We study perturbations of Hamiltonian systems of $n + 1$ degrees of freedom ($n \geq 2$) in the real-analytic case, such that in the absence of the perturbation they contain a partially hyperbolic (whiskered) $n$-torus with the Kronecker flow on it with a Diophantine frequency, connected to itself by a homoclinic exact Lagrangian submanifold (separatrix), formed by the coinciding unstable and stable manifolds (whiskers) of the torus. Typically, a perturbation causes the separatrix to split. We study this phenomenon as an application of the version of the KAM theorem, proved in [13]. The theorem yields the representations of global perturbed separatrices as exact Lagrangian submanifolds in the phase space. This approach naturally leads to a geometrically meaningful definition of the splitting distance, as the gradient of a scalar function on a subset of the configuration space, which satisfies a first order linear homogeneous PDE. Once this fact has been established, we adopt a simple analytic argument, developed in [15] in order to put the corresponding vector field into a normal form, convenient for further analysis of the splitting distance. As a consequence, we argue that in the systems, which are Normal forms near simple resonances for the perturbations of integrable systems in the action-angle variables, the splitting is exponentially small.

1. Introduction

This note deals with real-analytic exact Hamiltonian perturbations of real-analytic exact Hamiltonian systems only, which will henceforth pass as “systems”. In particular, all the functions, maps, and vector fields, that one encounters along the way will be real-analytic, unless specified the contrary. Discussing the “splitting problem”, we restrict ourselves solely to the case when in the unperturbed setting an $(n + 1)$-dimensional separatrix connects to itself an invariant $n$-torus.

The phenomenon of splitting of separatrices was discovered by Poincaré [11] and used by Arnold [1] in order to suggest a mechanism for global topological instability in near-integrable Hamiltonian systems in the action-angle variables with several degrees of freedom. The work of Chirikov [2] laid down the fundamental concepts of the subject, which the author baptized as “Arnold’s diffusion”. Its recent studies [3], [6], [5], [14], [4] (and others, referenced therein) were based on analyzing time-evolution of the trajectories on the separatrices (whiskers). However, solving the initial value problem in Hamiltonian perturbation theory has always been rather challenging, because in the phase space of the systems of interest, the trajectories with principally different alpha and omega destinations are jumbled in a very intricate way, due to small divisors. Historically, the new geometric methods, such as the KAM (Kolmogorov–Arnold–Moser) theory, would benefit by avoiding it. Much earlier, Poincaré [11] emphasized the advantages of the Hamilton–Jacobi formalism versus trying to integrate the equations of motion directly.

The Hamilton–Jacobi formalism was used to study the splitting phenomenon in the recent work of Sauzin [15]. Apropos of the “splitting distance function”, giving the quantitative measure of the extent of the damage, the author made the following important observations:

Mathematics Subject Classification 34C15, 34C20, 70H05, 26B05, 26B40
This research was partially supported by NSF Grant DMS-9704750
S1: The splitting distance function is a gradient of a scalar function (hereafter the “splitting potential”).

S2: The splitting potential satisfies a first-order linear homogeneous PDE (which is interpreted as the “characteristic vector field”).

From S2 one quickly gets optimal upper bounds on the splitting distance. But the essential feature of the model, studied in [15] was that the perturbation vanishes on the torus to the second order. This made it rather easy to arrive at the following fact, that appears to be the key statement implying S1 and S2 above.

S3: The whiskers can be represented as exact Lagrangian submanifolds in the original coordinates.

Without the above mentioned assumption on the perturbation, [15] might need a specially adapted version of the KAM theorem, yet unavailable.

The KAM theorem has had a “whiskered” version since Graff [7] (see also [18], [16]). But the latter formulation is rather general: dealing with any number of hyperbolic degrees of freedom, it requires localization in the “hyperbolic variables”, thus becoming not the sharpest tool in order to study the splitting, which is a global phenomenon, and so far has been explored exclusively in the near-integrable case of only one hyperbolic degree of freedom. Technically, in order to put the system into a suitable normal form, satisfying the conditions of Graff’s theorem, one would routinely apply a lemma due to Moser [10]. All the above quoted recent works on the splitting would begin this way for which we will use a “dummy” notation for the standard symplectic charts defined on the configuration space being a finite cylinder C = J × T^n. Any symplectic chart (X, Y) = (x, y, y), on M, such that X = (x, x) ∈ J × T^n, and Y = (y, y) ∈ B^{n+1}, with the canonical symplectic structure Ω_0 = dX ∧ dY = dx ∧ dy + dx ∧ dy will be referred to as a “standard symplectic chart” on M. Suppose, J = J^- ∪ J^+, where J^± are two open intervals with a nonempty intersection J_0, with e.g. π ∈ J, and 0 ∈ J^-, 2π ∉ J^-, 2π ∈ J^+, 0 ∉ J^+. Let also M_± = J_± × T^n × B^{n+1}, M = J × T^n × B^{n+1}. The indexing ± should read either “−” and “+”, or “−” or “+”, which shall be clear from the context. We will also use the notations (X^±, Y^±) = (x^±, x^±, y^±, y^±), (X, Y) = (x^±, y^±, y^±), for the standard symplectic charts defined on the M_± or M, or sometimes in order to avoid repeating ourselves, use (X, Y) = (x, x, y, y) as a set of “dummy variables”, defined on either of these manifolds, for which we will use a “dummy” notation M, and accordingly the notation C for the corresponding configuration space. Let (X, Y) = (x, x, y, y) be a fixed standard symplectic chart on M, which may be further restricted to M_± or M.

Consider a perturbation H_µ = H_0 + O(µ) of the following Hamiltonian H_0 on M, parametrized by a constant c_0, a function f_{10}(x) on J, and a symmetric (n + 1) × (n + 1) matrix-valued function Q_0(X) on C, namely

\[ H_0(X, Y) = c_0 + f_{10}(x)y + \langle \bar{\omega}, \dot{y} \rangle + \frac{1}{2} (Y, Q_0(X)Y) + O_3(X, Y), \] (1.1)
ON A HOMOCLINIC SPLITTING PROBLEM

\[ Q_0(X) = \begin{pmatrix} f_{20}(X) & f_{21}(X) \\ f_{20}(X) & F_{20}(X) \end{pmatrix}, \]

where in particular \( F_{20} \) is an \( n \times n \) matrix-valued function on \( \mathcal{C} \).

Henceforth for \( m = 2, 3, \ldots, \) we will use the notations \( O_m(X, Y), \) \( O_m(X, Y, \mu) \) for the part of the Taylor series in \( Y \) at zero, embracing terms of the \( m \)-th and higher powers of \( Y \). We denote \( N^\tau_0(c, f_1, Q) \) a family of real-analytic Hamiltonian functions on either \( \mathcal{M}^\tau \), affording a representation, similar to (1.1), with some parameter triple \((c, f_1, Q)\) serving respectively for \((\alpha_0, f_{10}, Q_0)\).

Suppose, the frequency vector \( \bar{\omega} \), fixed throughout this paper is Diophantine \((\gamma, \tau)\), namely

\[ \forall \bar{k} \in \mathbb{Z}^n \backslash \{0\}, \quad |\langle \bar{\omega}, \bar{k} \rangle| = \gamma |\bar{k}|^{-\tau}, \quad \gamma > 0, \quad |k| = \sum_{i=1}^{n} |k_i|, \quad \tau \geq n - 1. \]  

Also suppose, the function \( f_{10}(x) \) on \( \mathcal{I} \) has simple zeroes at \( x = 0, 2\pi \), and is strictly nonzero elsewhere, being in addition odd, as well as \( f_{10}(2\pi - x) \). Suppose, \( f_{10e}(0) > 0, f_{10e}(2\pi) < 0 \). Let the matrix function \( Q_0(X) \) on \( \mathcal{I} \times T^n \) satisfy the non-degeneracy condition: \( \det F_{200}(x)|_{x=0,2\pi} \neq 0 \). Hereafter the subscript 0 stands for the average in the angle variables \( \bar{x} \).

For \( \mu = 0 \) we have \( H_{\mu} = H_0 \), and in the phase space there is a pair of exact Lagrangian submanifolds \( \mathcal{W}_0^\tau \), defined in particular respectively on \( \mathcal{M}^\tau \), and given by \( Y = 0 \). These manifolds asymptote to the invariant partially hyperbolic tori \( \mathcal{T}_0^\tau \), where also \( x = 0 \), in negative time, and \( \mathcal{T}_0^\tau \), where also \( x = 2\pi \), in positive time. The manifolds \( \mathcal{W}_0^\tau \) coincide in particular in \( \mathcal{M} \), their intersection there denoted as \( \mathcal{W}_0 \). For \( \mu \) small, but nonzero, we will still have a pair of exact Lagrangian submanifolds \( \mathcal{W}_\tau \), living respectively in \( \mathcal{M}^\tau \), but they will generally no longer coincide in \( \mathcal{M} \), in other words, \( \mathcal{W}_0 \) splits.

The main theorem which we will use to study the splitting problem is the KAM theorem from [13].

This theorem applies to either whisker, so we will keep using the indexing \( \tau \), whenever possible. Suppose, for the moment, that a pair of standard symplectic charts \((X^\tau, Y^\tau)\) coincides with \((X, Y)\) on \( \mathcal{M}^\tau \) respectively. Denote

\[ \mathcal{W}^\tau \overset{\text{def}}{=} \{(X^\tau, Y^\tau) \in \mathcal{M}^\tau | Y^\tau = 0 \}, \]

\[ \mathcal{T}^- \overset{\text{def}}{=} \{(X^-, Y^-) \in \mathcal{M}^- | Y^- = 0, x^- = 0 \}, \]

\[ \mathcal{T}^+ \overset{\text{def}}{=} \{(X^+, Y^+) \in \mathcal{M}^+ | Y^+ = 0, x^+ = 2\pi \}. \]

Trivially, there exists a pair of exact symplectic diffeomorphisms of \( \mathcal{M}^\tau \), both identity, denoted \( \Xi_0 \), acting from \((X^\tau, Y^\tau)\) to \((X, Y)\), that is \( \mathcal{W}_0^\tau = \Xi_0 \mathcal{W}^\tau \), and \( \mathcal{T}_0^\tau = \Xi_0 \mathcal{T}^\tau \).

Recall that the notation \( \mathcal{M} \) (\( \mathcal{C} \)) may stand for either \( \mathcal{M}^\tau \), \( \mathcal{C}^\tau \), \( \mathcal{E}^\tau \) with the proper standard symplectic chart \((X, Y)\). Let \( \mathcal{A} \) be a diffeomorphism of \( \mathcal{C} \), and \( S \) be a scalar function on \( \mathcal{C} \), such that it admits the representation \( S(X) = (\xi, \bar{x}) + B(x, \bar{x}) \), with a constant \( \xi \), and a function \( B \), which is quasiperiodic in \( \bar{x} \). Let \( \mathcal{E}_\mathcal{M} = \mathcal{E}_\mathcal{M}(A, S) \) be the subgroup of the group of exact symplectic diffeomorphisms of \( \mathcal{M} \), parametrized as follows: for any \( \Xi \in \mathcal{E}_\mathcal{M} \)

\[ \Xi = \Xi(A, S): \begin{cases} X = A(X), \\ Y = (dA)^{-T}(Y + \partial_X S(X)). \end{cases} \]

The notation \( (dA)^{-T} \) will henceforth read “\( dA \) inverse, transposed”. The map \( \Xi \) is exact with the generating function \( G(X, Y) = S[A^{-1}(X)] + (A^{-1}(X), Y) \).

Let \( \mathcal{L} \) be an exact Lagrangian submanifold, represented as \((X, \partial_X \mathcal{L}(X))\) with the generating function \( S_{\mathcal{L}}(X) \). Then, under a transformation \( \Xi(A, S) \in \mathcal{E}_\mathcal{M} \), the generating function \( S_{\mathcal{L}} \) transforms into \( S_{\mathcal{L}'} = (S_{\mathcal{L}} + S) \circ A^{-1} \), namely \( \mathcal{L} \) can be represented as \((X, \partial_X \mathcal{L}'(X))\) in the chart \((X, Y)\).

If \( \mu \neq 0 \), the main theorem in [13] would say, that there exists a pair of standard symplectic charts \((X^\tau, Y^\tau)\) over \( \mathcal{M}^\tau \), a pair of symplectic near-identity diffeomorphisms \( \Xi_\tau \) of \( \mathcal{M}^\tau \), and a pair

---

REGULAR AND CHAOTIC DYNAMICS, V. 5, № 2, 2000
of triples \((c_\tau, f_1, Q_\tau)\) that differ from \((c_0, f_{10}, Q_0)\) by \(O(\mu)\), such that \(H_\mu \circ \Xi_\tau \in N_{\tau}^\tau(c_\tau, f_1, Q_\tau)\). Besides, the function \(f_1(x)\) is odd, as well as the function \(f_1(2\pi - x)\). Throughout the rest of the paper, the dependence on \(\mu\), if not specified explicitly, will be implied in the subscripts \(\tau\) or the “hat” notation for the functions, maps, and vector fields.

The transformations \(\Xi_\tau\) are described each by the proper pair of parameters \((A_\tau, S_\tau)\) and belong to the subgroups \(\mathcal{E}_{\mu, \tau}\) of the groups of exact symplectic diffeomorphisms of \(\mathcal{M}^\tau\), described above. The diffeomorphisms \(A_\tau = A_\tau\) are near-identity, the functions \(S_\tau\) admit the representations \(S_\tau(x^\tau, x^\tau) = \langle \xi^\tau, x^\tau \rangle + B_\tau(x^\tau, x^\tau), \) where \(\xi^\tau\) and the quasiperiodic in \(x^\tau\) functions \(B_\tau\) all vanish at \(\mu = 0\).

Then \(\mathcal{F}_\mu = \Xi_\tau \mathcal{F}^\tau\) is a pair of invariant tori, the flow whereupon is conjugated (by \(\Xi_\tau\)) to the Krylov-Kolmogorov flow with the frequency \(\omega\). Each of these tori is contained inside a corresponding exact Lagrangian submanifold \(\mathcal{W}_\mu = \Xi_\tau \mathcal{W}^\tau\) with the generating function \(\mathcal{F}_\tau = S_\tau \circ A_\tau^{-1}\), yielding in the original coordinates \((X, Y)\) the representation \(\mathcal{W}_\mu = \{(X, Y) \in \mathcal{M}^\tau | Y = \partial_X \mathcal{F}_\tau(X)\}, \) with
\[
\mathcal{F}_\tau(X) = \langle \xi^\tau, x^\tau \rangle + B_\tau(x^\tau, x^\tau).
\]
The constants \(\xi^\tau\) are the same as above, and the functions \(B_\tau(x, x)\) are such that they are quasiperiodic in \(x\) and vanish at \(\mu = 0\).

So far we have not really taken advantage of dealing with the homoclinic problem. This is taken care of by adding an extra condition (2.5) on the Hamiltonian \(H_\mu\), allowing to view the tori \(\mathcal{F}_\mu\) as a single geometric object. But all the argument given so far can be extended (see (2.6) below) to a slight generalization of (1.1), when instead of the term \(\langle \omega, y \rangle\) in one has \(\langle f_1(x), y \rangle\), with the function \(f_1\) on \(\mathcal{F}\), such that \(f_1(0) = \tilde{\omega}_-\), \(f_1(2\pi) = \tilde{\omega}_+\), both Diophantine, thus leading to a heteroclinic splitting problem.

As we mentioned earlier, knowing the generating function for a Lagrangian submanifold \(\mathcal{L}\) in \(\mathcal{M}\), in one standard symplectic chart, one can rewrite it in another, obtained from the former by applying a transformation \(\Xi(A, S) \in \mathcal{E}_M\). Hence, we introduce the “splitting potential” \(\mathcal{F}\) on \(\hat{\mathcal{C}}\) in the original coordinates \((X, Y)\) as
\[
\mathcal{F} \overset{\text{def}}{=} \mathcal{F}_- - \mathcal{F}_+.
\]
The functions \(\mathcal{F}_\tau(X)\) satisfy the Hamilton-Jacobi equations on \(\mathcal{C}^\tau\) (one for \(\tau-\) and the other for \(\tau+\)):
\[
H_\mu(X, \partial_X \mathcal{F}_\tau(X)) = 0.
\]
If we subtract the latter equation from the former, and recall that \(H_\mu(X, Y)\) ist a Taylor series in \(Y\), and that \(\mathcal{F}_\tau = O(\mu)\), we shall obtain the following first order linear PDE on \(\hat{\mathcal{C}}\):
\[
\hat{D} \mathcal{F} = c_- - c_+, \quad \text{with} \quad \hat{D} = (f_{10} + O(\mu)) \partial_x + \langle \tilde{\omega} + O(\mu), \partial_x \rangle,
\]
being a perturbation of
\[
D_0 \overset{\text{def}}{=} f_{10} \partial_x + \langle \tilde{\omega}, \partial_x \rangle.
\]
Most importantly, in (1.8) all the momenta-independent terms, that the perturbation \(H_\mu\) of \(H_0\) may have, are gone. Of further interest is only the case when \(c_- = c_+\), namely when the perturbed invariant tori \(\mathcal{F}_\mu\) (and their whiskers) lie on the same energy surface, otherwise the whiskers will never intersect. In the homoclinic case \(c_-\) always equals \(c_+\); this can also be insured in the heteroclinic case by imposing an extra condition, generalizing the following homoclinic condition (2.5) on the full Hamiltonian \(H_\mu\). Then the equation (1.8) can be interpreted as the existence of the “characteristic vector field” on \(\hat{\mathcal{C}}\), orthogonal to \(\partial_X \hat{\mathcal{F}}(X)\). This essentially proves S2 (S1 and S3 having been shown already).
Merely due to the fact that on $\hat{\mathcal{F}}$ the function $f_{10}(x)$ is bounded away from zero, by a simple application of the Implicit function theorem, fulfilled in [15], one can conjugate $\hat{D}$ and $D_0$ by a near-identity diffeomorphism $\hat{A}: \hat{X} \to X$ of $\hat{\mathcal{C}}$. With the vector field notation

$$V_0 \overset{\text{def}}{=} (f_{10}, \bar{\omega}), \quad \hat{\mathcal{F}} \overset{\text{def}}{=} (f_{10} + O(\mu), \bar{\omega} + O(\mu)) \equiv V_0 + V_1,$$

it means

$$(d\hat{A})^{-1} \hat{\mathcal{F}} \circ \hat{A} = V_0. \quad (1.11)$$

This implies that the function $\hat{\mathcal{F}}_s = \hat{\mathcal{F}} \circ \hat{A}$ satisfies the equation on $\hat{\mathcal{C}}$

$$D_0 \hat{\mathcal{F}}_s = 0. \quad (1.12)$$

The splitting potential $\hat{\mathcal{F}}$ has a simple geometric meaning. Let $\hat{\mathcal{F}}_{\pm}$ be the restrictions on $\hat{\mathcal{C}}$ of the generating functions $\mathcal{F}_\pm$ of the whiskers $\mathcal{W}_\pm$ in the original coordinates $(X, Y)$. Then the splitting distance is $\partial_X \hat{\mathcal{F}}_{\pm}(X)$, or the difference of the generalized momenta $Y$, given the same value of the generalized coordinates $X$ on $\hat{\mathcal{C}}$. Suppose, the diffeomorphism $\hat{A}$ satisfies (1.11). Let a transformation $\hat{\Xi}(\hat{A}, \hat{S}) \in \hat{E}_{\mathcal{F}}$ with some “free term” $\hat{S}$ act from $(\hat{X}, \hat{Y})$ to $(X, Y)$. Then the generating functions for the restrictions of whiskers $\mathcal{W}_\pm$ on $\hat{M}$, written in the chart $(\hat{X}, \hat{Y})$ are $\hat{\mathcal{F}}_{\pm} \circ \hat{A} - \hat{S}$, thus the difference of the generalized momenta $\hat{Y}$, corresponding to the same values of the generalized coordinates $\hat{X}$ becomes $\partial_{\hat{X}} \hat{\mathcal{F}}_s(\hat{X})$.

We call $(\hat{X}, \hat{Y})$ an optimal standard symplectic chart over $\hat{M}$. In such a chart the splitting potential $\hat{\mathcal{F}}_s = \hat{\mathcal{F}} \circ \hat{A}$ satisfies (1.12). There exists an infinite number of such symplectic charts: they can be obtained from the restriction of the original chart $(X, Y)$ to $\hat{M}$, by means of a transformation $\Xi(\hat{A}, \hat{S}) \in \hat{E}_{\mathcal{F}}$, where the diffeomorphism $\hat{A}$ of $\hat{\mathcal{C}}$ satisfies (1.11), and $\hat{S}$ is any scalar function on $\hat{\mathcal{C}}$.

The equation (1.1) is convenient in order to study the splitting further. As we mentioned earlier, the functions $\mathcal{F}_\pm(x, \bar{x})$ (and so will $\hat{\mathcal{F}}_\pm \circ \hat{A}$) have a part $\langle \xi_\pm, \bar{x} \rangle$, which is not quasiperiodic in $\bar{x}$, thus yielding a constant component $\xi_\pm - \xi_\pm$ to the splitting distance. In the homoclinic case the fact that the tori $\mathcal{T}_\pm$ are the same geometric object implies $\xi_\pm = \xi_\pm$ (Lemma 1). But we see no reason for it to be the case in the heteroclinic situation.

Thus, in the homoclinic set-up the function $\hat{\mathcal{F}}_s = \hat{\mathcal{F}}(\hat{x}, \hat{x}, \mu)$, satisfying the equation (1.12) is quasiperiodic in $\hat{x}$. Then it becomes a quasiperiodic function of $\hat{x} - \bar{\omega} \int_{\bar{\omega}}^{\hat{x}} \frac{d\zeta}{f_{10}(\zeta)}$, which leads to exponentially small estimates for the splitting distance near simple resonances, obtained for model problems in [5] and [15], and claimed erroneously in [14] — see Theorem 2 and the remarks following it.

2. Set-up

For the motivation we recall a Hamiltonian, representing a localization of a Hamiltonian system in the action-angle variables near a simple resonance (see [16]):

$$H(x, \bar{x}, y, \bar{y}, \mu) = \langle \bar{\omega}, \bar{y} \rangle + \frac{y^2}{2} + U(x) + y \langle \bar{f}(x), \bar{y} \rangle + \frac{1}{2} \langle \bar{y}, F_2(x, \bar{x}) \rangle + H_1(x, \bar{x}, y, \bar{y}, \mu). \quad (2.1)$$

This Hamiltonian will be real-analytic in some complex extension of $x, \bar{x}, y, \bar{y} \in \mathbf{T} \times \mathbf{T}^n \times \mathbf{B} \times \mathbf{B}^n$, $H_1(\cdot, \mu)$ vanishing for $\mu = 0$. Moreover, the frequency $\bar{\omega}$ will be “fast”: $\bar{\omega} = \frac{\omega}{\sqrt{\varepsilon}}$ for a small positive $\varepsilon$, whereas $\mu \sim \varepsilon^p$ for some $p > 1$. One usually adds the assumption of analyticity in $\mu$ upon the Hamiltonian (2.1).
Unless $\tilde{f}_2, F_2$ are identically zero\(^1\), one imposes a non-degeneracy condition
\[
\det \left( \begin{array}{cc} 1 & \tilde{f}_2(0) \\ f_{20}(0) & F_{20}(0) \end{array} \right) \neq 0.
\]

The function $U(x)$ reaches the absolute maximum on $T$ at $x = 0:
1. $U(0) = 0, U'(0) = 0, U''(0) < 0$,
2. $\forall x_0 \in T \setminus \{0\}: U'(x_0) = 0, U(x_0) < 0$.

The phase portrait of (2.1), projected upon the $(x,y)$ plane, provided that $(\ddot{y}, \mu) = (0,0)$, contains saddles $(x,y) = (0,0)$ and $(x,y) = (2\pi,0)$, connected by separatrices $y = \mp \sqrt{-2U(x)}$. The assumptions on the function $U(x)$ allow to define an analytic $4\pi$-periodic function
\[
f_i(x) = \begin{cases} \sqrt{-2U(x)}, & x \in [0, 2\pi), \\ -\sqrt{-2U(x)}, & x \in [2\pi, 4\pi), \end{cases}
\]
fixing the branch of the square root by $\sqrt{-1} = i$. For the simple pendulum, $U(x) = \cos x - 1$, $f_i(x) = 2\sin \left( \frac{x}{2} \right)$.

We localize (2.1) near a separatrix (upper or lower branch) by changing (symplectically) $y \to y + f_1(x)$ (or $y \to y - f_1(x)$), denoting $X = (x, \bar{x}), Y = (y, \ddot{y})$, and obtaining
\[
H(X,Y,\mu) = y f_1(x) + (\ddot{\omega} + f(x) \tilde{f}_2(x), \ddot{y}) + \frac{y^2}{2} + y(\ddot{f}_2(x), \ddot{y}) + \frac{1}{2}(\ddot{y}, F(x) \ddot{y}) + H_1(X,Y,\mu),
\]
with an abuse of notation, for the functions $H, H_1$ are different from those in (2.1). At this stage the Hamiltonian (2.4) is no longer $2\pi$-periodic in $x$, but instead satisfies the following translational property:
\[
H \circ T = H, \quad \text{with} \quad T: (x, \bar{x}, y, \ddot{y}, \mu) \to (x + 2\pi, \bar{x}, y + 2f_1(x), \ddot{y}, \mu).
\]
The translation $T$ acts locally from the neighborhood of $(0,0)$ to the neighborhood of $(0,2\pi)$ in the variables $(x,y)$. Being symplectic, it conjugates the dynamics in the neighborhoods of $x = 0$ and $x = 2\pi$. Hence, the system (2.4) for $\mu = 0$ has a pair of invariant tori $\mathcal{T}_0^\pm$, such that $x = 0$ on $\mathcal{T}_0^-$ and $x = 2\pi$ on $\mathcal{T}_0^+$, and $Y = 0$ on both of them, connected by a homoclinic manifold, $\mathcal{H}_0$, formed by coinciding the unstable whisker $\mathcal{W}_0^-$ of the torus $\mathcal{T}_0^-$ and the stable whisker $\mathcal{W}_0^+$ of the torus $\mathcal{T}_0^+$, given by $Y = 0$.

Let $\tilde{f}_1(x) = \ddot{\omega} + f(x) \tilde{f}_2(x)$ in (2.4). Then without loss of generality one can render $\tilde{f}_1(x) \equiv \ddot{\omega}$, for otherwise let
\[
\ddot{\omega}(x) = \int_0^x \frac{\tilde{f}_1(\zeta) - \ddot{\omega}}{\tilde{f}_1(\zeta)} d\zeta, \quad A': \begin{cases} x = x', \\ \bar{x} = \bar{x} + \ddot{\omega}(x'), \end{cases}, \quad \Xi': \begin{cases} X = A'(X'), \\ Y = [dA']^{-T} Y', \end{cases}
\]
and further dispense with the prime indices and study $H \circ \Xi'$, under relevant analyticity arguments. Note that $\Xi' = \Xi'(A',0) \in \mathcal{E}_H$ (see (1.5)), discussed in the Introduction.

We further generalize the problem and introduce the quantitative assumptions, starting from the notations for the domains (see also [13]), deeming $\mathcal{C}$ and $\mathbb{R}^2$ topologically equivalent in the usual sense. Given a pair of real positive parameters $(\sigma, \tau)$ we will use the following notations:
\[
\mathbb{T}_\sigma \defeq \{ z \in \mathbb{C} \mid \Re z \in T, | \Im z | < \sigma \}, \quad \mathbb{B}_r(z_0) \defeq \{ z \in \mathbb{C} \mid | z - z_0 | < r \}, \quad \mathbb{B}_r(0) \equiv \mathbb{B}_r,
\]
\[
\mathbb{T}_\sigma^n \defeq \mathbb{T}_\sigma \times \cdots \times \mathbb{T}_\sigma, \quad \mathbb{B}_r^n \defeq \mathbb{B}_r \times \cdots \times \mathbb{B}_r.
\]
\(^1\)Otherwise the perturbation may not depend on $\ddot{y}$, a so-called "isochronous" case, which is not conceptually different, and can be treated within the KAM framework by merely introducing an extra parameter, as [9] shows.
Let us start out with the Hamiltonian \( H_\mu(X, Y) \), which is real-analytic in \( \mu \), and can be represented as follows:

\[
H_\mu(X, Y) = H_0(X, Y) + H_1(X, Y, \mu),
\]

where \( H_0(X, Y) \) is given by (1.1), and

\[
H_1(X, Y, \mu) = H_{10}(X, \mu) + H_{11}(X, \mu) + \langle \tilde{H}_{11}(X, \mu), \tilde{y} \rangle + O_2(X, Y, \mu).
\]

For some \( r < 1, \sigma > 0 \) denote \( \mathscr{J}_{r, \sigma} \overset{\text{def}}{=} \{ x \in \mathbb{C} | \text{Re} x \in (-r, 2\pi + r), |\text{Im} x| < \sigma \} \). As far as the Hamiltonian (2.7), (2.8) is concerned, suppose the following.

**Assumption 1.** There exists a set of real positive numbers \( \sigma, \kappa, \lambda, r, R \): \( \kappa, R < 1, r < \min(\lambda^{-1}, \sigma, 1) \), such that

1. **The function \( f_{10}(x) \) is such that**
   1.1 (a) its Taylor expansion at \( x = 0 \) contains only odd powers of \( x \), with \( f_{10}(0) = \lambda \),
   1.1 (b) \( \forall x \in \mathscr{J}_{r, \sigma} \setminus \mathbb{B}_r \), \( |f_{10}(x)| > \frac{1}{4} \lambda r \),
   1.1 (c) \( \max_{\mathbb{B}_r} |u_{x x}| < \frac{1}{r} \).

2. **The frequency \( \bar{\omega} \) is Diophantine(\( \gamma, \tau \)).

3. \( |\det [F_{30}(0)]| \geq R \).

4. **The Hamiltonian (2.7)-(2.8) is analytic in \((x, \bar{x}, Y) \in \mathscr{J}_{r, \sigma} \times \Gamma \times \mathbb{B}^{n+1} \), continuous up to the boundary, and for all \((X, Y) \) in this domain and all \( \mu \) small enough \(|H_\mu| < 1, |H_1| \leq |\mu| \).**

5. **The Hamiltonian (2.7)-(2.8) satisfies (2.5), for \((x, y) \) in some neighborhood of \((0, 0) \).**

Assumption 1.5 plus real analyticity allow to translate all the proper transverse properties of the Hamiltonian near \( x = 0 \) to the neighborhood of \( x = 2\pi \). In particular, one gets \( f_{10}(x + 2\pi) = -f_{10}(x) \), \( \forall x \in \mathscr{J}_{r, \sigma} \), hence \( f_{10} \) can be continued \( 4\pi \)-periodically. Besides, from (2.4) it follows that \( \kappa < |\bar{\omega}|^{-1} \) and \( |f_{10}(x)| < \kappa^{-1}, \forall \bar{x} \) (for specific problems, pursuing sharper estimates, one may be interested in introducing a pair \((\kappa, \bar{\kappa})\) of different analyticity parameters in the variables \((y, \tilde{y}) \)). If \( f_{10} \) satisfies Assumption 1.1(a,b), with the parameters \( \lambda, r \), one can redefine \( r < \min(\frac{r}{2}, \frac{1}{8} \lambda \kappa r^2) \) in order to satisfy Assumption 1.1(c).

Of great importance for quantitative analysis will further be the function

\[
t_0(x) = \int_0^\pi \frac{d\bar{\zeta}}{f_{10}(\bar{\zeta})}, \quad x \in \mathscr{J}_{r, \sigma} \overset{\text{def}}{=} \mathscr{J}_{r, \sigma} \setminus ((-\infty, 0] \cup [2\pi, \infty)). \tag{2.9}
\]

Notationwise, further for any domain \( \mathcal{D} \subset \mathcal{C} \), we will use the notation \( \tilde{D} = \mathcal{D} \setminus ((-\infty, 0] \cup [2\pi, \infty)) \), as well as \( \tilde{\mathcal{D}} \) for the closure of \( \mathcal{D} \).

In order to make \( t_0 \): \( x \to t, t = t_0(x) \) a conformal map, we had to draw in (2.9) the branch cuts along the real axis for \( x < 0 \) and \( x > 2\pi \), because at \( x = 0, 2\pi \) the function \( t_0(x) \) has logarithmic singularities, the branch of the natural logarithm being chosen such that \( \text{Im} t_0(x) \equiv 0, x \in (0, 2\pi) \). On the shores of the branch cuts \( \text{Im} t_0(x) = \pm \frac{\pi}{2} \). One can easily check Taylor expanding \( f_{10}(x) \) at zero \((2\pi)\), that for any \( x \in \tilde{\mathbb{B}}_r \), there exists a segment of a level curve of the function

\[\text{Theorem 1 has been proved under this assumption in [13], although it is amenable to a weaker one that } f_{10} \text{ have a simple zero at the origin by a rather technical modification of Assumption 1.1(b).}\]
Im\(t_0(x)\), connecting \(x\) and the origin (the point \(2\pi\)) — see Proposition 2.2 in [13]. In the latter paper we developed a procedure, describing how one can cut the domain \(\mathcal{J}_{r,\sigma}\) in order to lead the KAM argument which requires that any \(x\) in the domain of the Hamiltonian (2.7) can always be connected to the origin (or the point \(2\pi\)) by a segment of a level curve of the function \(\text{Im} \, t_0(x)\), fully contained in this domain. For the splitting problem this is easier, for the splitting distance can be defined only for such \(x\), that there exists a level curve of \(\text{Im} \, t_0(x)\), connecting 0 and \(2\pi\) and passing through \(x\). Hence, let \(\Gamma_{\mathcal{J}_{r,\sigma}}\) be the set of all level curves of the function \(\text{Im} \, t_0(x)\), emanating from \(x = 0\) and terminating at \(x = 2\pi\) (the endpoints not being included), which are fully contained in \(\mathcal{J}_{r,\sigma}\). This set is not empty, for \((0, 2\pi)\) is such a curve, where \(\text{Im} \, t_0(x) = 0\); besides from the above described local properties near \(x = 0, 2\pi\) this set will contain other level curves \(\text{Im} \, t_0(x) = \rho^*\) with some small \(\rho^*\). Let \(\gamma_{\rho^*} \in \Gamma_{\mathcal{J}_{r,\sigma}}\) be an element of the above set, such that on it \(\text{Im} \, t_0(x) = \rho^*\). For simplicity we will further use the symmetry, coming from the fact that the function \(f_{10}(x)\) is odd in \(x\) and \(f_{10}(x + 2\pi) = -f_{10}(x)\), although one can certainly do without it. Define
\[
\rho = \sup_{\gamma_{\rho^*} \in \Gamma_{\mathcal{J}_{r,\sigma}}} \rho^* , \quad T = t_0(2\pi - r),
\]
\[
\mathcal{D}_\rho = \bigcup_{|\rho| \leq \rho} \gamma_{\rho^*} , \quad \mathcal{D}_{\rho,T} = \mathcal{D}_\rho \cup \mathcal{B}_r \cup \mathcal{B}_r(2\pi), \quad \mathcal{D}_{\rho,T}^+ = \mathcal{D}_{\rho,T} \cap \mathcal{D}_{\rho,T}^+ .
\]
(2.10)

In other words, \(\rho = \rho[f_{10}, \sigma]\) is the supremum of the width of a by-infinite strip in the \(t\)-plane, whose pre-image \(\mathcal{D}_\rho\) under the map \(t_0\) will be contained in \(\mathcal{J}_{r,\sigma}\). The set \(\mathcal{D}_\rho\) has the shape of a boat with the endpoints at 0, 2\(\pi\). We further put the balls \(\mathcal{B}_r, \mathcal{B}_r(2\pi)\) on it in order to obtain the set \(\mathcal{D}_\rho\), so that the image of \(\mathcal{D}_{\rho,T}\) under the map \(t_0\) gets wider and wider up to \(|\text{Im} \, t| \to \pi/\lambda\) as \(r \to \pm \infty\). The quantity \(T > 0\) is chosen such that the sets \(\mathcal{D}_{\rho,T}^\pm\) are respectively the pre-images under the map \(t_0\) of the semi-infinite strips \(\{t| | \text{Re} \, t < T, |\text{Im} \, t| < \rho\}, \{t| | \text{Re} \, t > -T, |\text{Im} \, t| < \rho\}\), \(\mathcal{D}_{\rho,T}^\pm\) being their intersection, and finally the sets \(\mathcal{D}_{\rho,T}^{\pm}\) are obtained by donning small balls around \(x = 0, 2\pi\) over respectively the sets \(\mathcal{D}_{\rho,T}^\pm\). These latter sets satisfy the assumptions of the KAM theorem in [13] (in particular, \(\mathcal{B}_r \cap \mathcal{D}_{\rho,T}^{\pm} = \mathcal{B}_r(2\pi) \cap \mathcal{D}_{\rho,T}^{\pm} = \emptyset\) — see Proposition 2.2 in [13]). In view of Assumption 1, which remains true if one decreases \(r\), this construction remains valid as well as all the notation of (2.10), e.g. \(\mathcal{D}_{\rho,T}^{\pm}\), remains unambiguous, for any \(0 < r' < r, 0 < \rho' < \rho, 0 < T' < T\).

For the sets defined in (2.10) let us construct the direct products:
\[
\mathcal{C} = \mathcal{C}_{r,\rho} \equiv \mathcal{D}_{\rho,T}^\pm \times \mathbb{T}_\rho, \quad \mathcal{M} = \mathcal{M}_{r,\rho,\sigma,\kappa} \equiv \mathcal{C}_{r,\rho,\sigma,\kappa} \times \mathcal{B}_\kappa^{n+1},
\]
(2.11)
\[
\mathcal{C}^\prime = \mathcal{C}_{r,\rho,\kappa}^\prime \equiv \mathcal{D}_{\rho,T}^\prime \times \mathbb{T}_\rho, \quad \mathcal{M}^\prime = \mathcal{M}_{r,\rho,\sigma,\kappa}^\prime \equiv \mathcal{C}_{r,\rho,\sigma,\kappa}^\prime \times \mathcal{B}_\kappa^{n+1}.
\]

All the scalar functions on the spaces \(\mathcal{C}_{r,\rho,\sigma,\kappa}\), \(\mathcal{C}_{r,\rho,\sigma,\kappa}^\prime\), \(\mathcal{C}_{r,\rho,\sigma,\kappa}^\prime\), \(\mathcal{C}_{r,\rho,\sigma,\kappa}\) will be members of the corresponding Banach spaces \(\mathcal{A}^j_{(1)}\) (the symbol \((\cdot)\) specifying the domain) of real-analytic functions on a corresponding domain continuous along with all their partial derivatives up to the order \(j = 0, 1, \ldots , \) up to the boundary, equipped with the norms \(|\mathcal{Q}^j|_{(1)}\), namely \(|\mathcal{Q}^j|_{(1)} \equiv \max_{k=0, \ldots, j} \left(\sup_{(1)} |D^k \mathcal{Q}|\right)\). We shall use the notation \(\mathcal{A}^j_{(1)}\) for the spaces of real-analytic \(l\)-vector functions on the domain \((\cdot)\). The functional spaces \(\mathcal{A}^j_{(1)}\) on \(\mathcal{M}_{r,\rho,\sigma,\kappa}\), \(\mathcal{M}_{r,\rho,\sigma,\kappa}^\prime\), \(\mathcal{M}_{r,\rho,\kappa}\), will be the Banach spaces of absolutely convergent in \(\mathcal{B}_\kappa^{n+1}\) Taylor series, whose coefficients are in \(\mathcal{A}^j_{(1)}\), where the subscript \((\cdot)\) stands respectively for \(\mathcal{C}_{r,\rho,\sigma,\kappa}\), \(\mathcal{C}_{r,\rho,\sigma,\kappa}^\prime\), \(\mathcal{C}_{r,\rho,\sigma,\kappa}^\prime\), the norm being the supremum of the Taylor series, whose coefficients are the.
corresponding norms of the corresponding Taylor coefficients, for the function itself and its partial
derivatives up to the order j. Zero superscripts will be omitted in the norm notation for j = 0, and
the maximum entry norm will be used for the vector (matrix) functions. We will imply real analyticity
and boundedness in the norm | |(·) by simply saying “a function on” a domain (·) For a function u(·, x),
analytic in x for x ∈ T^n we shall use its Fourier expansion: u(·, x) = \sum_{k∈\mathbb{Z}^n} u_k(·)e^{j\overline{k} \cdot x}, with the Fourier
coefficients u_k(·), in particular u_0(·) standing for the average in \overline{x}.

We will further restrict the Hamiltonian (2.7)–(2.8) upon M_{r,ρ,T,σ,κ}, Assumption 1 being satisfied
on it. If μ = 0 the Hamiltonian (2.7) possesses a pair of invariant partially hyperbolic tori \mathcal{J}_0^\mp with the
whiskers \mathcal{J}_0^\mp, emanating respectively from the tori \mathcal{J}_0^\pm. They are defined by the identity embeddings
of (1.4) into the phase space (hence the zero subscripts versus the notation of (1.4)), the whiskers
coinciding on \mathcal{J}_{ρ,T,σ,κ}, forming a homoclinic whisker \mathcal{J}_0. Under the condition (2.5), the tori \mathcal{J}_0^\pm can
be viewed as the same geometric object. Moreover, the following Theorem 1 establishes the existence
of the pair of perturbed tori \mathcal{J}_ρ^\pm, which also correspond to a single geometric object. Indeed, since T
is symplectic, then if \mathcal{J}_\perp is an invariant torus for the perturbed system, lying in the domain of T,
then T \mathcal{J}_\perp = \mathcal{J}_\perp would be also an invariant torus, lying on the same energy surface of H (in fact, in
order that this be true, T does not have to be symplectic, but one only needs non-degeneracy of its
differential dT — see Lemma 1).

3. KAM theorem, splitting potential, splitting distance

The Hamiltonian (2.7)–(2.8) restricted to M_{r,ρ,T,σ,κ} under Assumption 1.1–1.4, satisfies the hypotheses
of the KAM theorem proved in [13]. Although the theorem was proved for the unstable whisker only,
the modification of it in order to apply to the stable whisker is clearly straightforward. Denote
M_0^\perp ≜ M_{r,ρ,T,σ,κ}^\perp. We adapt the (quite technical) formulation of this theorem to the needs of the particular
application herein. Recall that the μ-dependence is implied in the notation for a function
(map, vector field), whenever it carries a subscript ± or wears a “hat”.

**Theorem 1.** Let δ be a real number, such that 0 < δ < min \left( \frac{σ}{4}, \frac{κ}{4}, \frac{ρ}{4}, \frac{T}{4}, r \right) let \overline{δ} = \frac{1}{4}δr, \overline{τ} = \max(τ, 2). There exists a constant C = C(\overline{r}, \overline{τ}) > 1, independent of the rest of the parameters of the
problem, such that if

\[ |μ| < μ_0(γ, δ, κ, λ, τ, R, τ, n) = C^{-1} γ^3 δ^{3λ+3} κ^9(λr)^{3}r^3R^2 \]  

(3.1)

there exists a pair of exact symplectic transformations Ξ_± : (X^±, Y^±) ∈ M_0^\perp ≜ M_{r,ρ,T,σ,κ}^\perp → (X, Y) ∈ M_0^\perp, such that

\[ Ξ_± : \begin{cases} X = A_±(X^±), \\ Y = (dA_±)^{-T}(Y^± + \partial_X S_±(X^±)) \end{cases} \]  

(3.2)

where

\[ |A_± - \text{Id}|_\mathcal{J}_0^\perp < C|μ| γ^{-2δ-2κ-6}(λr)^{-6}r^{-2}R^{-1}, \]  

(3.3)

\[ |S_±|_\mathcal{J}_0^\perp < C|μ| γ^{-1δ-τ-1κ-2}(λr)^{-2}rR^{-1} \]  

(3.4)

The functions S_± are such that

\[ S_±(X^±) = S_±(x^±, x^\mp) = \langle ξ_±, x^\mp \rangle + B_±(x^±, x^\mp) \]  

(3.5)

with the constants ξ_±, and the quasiperiodic in \overline{x} functions B_±(x, \overline{x}) both satisfying the estimate (3.4).
The numbers $c_x = O(\mu)$. The systems $H_\mu \circ \Xi_\mu$ on $M_\delta^+$ each possess an exact Lagrangian submanifold $\mathcal{W}_x$ (see (1.4)), containing an invariant torus $\mathcal{T}_x$, the flow whereupon is the Kronecker flow with the frequency $\tilde{\omega}$.

The system (2.7)-(2.8) possesses a pair of invariant exact Lagrangian submanifolds $\mathcal{W}_x = \Xi_x \mathcal{W}_x^+$, containing the invariant tori $\mathcal{T}_x = \Xi_x \mathcal{T}_x^+$, which for $X \in \mathcal{C}_m^+$, $\mathcal{C}_m^+$ can be described as

$$\mathcal{W}_x = \{(X, Y) | Y = \partial_X \mathcal{J}_x(X), \mathcal{J}_x(X) = (\xi_x, \tilde{x}) + \mathcal{B}_x(X)\},$$

(3.6)

where $\mathcal{J}_x = S_x \circ A_0^{-1}$, the constants $\xi_x$ are the same as above, the functions $\mathcal{B}_x(X) = \mathcal{B}_x(x, \tilde{x})$ are $O(\mu)$ in the norm $\|\cdot\|_{\mathcal{C}_m}$ and quasiperiodic in $\tilde{x}$.

The functions $\mathcal{J}_x(X)$ satisfy the Hamilton-Jacobi equations $H_\mu(X, \partial_X \mathcal{J}_x(X)) = c_x$ on $\mathcal{C}_m^+$.

Note, that the condition (3.4) specifies exactly what is meant further by the notation $O(\mu)$. The constant $C$, though, can be increased, if necessary, without notice. The functions $S_x, \mathcal{J}_x$ are defined modulo a constant.

The following Lemma takes into account the fact that we are dealing with the homoclinic problem, namely that the pair of tori $\mathcal{T}_x$ can be viewed as the same geometric object.

**Lemma 1.** Under Assumption 1.5, $\forall \mu$, $c_- = c_+, \xi_- = \xi_+$.

**Proof.**

The statement follows from the condition (2.5). Generally, since $dT$ is non-degenerate, no matter what $T$, the condition $H \circ T = H$ implies that if $\mathcal{T}_x$ is an invariant submanifold in the domain of $T$, then $T \mathcal{T}_x$ is an invariant submanifold in the range of $T$. This is true, because $T$ maps the energy surface of $H_\mu$ into itself, thus the flow of $H_\mu$ at the range of $T$, and the pull-forward under $T$ of the flow of $H_\mu$ from the domain of $T$ into the range of $T$ are tangent to one another.

Thus, $\mathcal{T}_x = T \mathcal{T}_x$, which implies that $\mathcal{T}_x$ lie on the same energy level of $H_\mu$, so $c_- = c_+.

In order to prove that $\xi_- = \xi_+$, let us restrict ourselves to the real setting to avoid arguing about the analyticity domains. Then the statement for complex $\mu$ will follow from real analyticity. Let $x$ be a small real number, e.g. $|x| < \delta$, and $\tilde{x} \in \mathbb{T}^n$. If $z_- = (x, \tilde{x}, \partial_x \mathcal{J}_-(x, \tilde{x}), \partial_{\tilde{x}} \mathcal{J}_-(x, \tilde{x}))$ is a point on the torus $\mathcal{T}_x$, where in particular $X = (x, \tilde{x}) = A_-(0, \tilde{x})$, for some $\tilde{x} \in \mathbb{T}^n$, then $z_+ = T z_- = (x + 2\pi, \tilde{x}, \partial_x \mathcal{J}_-(x, \tilde{x}) + 2 \mathcal{J}_-(0, \tilde{x}), \partial_{\tilde{x}} \mathcal{J}_-(x, \tilde{x}))$ is a point on $\mathcal{T}_x$, which also has to admit the representation $z_+ = (x + 2\pi, \tilde{x}, \partial_x \mathcal{J}_+(x + 2\pi, \tilde{x}), \partial_{\tilde{x}} \mathcal{J}_+(x + 2\pi, \tilde{x}))$, hence $(x + 2\pi, \tilde{x}) = A_+(2\pi, \tilde{x}^+)$ for some $\tilde{x}^+ \in \mathbb{T}^n$, and one shall have

$$\partial_x \mathcal{J}_-(x, \tilde{x}) = \partial_x \mathcal{J}_+(x + 2\pi, \tilde{x}),$$

where $(x, \tilde{x}) = A_-(0, \tilde{x}) = A_+(2\pi, \tilde{x}^+) = (2\pi, \tilde{0}).$

Let $A_-(0, \tilde{x}^-) \equiv (a_-(\tilde{x}^-), \tilde{\alpha}_-(\tilde{x}^-)), A_-(2\pi, \tilde{x}^+) \equiv (a_+(\tilde{x}^+), \tilde{\alpha}_+(\tilde{x}^+))$. Requiring that $\tilde{x}$ be the same in the parameterizations of the tori $\mathcal{T}_x$ defines a near-identity diffeomorphism $\tilde{x}^- = \tilde{x}^+(\tilde{x}^-)$ of $\mathbb{T}^n$, or the ‘scattering map’, which is defined implicitly by $\tilde{\alpha}_-(\tilde{x}^-) = \tilde{\alpha}_+(\tilde{x}^+(\tilde{x}^-)).$ The presence of the condition (2.5) insures that also $2\pi + a_-(\tilde{x}^-) = a_+(\tilde{x}^+(\tilde{x}^-)).$

Recall (3.2) and $S_x = \mathcal{J}_x \circ A_0$. Then by the chain rule:

$$(\mathcal{J}_- - \mathcal{J}_+)T \partial_{\tilde{x}} S_-(2\pi, \tilde{x}^-) = (\tilde{x}^+ - \tilde{x}^-)T \partial_{\tilde{x}} S_+(2\pi, \tilde{x}^-),$$

$$(\tilde{x}^+ - \tilde{x}^-)T \partial_{\tilde{x}} S_+(2\pi, \tilde{x}^-) = (\tilde{x}^+ - \tilde{x}^-)T \partial_{\tilde{x}} S_+(2\pi, \tilde{x}^-).$$

Since $\tilde{x}^+ = \tilde{x}^+(\tilde{x}^-)$ is a near-identity diffeomorphism of $\mathbb{T}^n$, namely one can write $\tilde{x}^+ = \tilde{x}^- + \tilde{x}^+\tilde{x}^-(\tilde{x}^-)$ for some quasiperiodic in $\tilde{x}^-$ function $\tilde{x}^+\tilde{x}^-(\tilde{x}^-)$, which also satisfies the norm estimate (3.3), and from the specific view (3.5) of the functions $S_x$, it follows that $\xi_- = \xi_+$, for all $\mu$ small enough (satisfying (3.1)).
Denote for $l = 2, 3, 4$: 
\[
\hat{\mathcal{E}}_{l\delta} \overset{\text{def}}{=} \hat{\mathcal{E}}_{p-\delta, T-\delta_1-\delta_2}, \quad \hat{\mathcal{H}}_{l\delta} \overset{\text{def}}{=} \hat{\mathcal{E}}_{l\delta} \times \mathbb{B}^{n+1}_{\kappa-l\delta}.
\]

Theorem 1 and Lemma 1 yield the following statement about the “splitting potential”, defined by (1.6):

**Corollary 1.** The splitting potential $\hat{\mathcal{F}}(x, \vec{x})$ is a quasiperiodic in $\vec{x}$ function on $\hat{\mathcal{E}}_{2\delta}$, and $|\hat{\mathcal{F}}|^1_{\hat{\mathcal{E}}_{2\delta}} = O(\mu)$.

Furthermore, restricting upon $\hat{\mathcal{E}}_{2\delta}$ the Hamilton-Jacobi equations (1.7) and subtracting the equation for the stable separatrix from the one for the stable separatrix, using $c_- = c_+$, we shall obtain on the above domain

\[
[f_{10}(x) + H_1(X, \mu)]\partial_x \hat{\mathcal{F}} + [\tilde{\omega} + H_1(X, \mu)]\partial_{\vec{x}} \hat{\mathcal{F}} + [O_2(X, \partial X, \mathcal{J}_-, \mu) - O_2(X, \partial X, \mathcal{J}_+, \mu)] = 0, \quad (3.7)
\]

where in particular the function $O_2(X, \partial X; \mathcal{J}_-, \mu)$ is at least quadratic in $Y$. Thus, $[O_2(X, \partial X; \mathcal{J}_-, \mathcal{J}_+, \mu)]$ can be rewritten as $\langle \bar{\Omega}(X, \partial X; \mathcal{J}_-, \mathcal{J}_+, \mu), \partial_x \hat{\mathcal{F}} \rangle$, the $(n+1)$-vector quantity $\bar{\Omega}(X, \partial X; \mathcal{J}_-, \mathcal{J}_+, \mu)$ such that

\[
\bar{\Omega}(X, \partial X; \mathcal{J}_-, \mathcal{J}_+, \mu) \leq \sup_{\vec{x} \in \hat{\mathcal{E}}_{2\delta}} \left[ \Omega_2 \left( X, |\partial_x \mathcal{J}_-|_{\hat{\mathcal{E}}_{2\delta}} + |\partial_x \mathcal{J}_+|_{\hat{\mathcal{E}}_{2\delta}} \right) \right] = O(\mu) .
\]

With the notation of (1.8), (1.9) the equation (3.7) reads

\[
\hat{\mathcal{F}} = 0.
\]

From Corollary 1 and the latter the equation (3.7) reads

\[
\hat{\mathcal{F}} = 0 .
\]

Hence, we interpret the differential operator of the above equation as a vector field $\hat{\mathcal{F}}$ on $\hat{\mathcal{E}}_{2\delta}$, given by (1.10). We shall further solve the conjugacy problem (1.11) on $\hat{\mathcal{E}}_{2\delta}$, thus casting the equation (3.8) into the form (1.12).

The map $t_0$, defined by (2.9) is conformal on $\hat{\mathcal{G}}_{p-2\delta, T-2\delta}$, its derivative being bounded between $\kappa$ and $4(\lambda r)^{-1}$. Let us extend it to $\hat{t}_0$: $\hat{\mathcal{E}}_{2\delta} \to \Pi_{2\delta}$, where $\Pi_{2\delta} \overset{\text{def}}{=} \{ t \in \mathbb{C} | \text{Re} t < T - 2\delta, |\text{Im} t| < \rho - 2\delta \} \times \mathbb{T}_{M-3\delta}$, by letting

\[
\hat{t}_0: \begin{cases} t = t_0(x), \\ \vec{x} = \tilde{x} . \end{cases}
\]

Then by (1.10), (2.9), let $\Omega_0 \overset{\text{def}}{=} d\hat{t}_0 V_0 \circ \hat{t}_0 = (1, \tilde{\omega})$, and let $d\hat{t}_0 V_1 \circ \hat{t}_0 = \Omega_1$, with $|\Omega_1|_{\Pi_{2\delta}} < O(\mu)(\lambda r)^{-1}$. In order to obtain the latter estimate we used Assumption 1.1 (b). Let also $\tilde{\Omega} = \Omega_0 + \Omega_1$. We shall prove that there exists a near-identity transformation $\Phi = \text{Id} + \Phi_1$, where $\Phi_1 \in \mathcal{C}^1(\Pi_{2\delta}, \mathbb{R}^{n+1})$, with $\Pi_{2\delta} = \{ t \in \mathbb{C} | |\text{Re} t| < T - 3\delta, |\text{Im} t| < \rho - 3\delta \} \times \mathbb{T}_{M-3\delta}$, such that $d\Phi^{-1} \tilde{\Omega} \circ \Phi = \tilde{\Omega}_0$. In other words, $\Phi_1$ satisfies the functional equation

\[
\Omega_1 \circ (\text{Id} + \Phi_1) - \partial \Phi_1 = 0 , \quad (3.9)
\]

where

\[
\partial = \frac{\partial}{\partial t} + \langle \tilde{\omega}, \frac{\partial}{\partial \vec{x}} \rangle.
\]

Then $\hat{A} = \hat{t}_0^{-1} \Phi \hat{t}_0$ solves (1.11).

The construction of the transformation $\Phi$ is postponed to the Appendix. It is done by a single application of the Implicit function theorem of [17], and is nothing new compared to the proof of the analogous statement in [15].

Then we have the following theorem.
Theorem 2. Suppose (3.1) is satisfied. There exists a diffeomorphism \( \hat{A} : \hat{\mathcal{E}}_{35} \to \hat{M}_{35} \), with
\[
|\hat{A} - \text{Id}|^{1}_{35} < C|\mu|\gamma^{-1}\delta^{-x_3}K^{-2}(\lambda r)^{-1},
\]
(3.11)
such that for any function \( \hat{S} \in \mathcal{A}^{1}_{35} \), with \( |\hat{S}|^{1}_{35} < \delta \), and an exact symplectic diffeomorphism \( \hat{\xi} = \hat{\xi}(\hat{A}, \hat{S}) : (\hat{X}, \hat{Y}) \in \hat{M}_{35} \to (X, Y) \in M_{35} \), namely
\[
\hat{\xi} : \begin{cases}
X = \hat{A}(\hat{X}), \\
Y = (\text{d}\hat{A})^{-T} \left( \hat{Y} + \partial_{\hat{X}} \hat{S}(\hat{X}) \right),
\end{cases}
\]
(3.12)
the standard symplectic chart \((\hat{X}, \hat{Y})\) on \( \hat{M}_{35} \) is such that \( \hat{M}_{35} \) contains a pair of exact Lagrangian submanifolds \( \hat{\mathcal{W}} \), for which \( \hat{X} \in \hat{\mathcal{C}}_{35} \) can be described by the generating functions \( \hat{\mathcal{S}}_x \) as \( \hat{\mathcal{W}} = \{ (\hat{X}, \hat{Y}) | \hat{Y} = \partial_{\hat{X}} \hat{\mathcal{S}}(\hat{X}) \} \). Given \( \hat{X} \), the distance between these submanifolds
\[
\hat{\Delta}(\hat{X}) \equiv \partial_{\hat{X}} \left( \hat{\mathcal{S}}_x(\hat{X}) - \hat{\mathcal{S}}_x(\hat{X}) \right) \equiv \partial_{\hat{X}} \hat{\mathcal{S}}_x(\hat{X}),
\]
is such that
\[
|\hat{\Delta}|^{1}_{35} < C|\mu|\gamma^{-1}\delta^{-x_3}K^{-2}(\lambda r)^{-1}.
\]
The function \( \hat{\mathcal{S}}_x(\hat{X}) = \hat{\mathcal{S}}_x(x, \hat{x}) \) is quasiperiodic in \( \hat{x} \), moreover
\[
\hat{\mathcal{S}}_x(x, \hat{x}) = \hat{C}_0 + \sum_{\hat{k} \neq 0} \hat{C}_k e^{i(k, \hat{x} - \hat{x}(\hat{x}))},
\]
(3.13)
where \( \hat{x}(\hat{x}) \equiv t_0(\hat{x}) \) (see (2.9)). The Fourier coefficients \( \hat{C}_k \) satisfy the upper bounds
\[
|\hat{C}_0| < C|\mu|\gamma^{-1}\delta^{-x_3}K^{-2}(\lambda r)^{-1},
\]
\[
|\hat{k}||\hat{C}_k| < C|\mu|\gamma^{-1}\delta^{-x_3}K^{-2}(\lambda r)^{-1} e^{-(\sigma-35)|\hat{k}| - (\rho-35)|\hat{k}|, \hat{\omega}|}, \quad \hat{k} \neq 0.
\]
(3.14)

Proof.

We apply Proposition 1 to the functional (3.9), with \( f \equiv \Omega_1 \), \( v \equiv \Phi_1 \) to get the transformation \( \hat{\Phi} = \text{Id} + \Phi_1 \). Then the transformation \( \hat{A} = \hat{\Phi}^{-1} = \hat{\Phi} \) solves the conjugacy problem (1.11). For the estimates recall that \( |\Omega_1|_{35} < O(\mu)(\lambda r)^{-1} \), (3.4), and use Footnote 3 in the Appendix, along with (3.4) and the standard Fourier coefficient estimates for the constants \( \hat{C}_k \) yield (3.14), being a slight modification of (3.4).

Furthermore, \( \hat{\mathcal{S}}_x = \hat{\mathcal{S}}_x \circ \hat{\mathcal{A}} - \hat{\mathcal{S}} \). Then the rest of the statements follow from the representations (3.5), (3.6), Lemma 1, and the fact that the transformation \( \hat{A} \) is near-identity (compare (3.11) and (3.1)). Forsooth, the latter fact and Corollary 1 guarantee that \( |\hat{\mathcal{S}}_x|^{1}_{35} \) is also bounded by the right-hand side of (3.4). Further, by (3.5), (3.6), and Lemma 1, \( \hat{\mathcal{S}}_x(\hat{X}) = \hat{\mathcal{S}}_x(x, \hat{x}) \) is quasiperiodic in \( \hat{x} \).

Besides, the conjugation problem (1.11) and (3.9) was set up in order that \( \hat{\mathcal{S}}_x \) satisfy (1.12). Hence \( \hat{\mathcal{S}}_x(\hat{X}) \) with \( \hat{X} \in \hat{\mathcal{C}}_{35} \) satisfies the following PDE: \( f_{10}(\hat{x}) \hat{\mathcal{S}}_x + (\hat{\omega}, \hat{\mathcal{S}}_x) = 0 \). This implies (3.13) and the standard Fourier coefficient estimates for the constants \( \hat{C}_k \) yield (3.14), with the bound for the norm of \( \hat{\mathcal{S}}_x \) and its partial derivatives in \( \hat{\mathcal{C}}_{35} \) being available.

Note, that the constant \( \hat{C}_0 \) in (3.13) is totally meaningless, and can be thrown away. Or, redefine \( \hat{\mathcal{S}} \equiv \hat{\mathcal{S}}_x \circ \hat{\mathcal{A}}^{-1} = \hat{C}_0 \). As we mentioned, \( \hat{\mathcal{S}}_0 \) does not depend on \( x \). Then, at (3.6), the function \( \Delta(x) = \partial_X \hat{\mathcal{S}}(x) \) is defined on \( \hat{\mathcal{C}}_{35} \), and if one looks back at (3.6), it gives the splitting distance in the original coordinates \( (X, Y) \), and Theorem 2 affords the following corollary.
Corollary 2. The function $\hat{\mathcal{F}}(X)$, restricted the projection $\text{Re} \hat{e}_{45}$ of the complex domain $\hat{e}_{45}$ on the real space satisfies the bound

$$|\hat{\mathcal{F}}|_{\text{Re} \hat{e}_{45}} < C|\mu|^{\gamma^{-1}T^{-1}K^{-1}}(\gamma T)^{-2}(\gamma T)^{-1}e^{-|\sigma-3\delta||k|-(\rho-3\delta)||k,\hat{\omega}|}. \quad (3.15)$$

We add a heuristic remark about how Theorem 2 can be applied to so-called “two time-scale” problems. Suppose, the frequency $\hat{\omega} = \frac{2n}{\sqrt{\varepsilon}}$, where $\varepsilon$ is a small parameter, which is the case in the Hamiltonians, which are Normal forms near simple resonances for the Hamiltonian systems in the action-angle variables. As we have shown at the beginning of Section 2 (see (2.1)-(2.6)), such a system can be converted to a particular case of the Hamiltonian (2.7)-(2.8), studied in this paper. In particular, it can be easily done for the models, studied in [3], [5], [14], [4], [15], involving simple pendulum, the function $f_{10}(x) = 2 \sin \left(\frac{x}{T}\right)$ (so, one can take $\rho = \frac{3}{2}$ and $T = 100$). Then, if the conditions of Theorem 1, namely Assumption 1, can be satisfied (although in general checking Assumption 1.4 may cause certain difficulties, it will along with the rest assumptions be satisfied in the pendulum-rotator models, studied in the above-mentioned references), after letting in (3.15) $\kappa \sim \sqrt{\varepsilon}$, $\delta \sim \varepsilon$, $\lambda, T, \rho, T, R = O(1)$, one will arrive at the exponentially small upper bounds for the splitting potential $\mathcal{F}$ (from which a constant has been extracted):

$$|\hat{\mathcal{F}}|_{\hat{\rho}, \hat{T}, \hat{\sigma}, \hat{\omega}} < C|\mu|^{e^{-\eta}}e^{-|\sigma-3\delta||k|-(\rho-3\delta)||k,\hat{\omega}|} \frac{1}{\sqrt{\varepsilon}},$$

for some finite exponent $p$. The parameters $\hat{\rho}, \hat{T}, \hat{\sigma}$ can be defined based upon (2.1) and its analyticity properties, using (2.3), (2.6), and then following the domain cutting procedure, described in Section 2, and will not depend on $\varepsilon$.

Appendix. Conjugation of vector fields

Assumption 1 in its full strength is unnecessary in order to obtain the solution $\Phi_1 = \Phi_1(\Omega_1)$ of the equation (3.9), to solve which we need only real-analyticity and the properties (in particular, precompactness) of the domain. The solution $\Phi_1$ will result from an application of Theorem 1.1 of [17] to the vector fields in question. Thus, the notation further will be somewhat independent, considerably correlating with the notation of [17].

Let $(\rho, T, \sigma)$ be a set of fixed positive parameters, the parameters $\delta, \delta'$, such that $0 < \delta' < \delta < \frac{3}{2}$ will vary. Let $\tilde{\omega} \in \mathbb{R}^n$ be a fixed vector.

Let $\Pi_{\rho, T, \sigma} \triangleq \{t \in \mathbb{C} | \text{Re} t < T, \text{Im} t < \rho\} \times \mathbb{T}^n_\sigma$, $\Pi_\delta \triangleq \Pi_{\rho-\delta, T-\delta, \sigma, \sigma}$. Let us partition $\mathbb{Z}^n = \mathbb{Z}^n_0 \cup \mathbb{Z}^n_- \cup \mathbb{Z}^n_+ = \mathbb{Z}^n_0 \cup \mathbb{Z}^n_-$, where $\mathbb{Z}^n_0 \triangleq \{k \in \mathbb{Z}^n_0 | \langle k, \tilde{\omega} \rangle = 1\}$, $\mathbb{Z}^n_- \triangleq \{k \in \mathbb{Z}^n_- | \langle k, \tilde{\omega} \rangle < 0\}$, $\mathbb{Z}^n_+ = \mathbb{Z}^n \setminus (\mathbb{Z}^n_0 \cup \mathbb{Z}^n_-)$. Let $p^{+} = \pm(i(\rho - \delta)$, $p^{0} = 0$ be complex numbers.

Consider the following functional equation $F(f, v) = 0$, with the functional

$$F(f, v) = f \circ (v + v) - \partial v, \quad (3.16)$$

the notation $\partial$ being defined by (3.10). This functional is a continuous map

$$F : \mathcal{A}(\Pi_0, \mathcal{C}^{n+1}) \times \mathcal{A}^1(\Pi_0, \mathcal{C}^{n+1}) \to \mathcal{A}(\Pi_0, \mathcal{C}^{n+1}),$$

clearly Lipschitz in $f$ and differentiable in $v$. We will show that given $f \in \mathcal{A}(\Pi_0, \mathcal{C}^{n+1})$ small enough (see Proposition 1 below), then for any $\delta$ we can find $v(f) \in \mathcal{A}^1(\Pi_0, \mathcal{C}^{n+1})$, such that $F(f, v(f)) = 0$.
(certainly $F(0,0) = 0$). This is a model problem of conjugation of vector fields, and except for the following Lemma 2, it follows the steps developed in Chapter 4 of [17] in order to apply Theorem 1.1 of [17] to a perturbation of a vector field on a torus. The purpose is showing that the conditions H.1-H.3 of that Theorem 1.1, for which suffice the following estimates (3.17) and (3.19). The norms of the functions on $\Omega_\delta$ will be denoted simply as $| \cdot |_\delta$, but we will keep in mind that if $F(f,v) = z$, then for $f, z$ we use the sup norms $| \cdot |_\delta \equiv | \cdot |_{\Omega_\delta}$, whereas for $v$ the notation $| \cdot |_{\Omega_\delta}$ will stand for $| \cdot |_{\Omega_\delta}$.

It is easy to verify that the functional $F$ satisfies the following estimate, if Taylor-expanded in $v$: for all $\delta, \delta'$, with the notation $D_2F(f,v)$ for the Frechet derivative of $F(f,v)$ in $v$ (the subscript 2 shall read “the second argument”, rather than “the second derivative”):

$$|F(f,v_2) - F(f,v_1) - D_2F(f,v_1)(v_2 - v_1)|_\delta < \frac{C}{(\delta - \delta')^2}|v_2 - v_1|^2,$$

(3.17)

where $v_1, v_2 \in \mathcal{A}^1(\Omega_\delta, \mathbb{R}^m)$, and the constant $C = C(\rho, T, n) > 1$, which will be further increased without notice. Indeed, for any $\hat{v} \in \mathcal{A}^1(\Omega_\delta)$ one gets

$$D_2F(f,v)\hat{v} = df \circ (\text{id} + v)\hat{v} - \partial \hat{v},$$

(3.18)

and then (3.17) is an application of the Cauchy inequality and the Taylor formula.

It remains to verify that given $f \in \mathcal{A}^1(\Omega_\delta, \mathbb{R}^m)$, the derivative $D_2F(f,v)$, as a linear map from $\mathcal{A}^1(\Omega_\delta, \mathbb{R}^m)$ to $\mathcal{A}^1(\Omega_\delta, \mathbb{R}^m)$, has an approximate right inverse $\eta(f,v)$, such that for any $\delta, \delta'$ the linear map $\eta(f,v): \mathcal{A}^1(\Omega_\delta, \mathbb{R}^m) \to \mathcal{A}^1(\Omega_\delta, \mathbb{R}^m)$, and for $z \in \mathcal{A}^1(\Omega_\delta, \mathbb{R}^m)$

$$|\eta(f,v)z|_\delta < \frac{C}{(\delta - \delta')^2}|z|_\delta,$$

$$|(D_2F(f,v) \circ \eta(f,v) - 1)z|_\delta < \frac{C}{(\delta - \delta')^2}|F(f,v)||z|_\delta.$$  

(3.19)

In the second inequality I stands for an $(n + 1) \times (n + 1)$ unit matrix.

If (3.19) is true, then the application of Theorem 1.1 of [17] results in the following Proposition.

**Proposition 1.** Let $|f|_0 < \epsilon < C^{-1}\delta^4$, where the constant $C = C(\rho, T, n)$. Then for every $f$ there exists $v \in \mathcal{A}^1(\Omega_\delta, \mathbb{R}^m)$, such that $|v|_\delta < C\delta^{-2}\epsilon$.

In order to prove (3.19) we shall first show in the following Lemma 2 that the operator $\partial = D_2F(0,0)$ has an exact right inverse $\eta(0,0)$, and will further follow [17], which used it in order to construct the approximate right inverse $\eta(f,v)$ of $D_2F(f,v)$.

**Lemma 2.** Let $\omega \in \{\pm i\partial, \pm i\partial, 0\}$, $g_\omega \in \mathcal{A}^1(\Omega_\delta)$ such that $g_\omega(t, \vec{x}) = \sum_{k \in \mathbb{Z}^m} g_{\omega k}(t)e^{i(k, \vec{x})}$. Then there exists a unique solution $u_\omega \in \mathcal{A}^1_{\Omega_\delta}$ of the Cauchy problem

$$\partial u_\omega = g_\omega,$$

$$u_\omega(p_\omega, \vec{x}) \equiv 0,$$

satisfying $|u_\omega|_\delta < \frac{C}{\delta - \delta'}|g_\omega|_{\Omega_\delta}$.

**Proof.**

After expanding $g_\omega$ in the Fourier series in $\vec{x}$, and accordingly seeking $u_\omega$ as a Fourier series in $\vec{x}$, we obtain the system of equations for the Fourier components for all $k \in \mathbb{Z}^m$:

$$u_{\omega \vec{k}} + i(k, \omega) u_{\omega \vec{k}} = g_{\omega k},$$

A closer look easily tells that one can choose $C = C(n)(\rho^2 + T^2)$, the constant $C(n)$ depending only on $n$. 

240 REGULAR AND CHAOTIC DYNAMICS, V. 5, № 2, 2000
thus the solution of the Cauchy problem is
\[ u_\varpi(t, \xi) = \sum_{k \in \mathbb{Z}^n} \int g_{-k}(\zeta) e^{i(k,\xi+\zeta(t))}, \]
where the integrals are taken along the straight lines, connecting \( p_\varpi \) and \( t \). Since along these lines \( |e^{i(k,\xi+\zeta(t))}| \leq 1 \), the statement of the Lemma follows after one applies the Cauchy inequality to estimate the derivatives of \( u_\varpi \).

Clearly, if \( g = g_0 + g_- + g_+ \) has a full Fourier spectrum, one takes the sum \( u = u_0 + u_- + u_+ \) for the solution of the Cauchy problem \( \partial u = g \), \( u(0, \xi) = u_0(0, \xi) + u_-(0, \xi) \).

Therefore, there exists the right inverse \( \eta(0,0) : A^1(\Pi_{\delta}, \mathcal{C}^{n+1}) \rightarrow A^1(\Pi_{\delta}, \mathcal{C}^{n+1}) \) of the operator \( \partial = D_2F(0,0) \), the operator \( \eta(0,0) \) satisfying the first relation of (3.19). We will further show how it can be used in order to construct the approximate right inverse \( \eta(f,v) \) of \( D_2F(f,v) \).

Let us formally differentiate \( F(f(X), v(X)) \) as a vector-function of \( X \):
\[ d\mathcal{F} = df \circ (id + v)V - \partial dv, \]
where \( V \overset{\text{def}}{=} 1 + dv \). For any \( \hat{\eta} \in A^1(\Pi_{\delta}, \mathcal{C}^{n+1}) \), if we multiply the above equation by \( \hat{\eta} = V^{-1}\hat{\eta} \), using \( \partial dv \hat{\eta} = (\partial V)\hat{\eta} = \partial(V\hat{\eta}) - V\partial\hat{\eta} \), we get
\[ d\mathcal{F} V^{-1}\hat{\eta} = df \circ (id + v)\hat{\eta} - \partial\hat{\eta} - \partial\hat{\eta} = d\mathcal{F} \hat{\eta} = D_2\mathcal{F}(f,v)\hat{\eta} - \partial\hat{\eta}, \]
where \( \partial = V\partial V^{-1} \). This entails the functional identity
\[ D_2\mathcal{F}(f,v)\hat{\eta} = -\partial\hat{\eta} = \mathcal{R}_{f,v}(\hat{\eta}), \tag{3.20} \]
where \( \mathcal{R}_{f,v}(\hat{\eta}) = d\mathcal{F} V^{-1}\hat{\eta} \).

The operator \( -\partial \hat{\eta} \) has a right inverse \( \eta_{\partial} = -V\eta(0,0)V^{-1} \), where the inverse \( \eta(0,0) \) of the operator \( \partial \) is stipulated by Lemma 2. The (linear) operator \( \eta_{\partial} : A^1(\Pi_{\delta}, \mathcal{C}^{n+1}) \rightarrow A^1(\Pi_{\delta}, \mathcal{C}^{n+1}) \). If \( z \in A^1(\Pi_{\delta}, \mathcal{C}^{n+1}) \), the first estimate of (3.19) is satisfied, moreover let \( \hat{\eta} = \eta_{\partial}(z) \) in (3.20), then
\[ D_2\mathcal{F}(f,v) \circ \eta_{\partial}(z) - z = \mathcal{R}_{f,v}(\eta_{\partial}(z)). \]

The right hand side can be estimated:
\[ |\mathcal{R}_{f,v}(\eta_{\partial}(z))|_\delta < |dF(f,u)|_\delta|V^{-1}|_\delta|\eta_{\partial}|_\delta < \frac{C}{{(\delta - \delta')^2}}|F(f,u)|_\delta|z|_\delta, \]
thus, if one lets \( \eta(f,v) \equiv \eta_{\partial} \), the second inequality of (3.19) becomes satisfied. Hence, Theorem 1.1 of [17] can be applied, yielding Proposition 1.

References


