \to 0$. The same arguments as for $d_\epsilon^+$ prove (4.19) for $d_\epsilon^-$ and $\tilde{B}^-(\epsilon)$.

To prove (4.19) for $\tilde{A}^+(\epsilon)$, we must establish that $\frac{d}{d\epsilon} y(\tilde{d}_\epsilon^+) = \frac{d}{d\epsilon} P_{\epsilon}(\tilde{d}_\epsilon^+) has a nonpositive lower limit (in fact this limit is $-\infty$). This is done by using the same arguments as for $R_\epsilon$. This proves statement 1 of Lemma 2.
To prove statement 2, we must estimate the magnitude of $R_\varepsilon$. Between the sections $\Gamma^+ = \{ y = -\tau + \delta \}$ and $\Gamma_1 = \{ y = -\tau + \delta^2 \}$, the $x$ coordinate of the trajectories does not change by more than $2\pi$ because they are trapped by slow curves. The time of transition from $\Gamma_1$ to $\Gamma_2 = \{ y = -\tau - \delta^2 \}$ is $2\delta^2/\varepsilon$. Therefore, the magnitude of $R_\varepsilon$ is

$$R_\varepsilon(x) = O\left( \frac{\delta^2}{\varepsilon} \right) + O(1).$$

Combining this estimate with (4.20), we see that $\varepsilon_n$ satisfies the equation

$$-n\pi = \frac{C + O(\delta^2) + O(\varepsilon)}{\varepsilon}$$

with $C < 0$. The solution of this equation is of magnitude $O(1/n)$. This completes the proof of Lemma 2.

\[\square\]

References


Mathematics Department, Cornell University, Ithaca, NY 14853

Independent University of Moscow, B. Vlashevsy per. 11, Moscow 121002, Russia,
Mathematics Department, Cornell University, Ithaca, NY 14853,
Moscow State University, and
Steklov Mathematical Institute

E-mail address: yulijs@mccme.ru, yulij@math.cornell.edu