Almost invariant elliptic manifolds in a singularly perturbed Hamiltonian system

V. Gelfreich*, L. Lerman†

* Institut für Mathematik I, Freie Universität Berlin
Arnimallee 2-6, D-14195 Berlin, Germany
E-mail: gelf@math.fu-berlin.de

† Institute for Applied Math. & Cybernetics
10 Ul'yanov St., Nizhny Novgorod, 603005 Russia

April 4, 2001

Abstract

We consider a singularly perturbed Hamiltonian system, which loses one degree of freedom at $\varepsilon = 0$. Assume the slow manifold to be normally elliptic. In the case of an analytic Hamilton function it is shown that the slow manifold persists up to an exponentially small error term.

1 Introduction

The main object of this paper is a singularly perturbed Hamiltonian system

$$\begin{aligned} \varepsilon\dot{x} &= \frac{\partial H}{\partial y}, & \dot{u}_i &= \frac{\partial H}{\partial v_i}, \\ \varepsilon\dot{y} &= -\frac{\partial H}{\partial x}, & \dot{v}_i &= -\frac{\partial H}{\partial u_i}, \end{aligned} \qquad i = 1, \dots, n. \tag{1}$$

Here x and y are scalar variables, and $u, v \in \mathbb{R}^n$. Many model equations from applications can be transformed to this form. For example, see [AK89], where the equation was derived to describe water waves in the presence of surface tension; or [AEL98], where a system of two second order equations was used to describe soliton solutions to non-local sine-Gordon equation; or [BCKV93], where a so-called Hamiltonian elliptic Hopf bifurcation was studied.

If we formally put $\varepsilon = 0$, the system (1) is reduced to a pair of algebraic equations,

$$\frac{\partial H}{\partial y} = \frac{\partial H}{\partial x} = 0, \tag{2}$$

and a Hamiltonian system in (u, v) variables, which is obtained from (1) after exclusion of (x, y) by means of (2). One can say the system (1) loses one degree of freedom at $\varepsilon = 0$.

The system (2) defines the *slow* manifold. Here we assume that (2) can be solved at $\varepsilon = 0$ with respect to x and y,

$$x = f(u,v), \qquad y = g(u,v) \tag{3}$$

in some domain D_0 of variables (u,v). Our goal is to study the dynamics near the slow manifold. Here one of the main questions is: does the slow manifold persists for small ε as an invariant manifold or a manifold sufficiently close to invariant one, and if yes, then in which sense? The answer is positive provided the slow manifold is *normally hyperbolic.* In order to explain, what normal hyperbolicity means here, it is convenient to rewrite the system (1) using the fast time $\tau = \varepsilon^{-1}t$:

$$\begin{aligned} x' &= \frac{\partial H}{\partial y} & \qquad u' &= \varepsilon\frac{\partial H}{\partial v} \\ y' &= -\frac{\partial H}{\partial x} & \qquad v' &= -\varepsilon\frac{\partial H}{\partial u}. \end{aligned} \tag{4}$$

If we formally put $\varepsilon = 0$, then we obtain an one-degree-of-freedom Hamiltonian system in (x,y) variables (*fast* system) which parametrically depends on u and v. In this representation each point of the slow manifold corresponds to an equilibrium of the fast system. The slow manifold is normally hyperbolic if any of these equilibria is a saddle. In the whole $(2n+2)$-dimensional phase space of the system (4) we get the following picture: at $\varepsilon = 0$ the system has an invariant manifold (namely, the slow manifold) filled with equilibria, and this manifold is normally hyperbolic in the usual sense, see, for example, [HPSh77, Fen79]. The normal hyperbolicity implies that this submanifold persists for $\varepsilon > 0$ small enough being a symplectic (with respect to the 2-form $\mathrm{d}y \wedge \mathrm{d}x + \varepsilon^{-1}\sum_{i=1}^{n} \mathrm{d}v_i \wedge \mathrm{d}u_i$) invariant normally hyperbolic submanifold, and the restriction of the system (4) to this submanifold is an n-degrees-of-freedom Hamiltonian system with a Hamiltonian being the restriction of H to this submanifold.

In this paper we will study the opposite case, when each equilibrium is elliptic instead. In this case there is not rigorous persistence results for the slow manifold, as we are aware of. We assume the Hamiltonian $H = H(x,y,u,v;\varepsilon)$ to be analytic in all its arguments and show that there is a slow manifold being invariant one up to an error exponentially small in ε. This follows from the following theorem, which additionally describes dynamics in a neighborhood of the slow manifold.

Let us assume the functions f and g in (3) are analytic in a complex δ-neighborhood $D_0 + \mathrm{i}\delta$ of the real domain D_0. Consider the slow manifold

$$W_0^{\mathrm{slow}} = \{\, (x,y,u,v) : x = f(u,v),\ y = g(u,v),\ (u,v) \in D_0 \,\}$$

and denote

$$\Delta(u,v) = \det \begin{pmatrix} \partial^2_{xx}H & \partial^2_{xy}H \\ \partial^2_{yx}H & \partial^2_{yy}H \end{pmatrix}\Bigg|_{x=f(u,v),y=g(u,v),\varepsilon=0}.$$

Denote $D_1(\delta) = W_0^{\mathrm{slow}} + \mathrm{i}\delta$.

Theorem 1 *Let the Hamiltonian H be analytic in $D_1(\delta)$ for some $\delta > 0$. Assume also $\Delta(u,v) > C > 0$ for all $(u,v) \in D_0$ and some C. Then there is a canonical change of coordinates $\Phi : (x,y,u,v) \mapsto (X,Y,U,V)$ and constants $\tilde{\delta}, c > 0$ such that in the new coordinates the Hamiltonian H takes the form*

$$H \circ \Phi^{-1} = H_0(I,U,V;\varepsilon) + R(X,Y,U,V;\varepsilon), \tag{5}$$

where $I = \frac{1}{2}(X^2 + Y^2)$ and

$$R = O\left(\mathrm{e}^{-\frac{c}{\varepsilon}}\right). \tag{6}$$

The change of coordinates and the new Hamiltonian are analytic with respect to (X,Y,U,V) in $D_1(\delta/2)$, and bounded in ε.

In particular, the theorem implies that I is almost an integral of the system: $I' = O\left(\mathrm{e}^{-\frac{c}{\varepsilon}}\right)$. The equation $X = Y = 0$ defines an almost invariant (normally) elliptic slow manifold.

In particular, in the case $n = 1$, the system (1) has two degrees of freedom and in some neighborhood of the slow manifold it is exponentially close to an integrable system. This allows to use powerful methods of Hamiltonian perturbation theory to a further study of the system.

The rest of the paper is devoted to a proof of Theorem 1. In Sections 2 and 3 we prepare the necessary technical tools. Then we prove the theorem in two steps in Sections 4 and 5. The main goal is to straighten the slow manifold and separate the rapid motions by means of canonical transformations. This is a common approach in the theory of Hamiltonian systems. Of course, a complete separation of variables is not possible, and we will show that it can be done up to an exponentially small remainder. The system (4) formally has two fast variables. In the elliptic case a trajectory in a neighborhood of the slow manifold rapidly rotates, we will see that it is possible to introduce new coordinates in such a way that the phase of this rotation remains the only rapid variable. We use an averaging procedure similar to [Nei81, Nei84] to eliminate the dependence of the Hamiltonian on the fast phase up to an exponentially small remainder.

2 Canonical transformations

The system (4) is Hamiltonian with respect to the symplectic form

$$\Omega = \mathrm{d}y \wedge \mathrm{d}x + \varepsilon^{-1} \sum_{i=1}^{n} \mathrm{d}v_i \wedge \mathrm{d}u_i .$$

So the symplectic form Ω is not the standard one but reflects the singular nature of the perturbation instead. For this reason we remind the definition and some basic properties of canonical transformations related to our Ω. A transformation is called *canonical* one if it preserves the symplectic form or, equivalently, the form (4) of a Hamiltonian system with an arbitrary H. It is convenient to define a canonical transformation using a generating function $S = S(x,Y,u,V)$, so the transformation $(x,y,u,v) \mapsto (X,Y,U,V)$ is defined implicitly by the system

$$X = \frac{\partial S}{\partial Y}, \qquad y = \frac{\partial S}{\partial x}, \qquad U = \varepsilon \frac{\partial S}{\partial V}, \qquad v = \varepsilon \frac{\partial S}{\partial u}. \tag{7}$$

This systems (locally) defines a coordinate change provided in addition

$$\frac{\partial^2 S}{\partial x \partial Y}\frac{\partial^2 S}{\partial u \partial V}-\frac{\partial^2 S}{\partial x \partial V}\frac{\partial^2 S}{\partial u \partial Y}\neq 0.$$

The domain, where the transformation $(x,y,u,v)\mapsto(X,Y,U,V)$ is correctly defined, may depend on the parameter ε.

Later on we mainly use coordinate transformations with generating functions of the following special form

$$S = xY + \frac{uV}{\varepsilon} + \varepsilon s(x,Y,u,V), \tag{8}$$

which generate close-to-identity transformations

$$X = x+\varepsilon\frac{\partial s}{\partial Y},\quad y = Y+\varepsilon\frac{\partial s}{\partial x},\quad U = u+\varepsilon^2\frac{\partial s}{\partial V},\quad v = V+\varepsilon^2\frac{\partial s}{\partial u}. \tag{9}$$

Note, that despite the singular form of the generating function S the transformation is nonsingular at $\varepsilon=0$.

For an analytic function s given in a bounded domain G, and continuous in the closure $\bar{G}$ we denote $\|s\| = \sup_G |s|$. For a vector-valued function the norm equals to the maximal norm among its partial components.

Lemma 2 *Let* $0<\varepsilon<1$*,* $\delta>0$*,* $0<\xi<\delta$*,* $D\subset\mathbb{R}^{2n+2}$*. Let the function* s *be analytic in* $D+\mathrm{i}\delta$ *and*

$$\|s\|_\delta \le \frac{\xi^2}{8\varepsilon}. \tag{10}$$

Then (9) *defines a canonical transformation* $D+\mathrm{i}(\delta-\xi)\to D+\mathrm{i}(\delta-\xi/2)$ *of the form*

$$\begin{array}{rclcrcl} x &=& X+\varepsilon\alpha_1(X,Y,U,V;\varepsilon), && u &=& U+\varepsilon^2\beta_1(X,Y,U,V;\varepsilon),\\ y &=& Y+\varepsilon\alpha_2(X,Y,U,V;\varepsilon), && v &=& V+\varepsilon^2\beta_2(X,Y,U,V;\varepsilon), \end{array}$$

where

$$\|\alpha_k\|_{\delta-\xi},\ \|\beta_k\|_{\delta-\xi}\le\frac{2\|s\|_\delta}{\xi},\qquad k=1,2.$$

Proof. We prove the lemma using a convergent iteration scheme. It is identical to the usual proof of Implicit Function Theorem, and we repeat it only to check the quantitative estimates, which will be essential for the proof of the main theorem.

As it is easy to see from the first and third equations of (9)

$$\begin{array}{rcl} \alpha_1 &=& -\dfrac{\partial s}{\partial Y}(X+\varepsilon\alpha_1, Y, U+\varepsilon^2\beta_1, V;\varepsilon),\\ \beta_1 &=& -\dfrac{\partial s}{\partial V}(X+\varepsilon\alpha_1, Y, U+\varepsilon^2\beta_1, V;\varepsilon). \end{array}$$

We solve this system by iterations

$$\begin{array}{rcl} \alpha^{(n+1)} &=& -\dfrac{\partial s}{\partial Y}(X+\varepsilon\alpha^{(n)}, Y, U+\varepsilon^2\beta^{(n)}, V;\varepsilon),\\ \beta^{(n+1)} &=& -\dfrac{\partial s}{\partial V}(X+\varepsilon\alpha^{(n)}, Y, U+\varepsilon^2\beta^{(n)}, V;\varepsilon). \end{array}$$

starting with $\alpha^{(0)}, \beta^{(0)} \equiv 0$. The first iteration gives

$$\begin{aligned}
\alpha^{(1)} &= -\frac{\partial s}{\partial Y}(X, Y, U, V; \varepsilon), \\
\beta^{(1)} &= -\frac{\partial s}{\partial V}(X, Y, U, V; \varepsilon).
\end{aligned}$$

Using Cauchy estimates for the derivatives we obtain:

$$\|\alpha^{(1)}\|_{\delta-\xi}\,,\ \|\beta^{(1)}\|_{\delta-\xi} \le \frac{\|s\|_\delta}{\xi}.$$

We continue by induction. Obviously,

$$\begin{aligned}
\alpha^{(n+1)} - \alpha^{(n)} &= -\frac{\partial s}{\partial Y}(X + \varepsilon\alpha^{(n)}, Y, U + \varepsilon^2\beta^{(n)}, V; \varepsilon) \\
&\quad + \frac{\partial s}{\partial Y}(X + \varepsilon\alpha^{(n-1)}, Y, U + \varepsilon^2\beta^{(n-1)}, V; \varepsilon), \\
\beta^{(n+1)} - \beta^{(n)} &= -\frac{\partial s}{\partial U}(X + \varepsilon\alpha^{(n)}, Y, U + \varepsilon^2\beta^{(n)}, V; \varepsilon) \\
&\quad + \frac{\partial s}{\partial U}(X + \varepsilon\alpha^{(n-1)}, Y, U + \varepsilon^2\beta^{(n-1)}, V; \varepsilon).
\end{aligned}$$

Let us assume that

$$\varepsilon\|\alpha^{(i)}\|_{\delta-\xi},\ \varepsilon\|\beta^{(i)}\|_{\delta-\xi} \le \xi/2$$

for all $i \le n$. We estimate the differences using the maxima of the derivatives, and bound the derivatives by Cauchy estimates. We obtain

$$\|\alpha^{(n+1)} - \alpha^{(n)}\|_{\delta-\xi} \le \frac{4\|s\|_\delta}{\xi^2}\left(\varepsilon\|\alpha^{(n)} - \alpha^{(n-1)}\|_{\delta-\xi} + \varepsilon^2\|\beta^{(n)} - \beta^{(n-1)}\|_{\delta-\xi}\right).$$

The same upper bound holds for $\|\beta^{(n+1)} - \beta^{(n)}\|_{\delta-\xi}$. From these upper bounds we conclude inductively that

$$\|\alpha^{(n+1)} - \alpha^{(n)}\|_{\delta-\xi}\,,\ \|\beta^{(n+1)} - \beta^{(n)}\|_{\delta-\xi} \le a_0 q^n,$$

where

$$a_0 = \frac{\|s\|_\delta}{\xi} \le \frac{\xi}{8\varepsilon}, \qquad q = \frac{4\varepsilon\|s\|_\delta}{\xi^2}(1+\varepsilon) \le \frac{1}{2}.$$

In particular, $\varepsilon\|\alpha^{(n+1)}\|_{\delta-\xi}$, $\varepsilon\|\beta^{(n+1)}\|_{\delta-\xi} \le \xi/2$, which justifies the induction assumption.

Consequently, the sequence $\alpha^{(n)}$, $\beta^{(n)}$ converges uniformly in $D + \mathrm{i}(\delta - \xi)$ and the limit functions α_1 and β_1 are bounded from above by $2a_0$.

Finally the functions α_2 and β_2 can be bounded directly from the second and fourth equations of (9) since they are explicitly given after α_1 and β_1 have been found.

3 Solutions of the homological equation

In this section we study the homological equation

$$x\partial_y s - y\partial_x s = f(x,y), \tag{11}$$

which we have to solve at each step of the averaging procedure described in Section 5. Let us describe the suitable class of functions. Let f be an analytic function of two complex variables in a ball $B_R = \{\, |x|^2 + |y|^2 < R^2 \,\}$. We connect with f a function $\bar{f}$ of a single complex variable $I = \frac{1}{2}(x^2 + y^2)$. We call $\bar{f}$ the *average* of f. This $\bar{f}$ is defined as follows. We introduce symplectic polar coordinates in $B_R \setminus \{\, x^2 + y^2 = 0 \,\}$,

$$x = \sqrt{2I}\cos\varphi, \qquad y = \sqrt{2I}\sin\varphi, \tag{12}$$

where we use the main branch of the square root. In the new coordinates B_R is determined by the inequality

$$|x|^2 + |y|^2 = 2|I|\big(|\cos\varphi|^2 + |\sin\varphi|^2\big) = 2|I|\cosh(2\operatorname{Im}\varphi) < R^2.$$

The intersection of B_R with the level set $\{\, I = I_0 \,\}$, where $0 < |I_0| < \frac{1}{2}R^2$, is a cylinder $\{\, |\operatorname{Im}\varphi\,| < \rho \,\}$, where $\rho \equiv \rho(I_0) = \frac{1}{2}\cosh^{-1}(R^2/2|I_0|)$. This cylinder shrinks into a circle as $|I_0| \to \frac{1}{2}R^2$. For a fixed I from the punctured disk $0 < |I| < \frac{1}{2}R^2$ the function $f(x,y) = f(\sqrt{2I}\cos\varphi, \sqrt{2I}\sin\varphi)$ is analytic and 2π-periodic in φ in the strip $|\operatorname{Im}\varphi| < \rho(I)$ and continuous in its closure. We call its mean value

$$\bar{f}(I) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\sqrt{2I}\cos\varphi, \sqrt{2I}\sin\varphi)\, d\varphi \tag{13}$$

the average of the function f. It is an analytic function of I. The analyticity is evident for $I \neq 0$. In principle, $\bar{f}$ might have a branching at $I = 0$. But, expanding $f(x,y)$ in the power series in (x,y), inserting polar coordinates and integrating, we get a convergent power series in $\sqrt{2I}$ with coefficients being averages of the related trigonometric monomials. The average of an odd monomial vanishes, so in the power series only even degrees of $\sqrt{2I}$ remain. Thus, one gets a convergent power series in I, that is, the average is an analytic function at $I = 0$ as well. Thus, $\bar{f}$ is analytic in the disk $|I| < \frac{1}{2}R^2$.

Let $\mathcal{X}$ be the Banach space of all functions analytic in B_R, continuous in its closure, and having zero average. The space is equipped with supremum norm.

Lemma 3 *The operator $L = x\partial_y - y\partial_x$ has a continuous left inverse $L^{-1} : \mathcal{X} \to \mathcal{X}$, defined on all $\mathcal{X}$, and $\|L^{-1}\| \leq 2\pi$.*

Proof. Let $f \in \mathcal{X}$. We check that (11) has a unique solution $s \in \mathcal{X}$ and establish the following upper bound

$$\|s\|_{B_R} \leq 2\pi\, \|f\|_{B_R}. \tag{14}$$

In the polar coordinates (12) Equation (11) takes the following simple form

$$\partial_\varphi s = f(\sqrt{2I}\cos\varphi, \sqrt{2I}\sin\varphi). \tag{15}$$

This equation can be easily solved. Existence, uniqueness and the upper bound for the solution are consequences of the following lemma.

Lemma 4 *Let a 2π-periodic function f be real-analytic in a strip $\Pi_\rho = \{\, |\operatorname{Im} z| < \rho \,\}$ and continuous in its closure, and $\int_0^{2\pi} f(z)\,\mathrm{d}z = 0$. Then there is a unique primitive F, $F' = f$, with $\int_0^{2\pi} F(z)\,\mathrm{d}z = 0$. Moreover, F is real analytic in Π_ρ and*

$$\sup_{\Pi_\rho} |F| \le 2\pi \sup_{\Pi_\rho} |f|.$$

If f analytically depends on a complex parameter I at $I = I_0$, and $\rho = \rho(I)$ is continuous, then F is analytic in I at $I = I_0$.

Proof. Existence and uniqueness of the primitive with zero mean is obvious. We have to establish the upper bound only. It is remarkable that the upper bound does not depend on ρ. Actually, it is sufficient to estimate F on the boundary $\partial\Pi_\rho$, where it reaches its maximum value, due to analyticity and continuity. The restriction of F onto the boundary $\partial\Pi_\rho$ is a continuous complex function of the real variable t, if we let $z = t + \mathrm{i}\rho$. It is sufficient to consider only the upper component of the boundary since F is symmetric with respect to complex conjugation due to its reality.

Let $u(t)$ be a real 2π-periodic continuous function of a real variable t with zero mean value $\int_0^{2\pi} u(t)\,\mathrm{d}t = 0$. Then there is the unique primitive U, $\dot U = u$, with zero mean value. It is easy to see that $|U(t)| \le \pi\|u\|$, where $\|u\| := \sup|u|$. Now we consider a complex-valued function $g(t) = u(t) + \mathrm{i}v(t)$, where real functions $u(t)$ and $v(t)$ have the above-mentioned properties. Then g has a unique primitive G with zero mean, and

$$|G(t)|^2 = |U(t)|^2 + |V(t)|^2 \le \pi^2\left(\|u\|^2 + \|v\|^2\right) \le 2\pi^2\|g\|^2.$$

Finally, we apply the last estimate to the function $g(t) = f(t + \mathrm{i}\rho)$, and obtain

$$|F(t+\mathrm{i}\rho)| = |G(t)| \le 2\pi\|g\| = 2\pi\|f\|,$$

which implies the desired estimate.

Our problem is complicated by the fact that the complex polar coordinates are not well defined in the whole B_R. Lemma 4 provides a solution s of Equation (15), which is analytic in

$$\tilde B_R = \{\, (\varphi, I) : I \ne 0,\ |I|\cosh(2\operatorname{Im}\varphi) < \tfrac12 R^2 \,\}.$$

Moreover, it does not imply single-valuedness of s as a function of I. For this reason, we have to check analyticity of $s = s(x,y)$ in the original Cartesian coordinates. Consider the Taylor series

$$f(x,y) = \sum_{j,k} a_{j,k} x^j y^k,$$

and construct the Taylor series for s. We have

$$f(\sqrt{2I}\cos\varphi, \sqrt{2I}\sin\varphi) = \sum_{j,k=0}^{\infty} a_{j,k}(2I)^{(j+k)/2}\cos^j\varphi\sin^k\varphi$$

or equivalently

$$f = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} b_{lm} I^{l/2} \mathrm{e}^{\mathrm{i}m\varphi}, \tag{16}$$

where

$$b_{lm} = \sum_{k=0}^{l} \frac{a_{l-k,k}}{2^l \mathrm{i}^k} \sum_{j_1+k_1=\frac{m+l}{2}} \left(\begin{array}{c} l-k \\ j_1 \end{array} \right) \left(\begin{array}{c} k \\ k_1 \end{array} \right).$$

It is easy to see that $b_{lm} = 0$ for all odd $m+l$. The assumption $\bar{f} = 0$ is equivalent to $b_{l0} = 0$ for all l. And the only solution of the equation $\partial_\varphi s = f$ with zero average has the form:

$$s = \sum_{l=1}^{\infty} \sum_{m=-l,\, m\neq 0}^{l} \frac{b_{lm}}{m} I^{l/2} \mathrm{e}^{\mathrm{i}m\varphi}. \tag{17}$$

At inner points of $\tilde{B}_R$ this series obviously converges (since the series (16) is convergent in $\tilde{B}_R$), and we have only to check that they define an analytic function at $(x, y) = (0, 0)$ in spite of the apparent presence of the non-integer powers of I. Indeed, we have

$$\mathrm{e}^{\mathrm{i}\varphi} = \frac{x + \mathrm{i}y}{\sqrt{2I}}, \qquad \mathrm{e}^{-\mathrm{i}\varphi} = \frac{x - \mathrm{i}y}{\sqrt{2I}},$$

for all $(x, y) \in B_R$ with $x^2 + y^2 \neq 0$. Thus

$$s = \sum_{l=1}^{\infty} \sum_{m=1}^{l} \frac{(x^2 + y^2)^{(l-m)/2}}{2^{l/2} m} \Big(b_{l,m}(x + \mathrm{i}y)^m - b_{l,-m}(x - \mathrm{i}y)^m \Big),$$

where all b_{lm} with odd $l - m$ vanish. Consequently, the solution s defined by (17) is single-valued and analytic outside of the set $x^2 + y^2 = 0$. In fact, this function has a unique analytic continuation on the planes $x = \pm \mathrm{i}y$. Indeed, according to the estimate of Lemma 4 s is bounded near the set K being the union of two disks $x = \pm \mathrm{i}y$. Furthermore, set K is fine (see, [GR65]), that is, it is locally represented as a zero level of some analytic function, here $I = x^2 + y^2$. So, by the theorem Osgood-Brown on the erasing singularities, this function is uniquely extended onto the set K as the analytic function.

Since the solution is unique this finishes the proof of Lemma 3.

4 Straightening of the slow manifold

Now we are ready to prove Theorem 1. As a first step we straighten the slow manifold of the Hamiltonian system (4) using a change of coordinates defined by the generating function (8) with

$$\varepsilon s = x g(u, V) - Y f(u, V).$$

Here f and g define the slow manifold by (3). Then in the new variables the Hamiltonian takes the form

$$H = H_0(x, y, u, v) + \varepsilon R_0(x, y, u, v; \varepsilon),$$

where

$$\left.\frac{\partial H_0}{\partial x}\right|_{x=y=0} = \left.\frac{\partial H_0}{\partial y}\right|_{x=y=0} = 0.$$

As it is usual we denote the new variables by the same letters as the old ones after the change of coordinates. We hope this will not lead to a confusion. In the new coordinates the slow manifold is given by $x = y = 0$.

Let us fix u and v for a moment and make a coordinate change $(x, y) \mapsto (X, Y)$ defined by a generating function $s = s(x, Y, u, v, \varepsilon)$. Then the generating function

$$S = \frac{uV}{\varepsilon} + s(x, Y, u, V, \varepsilon)$$

defines due to (7) a symplectic change of coordinates $(x, y, u, v) \mapsto (X, Y, U, V)$, which is close to identical transformation in u and v. Moreover, the transformations defined by s and S coincide on the (x, y) component.

Let us fix u and v and transform the Hamiltonian H_0 to its angle-action variables near $x = y = 0$. This is always possible since the corresponding system has one degree of freedom only. Then we extend this change of variables to the $(2n + 2)$-dimensional space. Since it is ε-close to identity in u and v, we have:

$$H(x, y, u, v; \varepsilon) = h_0(I, u, v) + \varepsilon r_0(x, y, u, v; \varepsilon), \tag{18}$$

where $I = \frac{1}{2}(x^2 + y^2)$. It is easy to see that $I' = O(\varepsilon)$, thus we found a hidden slow variable in our system.

5 Averaging method

Now we try to decrease the remainder term r_0 in (18) by a sequence of close-to-identity canonical transformations $\Phi_n : \tilde{D}_{\delta_{n-1}} \to \tilde{D}_{\delta_n}$, where

$$\tilde{D}_{\delta_n} = B_{\delta_n} \times (D_0 + \mathrm{i}\delta_n).$$

The sequence of $\tilde{D}_{\delta_n} \subset \mathbb{C}^{2n+2}$ is strictly decreasing. After the n^{th} step the Hamiltonian $H = H \circ \Phi_1 \circ \ldots \circ \Phi_n$ takes the form

$$H = h_n(I, u, v; \varepsilon) + \varepsilon r_n(x, y, u, v; \varepsilon), \qquad I = \tfrac{1}{2}(x^2 + y^2). \tag{19}$$

First we study a single step and then continue recurrently. On the n^{th} step we use a coordinate change defined by a generating function of the form (8), where we define s to be the unique (see Lemma 3) solution of the equation

$$y\partial_x s - x\partial_y s = \frac{r_n - \overline{r}_n}{\partial_I h_n} \tag{20}$$

with $\overline{s} = 0$, where the bar denotes an average of a function defined by (13). Let

$$\begin{aligned} h_{n+1} &= h_n + \varepsilon\overline{r}_n \ , && (21)\\ \varepsilon r_{n+1} &= H - h_{n+1} \ . && (22) \end{aligned}$$

Then the Hamiltonian H takes the form (19) with n replaced by $n+1$.

This procedure is similar to the classical averaging [BM61]. The classical estimates give upper bounds $r_n = O(\varepsilon^n)$ and $h_n = h_{n-1} + O(\varepsilon^n)$. These estimates are not uniform in n. It is usually possible to make more averaging steps for smaller ε. But the classical approach does not give indications on the optimal choice for $n = n(\varepsilon)$. Technically the upper bound on the maximum number of steps comes from the reduction of the domain on each step, which is necessary to apply Cauchy estimates for derivatives.

In a similar situation Neishtadt observed that it is possible to make up to $\sim \varepsilon^{-1}$ averaging steps. We will use Neishtadt-type estimates [Nei81, Nei84] to show that

$$\|r_{n+1}\|_{\delta_{n+1}} \le \|r_1\|_{\delta_1} \mathrm{e}^{-n}$$

for all $n < c\varepsilon^{-1}$. Thus after $n = [c\varepsilon^{-1}]$ steps the remainder is exponentially small in ε. Here we denote by $\|\cdot\|_\delta$ the supremum norm of a function on $\tilde{D}_\delta$.

Both approaches use the same changes of coordinates. The only difference between the classical and Neishtadt approaches is related to the choice of δ_n. In both cases the domain reduction $\xi = \delta_n - \delta_{n-1}$ is independent from n. The classical choice for ξ is a small constant, and Neishtadt used $\xi = A\varepsilon$ with sufficiently large constant A. In this way Neishtadt showed that a larger number of averaging steps is possible before the domain of the change becomes empty.

Proposition 5 (classical-type estimates) *Let $\xi \in (0, \delta_0)$ and $\delta_n = \delta_0 - n\xi$, $1 \le n \le \delta_0/(2\xi)$. There is $\varepsilon_0 > 0$ and $C_1, C_2 > 0$ such that for $0 < \varepsilon < \varepsilon_0$ the generating functions (20) define a sequence of canonical changes of coordinates $\Phi_n : \tilde{D}_{\delta_n} \to \tilde{D}_{\delta_{n-1}}$, such that $H \circ \Phi_1 \circ \ldots \circ \Phi_n$ has the form (19) with*

$$\|h_n - h_0\|_{\delta_n} \le C_1\varepsilon, \qquad \|r_n\| \le C_2\varepsilon^n.$$

It is convenient to make a single averaging step before making "exponential" averaging. Namely we may assume $r_0 = O(\varepsilon)$.

Proposition 6 (Neishtadt-type estimates) *Let $\|r_0\|_{\delta_0} \le C_0\varepsilon$. There is $\varepsilon_0 > 0$, and a sufficiently large positive constant A, such that for $0 < \varepsilon < \varepsilon_0$, $\delta_n = \delta_0 - nA\varepsilon$, $1 \le n \le \delta_0/(2A\varepsilon)$ the generating functions (20) define a sequence of canonical changes of coordinates $\Phi_n : \tilde{D}_{\delta_n} \to \tilde{D}_{\delta_{n-1}}$, such that $H \circ \Phi_1 \circ \ldots \circ \Phi_n$ has the form (19) with*

$$\|h_n - h_0\|_{\delta_n} \le C_1\varepsilon^2, \qquad \|r_n\| \le \|r_0\| \,\mathrm{e}^{-n}.$$

The last proposition directly implies Theorem 1.

Proof of the propositions. The proof is done inductively. According to the condition of normal ellipticity $\partial_I h_0|_{I=0}$ is positive and uniformly bounded away from zero. Then by Lemma 3 Equation (20) uniquely defines the generating function s with $\overline{s} = 0$, and

$$\|s\| \le C_3\|r_n\|.$$

Then Lemma 2 says that the corresponding change of coordinates is well-defined and

$$|x-X|,\ |y-Y| \le \frac{2\varepsilon\|s\|_{\delta_n}}{\xi} \le \frac{\xi}{2} \quad \text{and} \quad |u-U|,\ |v-V| \le \frac{2\varepsilon^2\|s\|_{\delta_n}}{\xi} \le \frac{\varepsilon\xi}{2},$$

provided ϵ_0 is sufficiently small.

We estimate the new Hamiltonian. The upper bound for h_{n+1} is straightforward. Let us estimate r_{n+1}. From (22) and (19) we have

$$\begin{aligned} \varepsilon r_{n+1}(X,Y,U,V) &= h_n(\tfrac{1}{2}(x^2+y^2),u,v) - h_n(\tfrac{1}{2}(X^2+Y^2),U,V) \\ &\quad +\varepsilon r_n(x,y,u,v) - \varepsilon\overline{r}_n(X,Y,U,V), \end{aligned}$$

where we suppressed the explicit dependence on ε. It is easy to see that

$$\begin{aligned} h_n(\tfrac{1}{2}(x^2+y^2),u,v) &= h_n + Y(\partial_I h_n)\varepsilon\partial_x s + E_1, \\ h_n(\tfrac{1}{2}(X^2+Y^2),U,V) &= h_n + x(\partial_I h_n)\varepsilon\partial_Y s + E_2, \\ r_n(x,y,u,v) &= r_n + E_3, \\ \overline{r}_n(X,Y,U,V) &= \overline{r}_n + E_4, \end{aligned}$$

where all functions in the right hand sides depend on the "mixed" variables (x,Y,u,V). These equalities may be considered as definitions for E_k, $k=1,\dots,4$. Taking into account (20) we obtain

$$\varepsilon r_{n+1}(X,Y,U,V) = E_1 - E_2 + \varepsilon E_3 - \varepsilon E_4.$$

Using Cauchy estimates we obtain

$$\begin{aligned} \|E_1\|_{\delta_{n+1}},\ \|E_2\|_{\delta_{n+1}} &\le \varepsilon^2\|h_n''\|_{\delta_n-\xi/2}\frac{\|s\|_{\delta_n}^2}{\xi^2} + \varepsilon^2\|h_n'\|_{\delta_n-\xi/2}\frac{\|s\|_{\delta_n}}{\xi}, \\ \|E_3\|,\ \|E_4\| &\le \varepsilon\,\frac{2\|r_n\|_{\delta_n}\|s\|_{\delta_n}}{\xi^2}. \end{aligned}$$

We check the uniform boundedness of the derivatives of h_n using Cauchy estimates and obtain

$$\|r_{n+1}\|_{\delta_{n+1}} \le \frac{\varepsilon C_3}{\xi}\,\|r_n\|_{\delta_n}.$$

The proposition follows immediately from this estimate and the corresponding choices of ξ and ε_0.

Acknowledgments

The authors thank D. Turaev for his helpful remarks. This research was partially supported by the INTAS grant 00-221. L. Lerman is thankful for a partial support to the Russian Foundation of Basic Research (grant 01-01-00905) and the program "Universities of Russia" (grant N 1905). V.G. work was also supported by the Deutsche Forschungsgemeinschaft.

References

[AK89] C.J. Amick, K. Kirchgässner. *A theory of solitary water-waves in the presence of surface tension.* Arch. Ration. Mech. Anal. 105, No.1, 1-49 (1989).

[AEL98] G.L. Alfimov, V.M. Eleonsky, L.M. Lerman. *Solitary wave solutions of nonlocal sine-Gordon equations.* Chaos 8, No.1, 257-271 (1998)

[BM61] N.N. Bogolyubov, Yu.A. Mitropolskij, *Asymptotic methods in the theory of non-linear oscillations*, New York: Gordon and Breach Science Publishers. V, 537 p. (1961).

[BCKV93] H.W. Broer, S.N. Chow, Y. Kim, G. Vegter. *A normally elliptic Hamiltonian Hopf bifurcation.* Z. angew. Math. Phys. 44 (1993), 389-432.

[Fen79] N. Fenichel. *Geometric singular perturbation theory for ordinary differential equations.* J. Diff. Equa., **31** (1979), pp.53-98.

[GLS01] V. Gelfreich, L. Lerman, J. Scheurle. *Exponentially small separatrix splitting in a singularly perturbed Hamiltonian system*, preprint (2001)

[GR65] R.C. Gunning, H. Rossi, Analytic functions of Several Complex Variables, Prentice-Hall, Inc., Englwood Cliffs, N.J., 1965.

[HPSh77] M. Hirsh, C. Pugh, M. Shub. *Invariant manifolds.* Lecture Notes in Math., 583, Springer-Verlag, New York. 1977.

[Nei81] A.I. Neishtadt. *On the accuracy of conservation of adiabatic invariant*, J. Appl. Math. Mech., vol. 45, no. 1 (1982) pp. 58–63 (translated from PMM 45, no. 1 (1981) pp. 1016–1025 (Russian))

[Nei84] A.I. Neishtadt. *The separation of motions in systems with rapidly rotating phase.* J. Appl. Math. Mech. 48 (1984), no. 2, 133–139 (1985); translated from Prikl. Mat. Mekh. 48 (1984), no. 2, 197–204 (Russian)