Normal Forms, Stability and Splitting of Invariant Manifolds I. Gevrey Hamiltonians

Abed Bounemoura*

Centre de Recerca Matemàtica, Campus de Bellaterra, Edifici C, 08193, Bellaterra, Barcelona, Spain

Received December 6, 2012; accepted April 8, 2013

Abstract—In this paper, we give a new construction of resonant normal forms with a small remainder for near-integrable Hamiltonians at a quasi-periodic frequency. The construction is based on the special case of a periodic frequency, a Diophantine result concerning the approximation of a vector by independent periodic vectors and a technique of composition of periodic averaging. It enables us to deal with non-analytic Hamiltonians, and in this first part we will focus on Gevrey Hamiltonians and derive normal forms with an exponentially small remainder. This extends a result which was known for analytic Hamiltonians, and only in the periodic case for Gevrey Hamiltonians. As applications, we obtain an exponentially large upper bound on the stability time for the evolution of the action variables and an exponentially small upper bound on the splitting of invariant manifolds for hyperbolic tori, generalizing corresponding results for analytic Hamiltonians.

MSC2010 numbers: 37J25, 37J40

DOI: 10.1134/S1560354713030040

Keywords: perturbation of integrable Hamiltonian systems, normal forms, splitting of invariant manifolds

1. INTRODUCTION AND MAIN RESULTS

1. Let $n \geq 2$ be an integer, $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ and $B_R$ be the closed ball in $\mathbb{R}^n$, centered at the origin, of radius $R > 0$ with respect to the supremum norm. For $\varepsilon > 0$, we consider a near-integrable Hamiltonian of the form

$$\begin{cases} H(\theta, I) = h(I) + f(\theta, I) \\ |f| \leq \varepsilon \ll 1, \end{cases}$$

where $(\theta, I) \in D_R = \mathbb{T}^n \times B_R$ are angle-action coordinates for the integrable part $h$ and $f$ is a small perturbation in some suitable topology defined by a norm $|\cdot|$. The phase space $D_R$ is equipped with the symplectic structure induced by the canonical symplectic structure on $\mathbb{T}^n \times \mathbb{R}^n = T^*\mathbb{T}^n$. For $\varepsilon = 0$, the system is trivially integrable: the action variables are integrals of motion and their level sets $\{I = I_0\}$, $I_0 \in B_R$, are invariant embedded tori on which the flow is linear with frequency $\nabla h(I_0)$. An important subject of study is the dynamical properties of these systems when $\varepsilon > 0$.

Here we will be interested in the local dynamics of the perturbed system around an invariant torus of the integrable system, which, without loss of generality, we assume is located at $I = 0$. The qualitative and quantitative properties of this invariant torus with a linear flow is then determined by its frequency vector $\omega = \nabla h(0) \in \mathbb{R}^n$. Let us say that a vector subspace of $\mathbb{R}^n$ is rational if it has a basis of vectors with rational (or equivalently, integer) components, and we let $F = F_\omega$ be the smallest rational subspace of $\mathbb{R}^n$ containing $\omega$. If $F = \mathbb{R}^n$, the vector $\omega$ is said to be non-resonant and the dynamics on the invariant torus $T_0 = \{I = 0\}$ is then minimal and uniquely ergodic. If $F$...
is a proper subspace of $\mathbb{R}^n$ of dimension $d$, the vector $\omega$ is said to be resonant and $d$ (respectively $m = n - d$) is the number of effective frequencies (respectively the multiplicity of the resonance): the invariant torus $T_0$ is then foliated into invariant $d$-dimensional tori (which are just images of translates of $F$ under the canonical projection $\pi_0 : \mathbb{R}^n \to T_0$) on which the dynamics is again minimal and uniquely ergodic. The example of a resonant vector to keep in mind, to which the general case can be actually reduced, is $\omega = (\tau, 0) \in \mathbb{R}^n = \mathbb{R}^d \times \mathbb{R}^m$ with $\tau \in \mathbb{R}^d$ non-resonant, for which $F = \mathbb{R}^d \times \{0\}$. From now on, we will assume that $1 \leq d \leq n$ (the case $d = 0$ corresponds to an invariant torus which consists uniquely of equilibrium solutions), and the particular case $d = 1$ will play a special role: in this case, writing $\omega = v$, the torus is foliated by periodic orbits and if $T$ denotes the minimal common period, then $Tv \in \mathbb{Z}^n$, and such a vector $v$ will be called periodic.

In order to study the local dynamics of the perturbed system around such an invariant torus, it is important to quantify the minimal character of this linear flow with frequency $\omega$ (on the whole invariant torus if $d = n$ or on each leaf if $d < n$). In general, to such a vector $\omega$ one can associate a constant $Q_\omega > 0$ and a real-valued function $\Psi_\omega$ defined for all real numbers $Q \geq Q_\omega$, which is non-decreasing and unbounded, by

$$\Psi_\omega(Q) = \max \left\{ |k \cdot \omega|^{-1} \mid k \in \mathbb{Z}^n \cap F, \ 0 < |k| \leq Q \right\},$$

(1.1)

where $\cdot$ denotes the Euclidean scalar product and $|.|$ is the supremum norm for vectors (see, e.g., [2] for much more detailed information, where this function is denoted by $\Psi^w$). Note that by definition, we have

$$|k \cdot \omega| \geq \frac{1}{\Psi_\omega(Q)}, \quad k \in \mathbb{Z}^n \cap F, \ 0 < |k| \leq Q.$$

Let us say that a vector $\omega$ is Diophantine if there exist constants $\gamma > 0$ and $\tau \geq d - 1$ such that $\Psi_\omega(Q) \leq \gamma^{-1}Q^\tau$, and let us denote by $\Omega_d(\gamma, \tau)$ the set of such vectors. For $d = 1$, $\omega = v$ is $T$-periodic and $F = F_v$ is just the real line generated by $v$, so that for any non-zero vector $k \in \mathbb{Z}^n \cap F$, as $k$ is collinear to $v$ and $k$ and $Tv$ are non-zero integer vectors, we have

$$|k \cdot v|^{-1} = |k|^{-1}|v|^{-1} = T|k|^{-1}|Tv|^{-1} \leq T$$

hence $\Psi_v(Q) \leq T$ and therefore any $T$-periodic vector belongs to $\Omega_1(T^{-1}, 0)$.
a second paper [8], we will prove the corresponding result for finitely differentiable Hamiltonians using a similar method (but with fairly different technical details).

We give two applications of our main result. As a first application, we prove local and exponentially long stability estimates for the evolution of the action variables for perturbations of non-linear integrable systems, without any condition on the integrable part. We also prove global and exponentially long stability estimates for perturbations of integrable linear systems, and show that the latter result is essentially optimal. If one is interested in global stability estimates for perturbations of non-linear integrable systems, one needs to impose a geometric condition on the integrable part. It is known that for convex integrable systems, an essentially optimal result can be obtained using only the case \( d = 1 \) (see [13] for analytic systems, [16] for Gevrey systems and [5] for finitely differentiable systems). However, it seems that the general case \( 1 \leq d \leq n \) is needed if one is interested in such global estimates for non-convex integrable Hamiltonians. In [4] and [6], special normal forms adapted to the problem were constructed using only periodic approximations; however, the quantitative results obtained there were very far from being optimal, and in particular our main theorem in this paper gives much more general and precise normal forms that should be useful in trying to also obtain essentially optimal results for perturbations of non-convex integrable systems. As a second application, we will prove an exponentially small upper bound for the splitting of invariant manifolds for a hyperbolic torus.

3. Let us now state precisely our results, starting with the regularity assumption. Recall that \( n \geq 2 \) and \( R > 0 \) have been fixed, and that our phase space is \( \mathcal{D}_R = \mathbb{T}^n \times B_R \), where \( B_R \) is the closed ball in \( \mathbb{R}^n \), centered at the origin, of radius \( R \) with respect to the supremum norm.

Given \( \alpha \geq 1 \) and \( L > 0 \), a real-valued function \( f \in C^\infty(\mathcal{D}_R) \) is \((\alpha,L)\)-Gevrey if, using the standard multi-index notation, we have

\[
|f|_{G^{\alpha,L}(\mathcal{D}_R)} = \sup_{(\mathbf{I},\mathbf{I}) \in \mathbb{N}^n} |f|_{\alpha,\mathbf{I},\mathbf{I},R} < \infty, \quad |f|_{\alpha,\mathbf{I},\mathbf{I},R} = \sum_{p=0}^L |\partial^\mathbf{I} f|_{C_\alpha(\mathcal{D}_R)},
\]

where \( |\cdot|_{C_\alpha(\mathcal{D}_R)} \) is the usual supremum norm for functions on \( \mathcal{D}_R \). The space of such functions, with the above norm, is a Banach space, which we denote by \( G^{\alpha,L}(\mathcal{D}_R) \). In the sequel, we shall forget the dependence on the domain in the notation and simply write \( |\cdot|_{\alpha,L} = |\cdot|_{G^{\alpha,L}(\mathcal{D}_R)} \) and \( |\cdot|_{\alpha,\mathbf{I},\mathbf{I},R} = |\cdot|_{G^{\alpha,L}(\mathcal{D}_R)} \). Next, for \( p \geq 1 \), a vector-valued function \( F \in C^\infty(\mathcal{D}_R,\mathbb{R}^p) \) is \((\alpha,L)\)-Gevrey if \( F = (f_1, \ldots, f_p) \) with \( f_i \in G^{\alpha,L}(\mathcal{D}_R) \) for \( 1 \leq i \leq p \). The space of such functions, which we denote by \( G^{\alpha,L}(\mathcal{D}_R,\mathbb{R}^p) \), is also a Banach space with the norm

\[
|F|_{G^{\alpha,L}(\mathcal{D}_R,\mathbb{R}^p)} = \max_{1 \leq i \leq p} |f_i|_{G^{\alpha,L}(\mathcal{D}_R)}.
\]

Similarly, we shall simply write \( |\cdot|_{\alpha,L} = |\cdot|_{G^{\alpha,L}(\mathcal{D}_R,\mathbb{R}^p)} \). By identifying functions with values in \( \mathbb{T} \) with 1-periodic real-valued functions, we can also define \( G^{\alpha,L}(\mathcal{D}_R,\mathbb{T}^n \times \mathbb{R}^n) \) and \( |\cdot|_{\alpha,L} = |\cdot|_{G^{\alpha,L}(\mathcal{D}_R,\mathbb{T}^n \times \mathbb{R}^n)} \). Finally, we will also use the notation \( G^n = \bigcup_{L \geq 0} G^{\alpha,L} \) to denote the space of functions which are \((\alpha,L)\)-Gevrey for some \( L > 0 \), on appropriate domains.

Note that analytic functions are a particular case of Gevrey functions, as one can check that \( G^{1,L}(\mathcal{D}_R) \) is exactly the space of bounded real-analytic functions on \( \mathcal{D}_R \) which extend as bounded holomorphic functions on the complex domain

\[
V_L(\mathcal{D}_R) = \{(\theta, I) \in (\mathbb{C}^n / \mathbb{Z}^n) \times \mathbb{C}^n \mid |I(\theta)| < L, \ d(I, B_R) < L\},
\]

where \( I(\theta) \) is the imaginary part of \( \theta \), \( |\cdot| \) the supremum norm on \( \mathbb{C}^n \) and \( d \) the associated distance on \( \mathbb{C}^n \).

Since we will be only interested in local properties, it will be enough to consider first linear integrable Hamiltonians, that is \( h(I) = \omega(I) = \omega \cdot I \). As we will see, our result for an arbitrary non-linear integrable Hamiltonian \( h \) will be obtained from the linear case by a straightforward localization procedure. Therefore we shall first consider a Hamiltonian \( H \in G^{\alpha,L}(\mathcal{D}_R) \) of the form

\[
\begin{cases}
H(\theta, I) = \omega(I) + f(\theta, I), \quad (\theta, I) \in \mathcal{D}_R, \\
|f|_{\alpha,L} \leq \varepsilon.
\end{cases}
\]

\[(G1)\]
We denote by \{\ldots\} the Poisson bracket associated to the symplectic structure on \(\mathcal{D}_R\). For any vector \(w \in \mathbb{R}^n\), let \(X^t_w\) be the Hamiltonian flow of the linear integrable Hamiltonian \(l_w(I) = w \cdot I\), and given any \(g \in C^1(\mathcal{D}_R)\), we define the average (along the linear flow of frequency \(w\)) of \(g\) by

\[
[g]_w = \lim_{s \to +\infty} \frac{1}{s} \int_0^s g \circ X^t_w dt.
\] (1.3)

Note that \(\{g, l_w\} = 0\) if and only if \(g \circ X^t_w = g\) if and only if \(g = [g]_w\). Recall that the function \(\Psi_{\omega}\) has been defined in (1.1), then we define the functions

\[
\Delta_{\omega}(Q) = Q \Psi_{\omega}(Q), \quad Q \geq Q_{\omega},
\]

\[
\Delta^*_\omega(x) = \sup\{Q \geq Q_{\omega} \mid \Delta_{\omega}(Q) \leq x, \ x \geq \Delta_{\omega}(Q_{\omega})\}.
\]

Our first result is the following.

**Theorem 1.** Let \(H\) be as in (G1). There exist positive constants \(c, c_1, c_2, c_3, c_4, c_5\) and \(C\) that depend only on \(n, R, \omega, \alpha\) and \(L\) such that if

\[
\Delta^*_\omega(ce^{-1}) \geq c_1,
\] (1.4)

then there exists a symplectic map \(\Phi \in G^\alpha, \tilde{L}(\mathcal{D}_{R/2}, \mathcal{D}_R)\), where \(\tilde{L} = CL\), such that

\[
H \circ \Phi = l_\omega + [f]_\omega + g + \tilde{f}, \quad \{g, l_\omega\} = 0
\]

with the estimates

\[
|\Phi - \text{Id}|_{\alpha, \tilde{L}} \leq c_2 \Delta^*_\omega(ce^{-1})^{-1}
\] (1.5)

and

\[
|g|_{\alpha, \tilde{L}} \leq c_3 \varepsilon \Delta^*_\omega(ce^{-1})^{-1}, \quad |\tilde{f}|_{\alpha, \tilde{L}} \leq c_4 \varepsilon \exp \left(-c_5 \Delta^*_\omega(ce^{-1})^{1/2}\right).
\] (1.6)

On the domain \(\mathcal{D}_{R/2}\), the above theorem states the existence of a symplectic Gevrey conjugacy, close to identity, between the original Hamiltonian and a Hamiltonian which is the sum of the integrable part, the average of the perturbation whose size is of order \(\varepsilon\), a resonant part which by definition Poisson commutes with the integrable part and whose size is of order \(\varepsilon(\Delta^*_\omega(ce^{-1}))^{-1}\), and a general part whose size is now exponentially small with respect to \(\Delta^*_\omega(ce^{-1})^{1/2}\). The first terms of this Hamiltonian, namely, \(l_\omega + [f]_\omega + g\), is what is called a resonant normal form, and the last term \(\tilde{f}\) is a “small” remainder.

Concerning the size of this remainder, in [7] an example in the analytic case for \(d = n\) is given to show that this result is “essentially” optimal: we will show below in \(\S 2\) that this example can be easily adapted to the Gevrey case and for any \(1 \leq d \leq n\), and that our estimate for the remainder in the Gevrey case is also “essentially” optimal. The word “essentially” should be understood as follows: given \(\omega\), one can always construct a sequence of positive real numbers \(\varepsilon_j = \varepsilon_j(\omega)\), going to zero as \(j\) goes to infinity, and a sequence of \(\varepsilon_j\)-perturbations \(f_j\), such that the estimate for the remainder \(\tilde{f}_j\) cannot be improved (see the next section for a precise statement). This, of course, does not preclude this estimate from being improved for other values of \(\varepsilon\).

Now in the Diophantine case, the estimates of Theorem 1 can be made more explicit. Indeed, in this case we have the upper bound \(\Psi_{\omega}(Q) \leq \gamma^{-1}Q^*\) which gives the lower bound \(\Delta^*_\omega(ce^{-1}) \geq (c_7\varepsilon^{-1})^{1/\gamma}\). The following corollary is then straightforward.

**Corollary 1.** Let \(H\) be as in (G1), and \(\omega \in \Omega_d(\gamma, \tau)\). There exist positive constants \(c, c_1, c_2, c_3, c_4, c_5\) and \(C\) that depend only on \(n, R, \omega, \alpha\) and \(L\) such that if

\[
\varepsilon \leq c_1^{-1}(1+\tau)\gamma,
\]

then there exists a symplectic map \(\Phi \in G^\alpha, \tilde{L}(\mathcal{D}_{R/2}, \mathcal{D}_R)\), where \(\tilde{L} = CL\), such that

\[
H \circ \Phi = l_\omega + [f]_\omega + g + \tilde{f}, \quad \{g, l_\omega\} = 0
\]
with the estimates
\[ |\Phi - \text{Id}|_{\alpha, \tilde{L}} \leq c_2(c^{-1} \gamma^{-1} \varepsilon)^{\frac{1}{1+r}} \]
and
\[ |g|_{\alpha, \tilde{L}} \leq c_3 \varepsilon (c^{-1} \gamma^{-1} \varepsilon)^{\frac{1}{1+r}}, \quad |f|_{\alpha, \tilde{L}} \leq c_4 \varepsilon \exp \left( -c_5(c \gamma \varepsilon^{-1})^{\frac{1}{\alpha(1+r)}} \right). \]

Even though the size of the remainder is essentially optimal, as we have already explained, in the Diophantine case we believe that the other estimates, concerning the size of the transformation \( \Phi \) and the resonant term \( g \), can be improved. As a matter of fact, these improvements are known in the analytic case for \( d = n \) (see the comments at the end of this section).

4. We can now state a local result for a perturbation of a general non-linear integrable Hamiltonian system, that is for a Hamiltonian \( H \in G^{a,d}(D_R) \) of the form
\[
\begin{cases}
H(\theta, I) = h(I) + f(\theta, I), \quad (\theta, I) \in D_R, \\
\nabla h(0) = \omega, \quad |h|_{\alpha, L} \leq 1, \quad |f|_{\alpha, L} \leq \varepsilon.
\end{cases}
\]

(G2)

For a “small” parameter \( r > 0 \), we will focus on the domain \( D_r = \mathbb{T}^n \times B_r \), which is a neighborhood of size \( r \) of the unperturbed torus \( T_0 = \mathbb{T}^n \times \{0\} \).

Since we are interested in \( r \)-dependent domains in the space of action, the estimates for the derivatives with respect to the actions will have a different size than the one for the derivatives with respect to the angles. To distinguish between them, we will split multi-integers \( l \in \mathbb{N}^{2n} \) as \( l = (l_1, l_2) \in \mathbb{N}^n \times \mathbb{N}^n \) so that \( \partial^l = \partial_{\theta}^{l_1} \partial_I^{l_2} \) and \( |l| = |l_1| + |l_2| \). Let us denote by \( \text{Id}_I \) and \( \text{Id}_\theta \) the identity map in the angle and action space, respectively, and for a function \( F \) with values in \( \mathbb{T}^n \times \mathbb{R}^n \), we shall write \( F = (F_{\theta}, F_I) \).

**Theorem 2.** Let \( H \) be as in (G2). There exist positive constants \( c, c_6, c_7, c_8, c_9 \) and \( c_{10} \) that depend only on \( n, R, \omega, \alpha \) and \( L \), such that if
\[ \sqrt{\varepsilon} \leq r \leq R, \quad \Delta^*_c(cr^{-1}) \geq c_6, \] (1.7)

then there exists a symplectic map \( \Phi \in G^a(D_{r/2}, D_r) \) such that
\[ H \circ \Phi = h + g + [f]_{\omega} + \tilde{f}, \quad \{g, l_\omega\} = 0. \]

Moreover, for \( \tilde{L} = CL \) and any \( l = (l_1, l_2) \in \mathbb{N}^{2n} \), we have the estimates
\[ |\Phi_I - \text{Id}_I|_{\alpha, \tilde{L}, l} \leq c_7 r^{-|l_2|} \Delta^*_c((cr^{-1})^{-1}), \quad |\Phi_\theta - \text{Id}_\theta|_{\alpha, \tilde{L}, l} \leq c_7 r^{-|l_2|} \Delta^*_c((cr^{-1})^{-1}) \]
and
\[ |g|_{\alpha, \tilde{L}, l} \leq c_8 r^2 r^{-|l_2|} \Delta^*_c((cr^{-1})^{-1}), \quad |\tilde{f}|_{\alpha, \tilde{L}, l} \leq c_9 r^2 r^{-|l_2|} \exp \left( -c_{10} \Delta^*_c((cr^{-1})^{-1}) \right). \] (1.8)

(1.9)

The proof of Theorem 2 is an easy consequence of Theorem 1: the Hamiltonian (G2) on the domain \( D_r \) is, after scaling, equivalent to the Hamiltonian (G1) on the domain \( D_1 \), and due to the non-linearity, the small parameter is here naturally \( r \geq \sqrt{\varepsilon} \) instead of \( \varepsilon \). Note that no assumptions on \( h \) are required, since no information on the vector \( \nabla h(I) \) for \( I \in B_r \) is used except of course at \( I = 0 \).

Let us note that the transformation \( \Phi \), and the functions \( g \) and \( \tilde{f} \) are not \((\alpha, \tilde{L})\)-Gevrey because of the factors \( r^{l_2} \) appearing in the estimates (nevertheless, they are still \((\alpha, \tilde{L}_r)\)-Gevrey for some \( \tilde{L}_r \) depending on \( \tilde{L} \) and \( r \), but this will not be used in the sequel). Even though these factors come from our method of proof (which consists in scaling the system so that it reduces to a perturbation of a linear integrable Hamiltonian), we believe this is not an artefact (similar factors appear in the analytic case when one uses Cauchy estimates to control the norm of the derivatives of a function from the norm of the function on a domain of size \( r \)).

Let us also note that in the statement of Theorem 2, one has the freedom to choose any \( r \) such that \( \sqrt{\varepsilon} \leq r \leq R \), provided \( r \) is sufficiently small so that the second part of (1.7) is satisfied. In the
sequel, we shall use the statement only for $r = \sqrt{\varepsilon}$, which appears to be the most natural value, but for some other purposes, for instance, if one is interested in global stability estimates for non-linear systems (as in the Nekhoroshev theorem), it is useful (and perhaps necessary) to use the statement for bigger values of $r$.

As before, in the Diophantine case we can give a more concrete statement.

**Corollary 2.** Let $H$ be as in (G2), and $\omega \in \Omega_\delta(\gamma, \tau)$. There exist positive constants $c$, $c_6$, $c_7$, $c_8$, $c_9$ and $c_{10}$ that depend only on $n, R, \omega, \alpha$ and $L$, such that if

$$\sqrt{\varepsilon} \leq r \leq R, \quad r \leq c_6^{-1+r}\gamma,$$

then there exists a symplectic map $\Phi \in C^\alpha(D_r, D_r)$ such that

$$H \circ \Phi = h + g + [f]_\omega + \tilde{f}, \quad \{g, l_\omega\} = 0.$$ 

Moreover, for $\tilde{L} = CL$ and any $l = (l_1, l_2) \in \mathbb{N}^{2n}$, we have the estimates

$$|\Phi|_{1} - Id_{L}|_{\tilde{\alpha}, \tilde{L}, l} \leq c_7 r^{-|l_2|} (c^{-1} \gamma^{-1} r)^{\frac{1}{1+r}}, \quad |\Phi_{\theta} - Id_{\theta}|_{\tilde{\alpha}, \tilde{L}, l} \leq c_7 r^{-|l_2|} (c^{-1} \gamma^{-1} r)^{\frac{1}{1+r}}$$

and

$$|g|_{\tilde{\alpha}, \tilde{L}, l} \leq c_8 r^2 r^{-|l_2|} (c^{-1} \gamma^{-1} r)^{\frac{1}{1+r}}, \quad |\tilde{f}|_{\tilde{\alpha}, \tilde{L}, l} \leq c_9 r^2 r^{-|l_2|} \exp \left( -c_{10} (c^r r^{-1})^{\frac{1}{1+r}} \right).$$

5. It is perhaps worthwhile to compare our results with previous results, which were restricted mostly to the analytic case $\alpha = 1$ or the periodic case $d = 1$.

In the analytic case $\alpha = 1$, results similar to Theorem 1 (and also to Theorem 2) are contained in [18] and [10] (see also [21]). Note, however, that our results are more accurate than those contained in [18]: in the estimates for $\Phi$ and $g$, the term $\Delta^{\ast}_\varepsilon(c^{-1})^{-1}$ we obtain here is replaced by 1. This does not lead to any improvement if one is interested in global stability estimates for perturbations of non-linear integrable Hamiltonians because the geometry of resonances prevails in this case, but this improvement on the estimate for $\Phi$ is visible for global stability estimates for perturbations of linear integrable Hamiltonians (compare Theorem 3 below with Theorem 5 of [18]) and for local stability estimates for perturbations of non-linear integrable Hamiltonians. Also, the improvement on the estimate for $g$ is important for the application to the splitting of invariant manifolds (see Remark 1). Now concerning Corollary 1 in the special case $\alpha = 1$, one can in fact find better results in [11]: still in the estimates for $\Phi$ and $g$, the term $(c^{-1} \gamma^{-1})^{-1}$ we find is replaced by $c^{-1} \gamma^{-1} \varepsilon$. This discrepancy can be explained as follows. The perturbation theory of quasi-periodic motions essentially relies on solving a certain equation (called homological) which consists in integrating a function along the linear flow of frequency $\omega$. Assuming that the function has zero average along the flow, a formal solution always exists, and it is a smooth convergent solution if and only if $\omega$ is Diophantine. So if one considers non-Diophantine vectors (which is important for several problems, such as Nekhoroshev type estimates), one can only solve an “approximate” equation. The usual approach is to approximate the function by a nicer function (usually a polynomial function), whereas here, as in [2], we just approximate the frequency by a nicer frequency (a periodic frequency). But if we restrict to Diophantine vectors, then one can actually solve the exact equation, and this ultimately leads to better estimates.

Now in the periodic case $d = 1$, the result is known for any $\alpha \geq 1$ (see [16]), and our proof of the general case $1 \leq d \leq n$ crucially relies on it. In fact, in [16], one has a more accurate result, namely, the existence of a formal transformation to a formal normal form, both having Gevrey asymptotics. Such a result is stronger than the existence of a transformation to a normal form with an exponentially small remainder, as the latter is implied by the former (but not the contrary). In the analytic case $\alpha = 1$, these asymptotics were first established in [20].

Finally, there are some related results for $d = n$ and any $\alpha \geq 1$. First in [15], the dynamics in the neighborhood of a Gevrey Lagrangian torus is considered, and a Gevrey normal form is constructed. The system studied there can be brought to a system of the form we considered, though with a special perturbation, and the result obtained in [15] is therefore close but different. Then, in an unpublished manuscript (“Moyennisation Gevrey multiphase”, D. Sauzin in collaboration...
with B. Braaksma and L. Stolovitch, 2006), a result analogous to [16], that is the construction of a
formal transformation to a formal normal form with Gevrey asymptotics, is proved in the case of
a Diophantine vector. As before, these two results give more information but in a different setting,
and, moreover, they are based on solving exact homological equations and so they cannot apply to
non-Diophantine vectors.

6. Finally, let us describe the plan of this paper. In Section 2, we will deduce from Theorem 1
and Theorem 2 exponentially large stability estimates for the evolution of the action variables,
which are global for perturbations of linear integrable systems and only local for perturbations
of non-linear integrable systems. In the first case, that is for perturbations of linear integrable systems,
we will show by an example that these estimates cannot be improved in general. Then in Section 3,
we will use Theorem 2 to prove a result of exponential smallness for the splitting of invariant
manifolds of hyperbolic tori. The proof of Theorem 1 and Theorem 2 will be given in Section 4.
Finally, we will gather in a appendix some technical estimates concerning Gevrey functions that
are used to prove our theorems.

To simplify the notations and improve the readability, when convenient, we shall replace
constants depending on \( n, R, \omega, \alpha \) and \( L \) that can, but need not, be made explicit by a \( \cdot \),
that is an expression of the form \( u \cdot v \) means that there exists a constant \( c > 0 \) that depends on the
above set of parameters, such that \( u \leq cv \). Similarly, we will use the notations \( u \cdot v \) and \( u = \cdot v \).

2. APPLICATION TO STABILITY ESTIMATES

In this section, we will give direct consequences of our normal form theorems (Theorem 1 and
Theorem 2) to the stability of the action variables.

Recall that we have defined \( F \) as the smallest rational subspace containing the vector \( \omega \). The
key point to obtain stability estimates is the following observation. By definition of the Poisson
Bracket, the equalities \( \{ [f]_\omega, l_\omega \} = \{ g, l_\omega \} = 0 \) imply that \( l_\omega \) is a first integral of the normal form
\( l_\omega + [f]_\omega + g \) in the statement of Theorem 1 (or a first integral of \( h + [f]_\omega + g \), since \( h \) is integrable,
in the statement of Theorem 2). Therefore the action variables of all solutions of the normal form
remain constant along \( F \), that is \( \Pi_F(I(t) - I_0) = 0 \) as long as they are defined, if \( I_0 = I(0) \) is an
arbitrary initial action and where we have denoted by \( \Pi_F \) the projection onto the subspace \( F \). Now
if we add a small remainder to the normal form, the quantities \( |\Pi_F(I(t) - I_0)| \), as long as they are
defined, will have small variations for an interval of time which is roughly the inverse of the size
of (the partial derivative with respect to \( \theta \) of) the remainder. Coming back to our original system,
since it is conjugated by a symplectic map which is close to identity to such a normal form with a
small remainder, the same property remains true as long as the solution stays on the domain where
the conjugacy is defined.

7. Let us state precisely the result, starting with a perturbation of a linear integrable
Hamiltonian as in (G1).

**Theorem 3.** Under the notations and assumptions of Theorem 1, let \((\theta(t), I(t))\) be a solution of
the Hamiltonian system defined by \( H \), with \((\theta_0, I_0) \in D_{R/4} \) and let \( T_0 \) be the smallest \( t \in [0, +\infty] \)
such that \((\theta(t), I(t)) \notin D_{R/4} \). Then we have the estimates

\[
|\Pi_F(I(t) - I_0)| \leq \delta, \quad |t| < \min \left\{ T_0, \delta \varepsilon^{-1} \exp \left( -\Delta_\omega^*(\varepsilon^{-1})^{\frac{1}{n}} \right) \right\}
\]

for any \( (\Delta_\omega^*(\varepsilon^{-1}))^{-1} < \delta < 1 \). Moreover, if \( F = \mathbb{R}^n \), then

\[
|I(t) - I_0| \leq \delta, \quad |t| < \delta \varepsilon^{-1} \exp \left( -\Delta_\omega^*(\varepsilon^{-1})^{\frac{1}{n}} \right).
\]

**Corollary 3.** Under the notations and assumptions of Corollary 1, let \((\theta(t), I(t))\) be a solution of
the Hamiltonian system defined by \( H \), with \((\theta_0, I_0) \in D_{R/4} \) and let \( T_0 \) be the smallest \( t \in [0, +\infty] \)
such that \((\theta(t), I(t)) \notin D_{R/4} \). Then we have the estimates

\[
|\Pi_F(I(t) - I_0)| \leq \delta, \quad |t| < \min \left\{ T_0, \delta \varepsilon^{-1} \exp \left( -\gamma \varepsilon^{-1} \frac{1}{\omega(n+2)} \right) \right\}
\]
for any \((\gamma^{-1} \varepsilon)^{1+r} < \delta < 1\). Moreover, if \(F = \mathbb{R}^n\), then
\[
|I(t) - I_0| \leq \delta, \quad |t| < \delta \varepsilon^{-1} \exp \left(\frac{(\gamma \varepsilon^{-1})^{1/(1+r)}}{\alpha}\right).
\]

Note that we have exponential stability for the action variables only in the case \(F = \mathbb{R}^n\), that is when \(\omega\) is non-resonant, since in this case we can ensure that \(T_0\) is at least exponentially large. In general, the stability result is only partial since it involves the projection of the action components onto \(F\), and the time \(T_0\), which is easily seen to be always at least of order \(\delta \varepsilon^{-1}\), cannot be much larger in general (it is easy to construct examples for which the action variables of a solution drift along the orthogonal complement of \(F\) with a speed of order \(\varepsilon\)).

Note that in the particular case where \(\delta = (\Delta^*_\omega(\varepsilon^{-1}))^{-1}\), we have
\[
|\Pi_F(I(t) - I_0)| < (\Delta^*_\omega(\varepsilon^{-1}))^{-1}
\]
for
\[
|t| < \min \left\{ T_0, (\Delta^*_\omega(\varepsilon^{-1}))^{-1} \varepsilon^{-1} \exp \left(\frac{(\Delta^*_\omega(\varepsilon^{-1}))^{1/4}}{\alpha}\right) \right\}
\]
and in the Diophantine case, for \(\delta = (\gamma^{-1} \varepsilon)^{1+r}\), we have
\[
|\Pi_F(I(t) - I_0)| < (\gamma^{-1} \varepsilon)^{1+r}
\]
for
\[
|t| < \min \left\{ T_0, -\gamma^{-1} \varepsilon^{-1} \exp \left(\frac{(\gamma \varepsilon^{-1})^{1/(1+r)}}{\alpha}\right) \right\}.
\]
Still in the Diophantine case, we expect that it may be possible to take \(\delta = \gamma^{-1} \varepsilon\) (see the comments at the end of the first section).

The deduction of Theorem 1.6 from Theorem 1 is very classical, but for completeness we give the details.

**Proof.** Using Theorem 1, it is enough to prove the statement for a solution \((\theta(t), I(t))\) of the Hamiltonian \(\tilde{H} = H \circ \Phi\), with \((\theta_0, I_0) \in \mathcal{D}_{R/2}\) and \(T_0\) being the smallest \(t \in [0, +\infty)\) such that \((\theta(t), I(t)) \notin \mathcal{D}_{R/2}\). Indeed, we have \(\tilde{H} = H \circ \Phi\), where \(\Phi : \mathcal{D}_{R/2} \to \mathcal{D}_{R}\) is symplectic. Moreover, from (1.5), we have the estimate
\[
|\Phi - \text{Id}|_{C^0(\mathcal{D}_{R/2})} \leq |\Phi - \text{Id}|_{|\alpha, \tilde{L}|} < (\Delta^*_\omega(\varepsilon^{-1}))^{-1}.
\]
so we can ensure that the image of \(\Phi\) contains the domain \(\mathcal{D}_{R/4}\), and since we have assumed that \((\Delta^*_\omega(\varepsilon^{-1}))^{-1} < \delta < 1\), the statement for \(\tilde{H}\) trivially implies the statement for \(H\) (only with different implicit constants). Note that in fact we will not use the first part of (1.6), but only the following obvious consequence of the second part of (1.6):
\[
|\partial_\theta \tilde{f}|_{C^0(\mathcal{D}_{R/2})} < |\tilde{f}|_{|\alpha, \tilde{L}|} \leq \varepsilon \exp \left(\frac{-(\Delta^*_\omega(\varepsilon^{-1}))^{1/4}}{\alpha}\right).
\]
(2.1)

Recall that the equations of motion of \(\tilde{H}\) are
\[
\dot{\theta} = \partial_I \tilde{H}(\theta, I), \quad \dot{I} = -\partial_\theta \tilde{H}(\theta, I)
\]
and using the mean value theorem and the fact that \(l_\omega\) is integrable,
\[
|I(t) - I_0| \leq |t| |\partial_\theta \tilde{H}|_{C^0(\mathcal{D}_{R/2})} \leq |t| |\partial_\theta (\{f\}_\omega + g + \tilde{f})|_{C^0(\mathcal{D}_{R/2})}, \quad |t| \leq T_0.
\]
But \(\{g, l_\omega\} = 0\) is equivalent to \(\partial_\theta g(\theta, I) \in F^\perp\), and similarly for \(\{f\}_\omega\), hence if we project onto \(F\), we have
\[
|\Pi_F(I(t) - I_0)| \leq |t| |\partial_\theta \tilde{f}|_{C^0(\mathcal{D}_{R/2})}, \quad |t| \leq T_0.
\]
Now using (2.1) and

$$|t| < \delta \varepsilon^{-1} \exp \left( - \frac{1}{2} (\Delta^*(\varepsilon^{-1}))^\frac{1}{n} \right)$$

we obtain

$$|\Pi_F(I(t) - I_0)| \leq \delta, \quad |t| < \min \left\{ T_0, \delta \varepsilon^{-1} \exp \left( - \frac{1}{2} (\Delta^*(\varepsilon^{-1}))^\frac{1}{n} \right) \right\}$$

for any $(\Delta^*(\varepsilon^{-1}))^{-1} < \delta < 1$. Finally, note that if $F = \mathbb{R}^n$, the map $\Pi_F$ is the identity so that

$$T_0 \cdot \delta \varepsilon^{-1} \exp \left( - \frac{1}{2} (\Delta^*(\varepsilon^{-1}))^\frac{1}{n} \right)$$

and hence

$$|I(t) - I_0| \leq \delta, \quad |t| < \delta \varepsilon^{-1} \exp \left( - \frac{1}{2} (\Delta^*(\varepsilon^{-1}))^\frac{1}{n} \right).$$

This concludes the proof.

8. Now we will show by an example that the estimates of Theorem 3, and therefore the estimate on the remainder in Theorem 1, are essentially optimal.

**Theorem 4.** Let $\omega \in \mathbb{R}^n \setminus \{0\}$. Then there exist a sequence $(\varepsilon_j)_{j \in \mathbb{N}}$ of positive real numbers and a sequence $(f_j)_{j \in \mathbb{N}}$ of functions in $G^{n,1}(\mathcal{D}_R)$, with

$$\varepsilon_j < |f_j|_{n,1} < \varepsilon_j, \quad \lim_{j \to +\infty} \varepsilon_j = 0,$$

such that for $j > 1$, the Hamiltonian system defined by $H_j = l_\omega + f_j$ has solutions $(\theta(t), I(t))$ which satisfy

$$|t| \varepsilon_j \exp \left( - \frac{1}{2} (\Delta^*(\varepsilon^{-1}))^\frac{1}{n} \right) \leq |\Pi_F(I(t) - I_0)| \leq |t| \varepsilon_j \exp \left( - \frac{1}{2} (\Delta^*(\varepsilon^{-1}))^\frac{1}{n} \right).$$

**Proof.** First recall the following fact from linear algebra: there exists an invertible square matrix $A$ of size $n$ with integer coefficients such that $\omega = A(\bar{\omega}, 0) \in \mathbb{R}^d \times \mathbb{R}^m = \mathbb{R}^n$ with $\bar{\omega} \in \mathbb{R}^d$ non-resonant. So replacing $H$ by $H \circ \Phi_A$, where $\Phi_A : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{T}^n \times \mathbb{R}^n$ is the linear symplectic map given by $\Phi_A(\theta, I) = (A\theta, (A^{-1})^t I)$, we can assume that $\omega = (\bar{\omega}, 0)$ as the statement for $H$ and $H \circ \Phi_A$ are equivalent (up to constants depending on $A$, and so on $n$ and $\omega$). Moreover, dividing $\bar{\omega}$ by $|\pm|\bar{\omega}|$ and reordering its components if necessary, we can also assume that $\bar{\omega} = (1, \bar{\omega}_1, \ldots, \bar{\omega}_{d-1}, 0)$, with $|\bar{\omega}_i| \leq 1$ for $1 \leq i \leq d-1$.

So we have reduced the general case to

$$\omega = (1, \bar{\omega}_1, \ldots, \bar{\omega}_{d-1}, 0) \in \mathbb{R}^d \times \mathbb{R}^m = \mathbb{R}^n, \quad F = \mathbb{R}^d \times \{0\}.$$ 

Let us denote by $(p_j/q_j)_{j \in \mathbb{N}}$ the sequence of the convergents of $\bar{\omega}_1$ for instance. We have the classical inequalities

$$(q_j + q_{j+1})^{-1} < |q_j \bar{\omega}_1 - p_j| < q_{j+1}^{-1}, \quad j \in \mathbb{N},$$

and since $q_{j+1} > q_j$, this gives

$$(2q_{j+1})^{-1} < |q_j \bar{\omega}_1 - p_j| < q_{j+1}^{-1}, \quad j \in \mathbb{N}. \tag{2.2}$$

Now by definition of $\Psi_{\omega}$, we obtain

$$q_{j+1} < \Psi_{\omega}(q_j) < 2q_{j+1}. \tag{2.3}$$

The perturbation $f_j$ will be of the form

$$f_j(\theta, I) = f_j^1(I) + f_j^2(\theta), \quad (\theta, I) \in \mathcal{D}_R.$$ 

First, we choose $f_j^1(I) = v_j \cdot I - \omega \cdot I$, where $v_j = (1, p_j/q_j, \bar{\omega}_2, \ldots, \bar{\omega}_{d-1}, 0)$. We set

$$\varepsilon_j = |\bar{\omega}_1 - p_j/q_j|.$$
From the inequalities (2.2) and (2.3), recalling the definitions of $\Delta_\omega$ and $\Delta^*_\omega$, we have

$$(\Delta_\omega(q_j))^{-1} \leq \epsilon_j < (\Delta_\omega(q_j))^{-1}, \quad \Delta^*_\omega(\epsilon_j^{-1}) < q_j < \Delta^*_\omega(\epsilon_j^{-1}).$$  \hspace{2cm} (2.4)

Now by definition, $|f_j|_{\alpha,L} = \epsilon_j$. Then, if we let $k_j = (p_j, - q_j, 0, \ldots, 0) \in \mathbb{Z}^n$, we define $f_j^2(\theta) = \epsilon_j \mu_j \sin(2\pi k_j \cdot \theta)$ with $\mu_j$ to be chosen. Note that $|k_j| = q_j$ since $|q_j| \geq |p_j|$ (as $|\varpi| \leq 1$). It is easy to estimate

$$\epsilon_j \mu_j \exp(|k_j|^{1/\alpha}) < |f_j^2|_{\alpha,L} \leq \epsilon_j \mu_j \exp(\cdot |k_j|^{1/\alpha})$$

and so

$$\epsilon_j \mu_j \exp(\cdot q_j^{1/\alpha}) < |f_j^2|_{\alpha,L} \leq \epsilon_j \mu_j \exp(\cdot q_j^{1/\alpha}).$$

Therefore we choose $\mu_j = \cdot \exp(- \cdot q_j^{1/\alpha})$ so that $\epsilon_j < |f_j|_{\alpha,L} < \epsilon_j$. Finally, $f_j = f_j^1 + f_j^2$ so $\epsilon_j < |f_j|_{\alpha,L} < \epsilon_j$, and $\epsilon_j \to 0$ when $j \to +\infty$. Now we can write the Hamiltonian

$$H_j(\theta, I) = I_\omega(I) + f_j(\theta, I)$$

$$= \omega \cdot I + v_j \cdot I - \omega \cdot I + \epsilon_j \mu_j \sin(2\pi k_j \cdot \theta)$$

$$= v_j \cdot I + \epsilon_j \mu_j \sin(2\pi k_j \cdot \theta)$$

and as $k_j, v_j = 0$, the associated system is easily integrated:

$$\begin{cases} 
\theta(t) = \theta_0 + tv_j \in \mathbb{Z}^n \\
I(t) = I_0 - t \epsilon_j \mu_j \sin(2\pi k_j \cdot \theta) 
\end{cases}$$

Choosing any solution with an initial condition $(\theta_0, I_0)$ satisfying $k_j, \theta_0 = 0$, $\cos(2\pi k_j \cdot \theta_0) = 1$ and using the fact that $|k_j| = q_j$ and $\mu_j = \cdot \exp(- \cdot q_j^{1/\alpha})$, we obtain

$$|I(t) - I_0| = |t| \epsilon_j q_j \exp(- \cdot q_j^{1/\alpha}).$$

For $j > 1$, $q_j > 1$ and we have $\exp(- \cdot q_j^{1/\alpha}) < q_j \exp(- \cdot q_j^{1/\alpha}) < \exp(- \cdot q_j^{1/\alpha})$ for well-chosen implicit constants, hence

$$|t| \epsilon_j \exp(- \cdot q_j^{1/\alpha}) < |I(t) - I_0| < |t| \epsilon_j \exp(- \cdot q_j^{1/\alpha}).$$

Using (2.4) gives

$$|t| \epsilon_j \exp\left(- \cdot \left(\Delta^*_\omega(\epsilon_j^{-1}) \right)^{\frac{1}{\alpha}}\right) < |I(t) - I_0| < |t| \epsilon_j \exp\left(- \cdot \left(\Delta^*_\omega(\epsilon_j^{-1}) \right)^{\frac{1}{\alpha}}\right)$$

and as the vector $k_j$ belongs to $F = \mathbb{R}^d \times \{0\}$, the solution we have constructed satisfies $\Pi_F(I(t) - I_0) = I(t) - I_0$. This concludes the proof.

9. For a perturbation of a non-linear integrable Hamiltonian as in (G2), we have a stability result similar to Theorem 3.

**Theorem 5.** Under the notations and assumptions of Theorem 2, let $(\theta(t), I(t))$ be a solution of the Hamiltonian system defined by $H$, with $(\theta_0, I_0) \in \mathcal{D}_{r/4}$ and let $T_0$ be the smallest $t \in [0, +\infty]$ such that $(\theta(t), I(t)) \notin \mathcal{D}_{r/4}$. Then we have the estimates

$$|\Pi_F(I(t) - I_0)| \leq \delta, \quad |t| < \min\left\{T_0, \delta r^{-2} \exp\left(\Delta^*_\omega(\epsilon_0^{1/\alpha})\right)\right\}$$

for any $r(\Delta^*_\omega(\epsilon_0^{-1}))^{-1} < \delta < r$. Moreover, if $F = \mathbb{R}^n$, then

$$|I(t) - I_0| \leq \delta, \quad |t| < \delta r^{-2} \exp\left(\Delta^*_\omega(\epsilon_0^{1/\alpha})\right).$$
Corollary 4. Under the notations and assumptions of Corollary 2, let \((\theta(t), I(t))\) be a solution of the Hamiltonian system defined by \(H\), with \((\theta_0, I_0) \in \mathcal{D}_{r/4}\) and let \(T_0\) be the smallest \(t \in [0, +\infty]\) such that \((\theta(t), I(t)) \notin \mathcal{D}_{r/4}\). Then we have the estimates
\[
|\Pi_r(I(t) - I_0)| < \delta, \quad |t| < \min \left\{ T_0, \delta r^{-2} \exp \left( (\gamma r^{-1})^{\frac{1}{\alpha + \tau}} \right) \right\}
\]
for any \(r(\gamma r)^{\frac{1}{\alpha + \tau}} < \delta < r\). Moreover, if \(F = \mathbb{R}^n\), then
\[
|I(t) - I_0| < \delta, \quad |t| < \delta r^{-2} \exp \left( (\gamma r^{-1})^{\frac{1}{\alpha + \tau}} \right).
\]

The proof of Theorem 5 is entirely similar to the proof of Theorem 3, replacing the use of Theorem 1 by the use of Theorem 2, since the only information we need to derive these estimates from the resonant normal form with a remainder are \(C^0\) estimates for the distance to the identity (with respect to the action variables) of the conjugacy \(\Phi\) and \(C^0\) estimates for the partial derivatives (with respect to the action variables) of the remainder \(\tilde{f}\), and this information is contained in Theorem 2. Therefore we shall not repeat the details.

These local estimates of stability (or their variants) are at the basis of global (in phase space) estimates of stability for perturbations of non-linear integrable systems, provided that the integrable part satisfies some geometric assumption. However, we do not know if it is possible to construct an example such as in Theorem 4 that would show that the estimates of Theorem 5 are “essentially” optimal.

3. APPLICATION TO THE SPLITTING OF INVARIANT MANIFOLDS

In this section, we apply our normal form results to a different but ultimately related problem, which is the so-called “splitting” of invariant manifolds (or “splitting” of “separatrices”). Roughly speaking, if a Hamiltonian system \(H\), as in (G2), has a suitable “hyperbolic” invariant torus for which the stable and unstable manifolds intersect, the problem is to evaluate in some sense the “angle” between these invariant manifolds at the intersection point. This is an important problem in itself, and this is also closely related to the problem of the speed of instability for the action variables (which is known as the speed of Arnold diffusion).

The general principle is that the “splitting” is exponentially small for analytic systems, and the literature on the subject is huge. Here we shall closely follow [12], Chapter §2, where an approach to obtaining exponentially small upper bounds in the analytic case is given based on normal form techniques. The results contained in [12] are quite general, as they are valid for any number of degrees of freedom, any Diophantine frequency vector and without any restriction on the perturbation, assuming that the torus exists and that its stable and unstable manifolds intersect. However, because of this great generality, these results are not very accurate, they only give an upper bound with a reasonable value for the exponent in the exponential factor. Much more accurate results, such as an asymptotic formula and reasonable values for the other constants involved, can be obtained in much more restricted situations (essentially for two degrees of freedom, specific frequency vectors and specific perturbations, see [3] for some recent results and references).

Our aim here is to generalize the results of [12] for Hamiltonian systems which are only Gevrey regular. We will also assume the existence of a “hyperbolic” torus, together with the property that its invariant manifolds intersect (that is, the existence of homoclinic orbits), simply because conditions that ensure the existence of these objects are well-known (we will quote some of these conditions).

10. We will not try to give an abstract definition of a hyperbolic torus for a Hamiltonian system, first because we will deal with rather concrete examples, but also because we could not find any satisfactory abstract definition (for instance, one which would ensure the existence and uniqueness of stable and unstable manifolds, we refer to [12] and [9] for some attempts).

Now let us consider the setting as described in (G2). Let \(T_0 = \mathbb{T}^n \times \{0\}\) be the
in the proof of Theorem 4), we may already assume that \( \omega = (\varpi, 0) \in \mathbb{R}^d \times \mathbb{R}^m = \mathbb{R}^n \) with \( \varpi \in \mathbb{R}^d \) non-resonant.

If \( \omega \in \Omega_n(\gamma, \tau) \), by KAM theory this torus will persist under any sufficiently small and regular perturbation, provided that \( h \) is non-degenerate (in the analytic case, the assumption that \( \omega \in \Omega_n(\gamma, \tau) \) can be weakened). This is not true if \( d \leq n - 1 \); however, under appropriate assumptions on the system, one can still apply KAM theory to prove the existence of an invariant \( d \)-dimensional torus, with frequency \( \varpi \), which is hyperbolic in the sense that it possesses stable and unstable manifolds. Moreover, such a tori will be isotropic and its asymptotic manifolds will be Lagrangian. If the stable and unstable manifolds intersect, following [12], we can define a symmetric matrix of size \( n \) at a given homoclinic point, called a splitting matrix, the eigenvalues of which are called the splitting angles. Our result is that there exist at least \( d \) splitting angles which are exponentially small (see Theorem 7). The proof of this result will be analogous to the one in [12]. However, in [12], it seems that the normal forms they used, which are taken from [18], are not accurate enough to derive the result they claimed (see Remark 1).

Before stating precisely the result, we need some rather lengthy preparations in order to introduce our definitions and assumptions (the validity of our assumptions will be briefly discussed after the statement of Theorem 7).

First let us split our angle-action coordinates \( (\theta, I) \in D_R = T^n \times B_R \) accordingly to \( \omega = (\varpi, 0) \in \mathbb{R}^d \times \mathbb{R}^m = \mathbb{R}^n \): we write \( \theta = (\theta_1, \theta_2) \in \mathbb{T}^d \times T^n \) and \( I = (I_1, I_2) \in B^d_R \times B^m_R \), where \( B^d_R = B_R \cap \mathbb{R}^d \) and \( B^m_R = B_R \cap \mathbb{R}^m \). We shall always assume here that \( m \geq 1 \), that is \( d \leq n - 1 \). Then Theorem 2 states that for the value \( r = 2\sqrt{\varepsilon} \), if

\[
\Delta_{\varepsilon}^{\text{sym}}(\cdot; \varepsilon^{-1}) \cdot 1,
\]

there exists a symplectic map \( \Phi \in G^s(D_{r/2}, D_r) \), satisfying the estimate (1.8), such that

\[
H \circ \Phi = h + [f]_\omega + g + \tilde{f}, \quad [f]_\omega(\theta_2, I) = \int_{\theta_1 \in \mathbb{T}^d} f(\theta, I) d\theta_1, \quad g(\theta, I) = g(\theta_2, I)
\]

with \( g \) and \( \tilde{f} \) satisfying the estimates (1.9). As in the proof of Theorem 2, it will be more convenient to use a rescaled version of the Hamiltonian \( H \circ \Phi \). Consider the map

\[
\sigma : (\theta, I) \mapsto (\theta, \sqrt{\varepsilon} I)
\]

which sends the domain \( D_1 \) onto \( D_{r/2} \), and let

\[
\mathcal{H} = \varepsilon^{-1}(H \circ \Phi \circ \sigma) = \varepsilon^{-1} h \circ \sigma + \varepsilon^{-1}[f]_\omega \circ \sigma + \varepsilon^{-1} g \circ \sigma + \varepsilon^{-1} \tilde{f} \circ \sigma
\]

be the rescaled Hamiltonian, which is defined on \( D_1 \). Note that in the proof of Theorem 2, we used the same scaling map but the scaling factor was \( \varepsilon^{-1} \) in order to bring the linear part of \( h \) to order one, while here the scaling factor is \( \varepsilon^{-1} \) since it is the quadratic part of \( h \) we want to bring to order one. Consequently, the solutions of the Hamiltonian system defined by \( \mathcal{H} \) coincide with those of \( H \circ \Phi \) only after scaling time by \( \sqrt{\varepsilon} \); since we will be only interested in invariant and asymptotic manifolds, it makes no difference to consider \( \mathcal{H} \) instead of \( H \circ \Phi \). Now let us define

\[
\lambda = \lambda(\varepsilon) = (\Delta^{\text{sym}}_{\varepsilon}(\cdot; \varepsilon^{-1}))^{-1}, \quad \mu = \mu(\varepsilon) = \exp(- \cdot \lambda(\varepsilon)^{-\frac{1}{2}}) = \exp(- \cdot \lambda^{-\frac{1}{2}}).
\]

For any \( l = (l_1, l_2) \in \mathbb{N}^{2n} \), we have

\[
\partial_l^j (g \circ \sigma) = \partial_j^l \partial_l^j (g \circ \sigma) = \sqrt{\varepsilon}^{l_2} |(\Delta \omega^1) | \partial_l^j g \circ \sigma
\]

so

\[
|\partial_l^j (g \circ \sigma)|_{C^0(D_1)} \leq \sqrt{\varepsilon}^{l_2} |\partial_l^j g|_{C^0(D_{r/2})}
\]

and therefore

\[
|g \circ \sigma|_{\alpha, \beta, l} \leq \sqrt{\varepsilon}^{l_2} |g|_{\alpha, \beta, l'}.
\]
Now using (1.9) gives
\[ |g \circ \sigma|_{\alpha,\tilde{L},I} < \sqrt{\varepsilon^2} \varepsilon (\sqrt{\varepsilon})^{-|I|} \lambda = \varepsilon \lambda \]
and hence
\[ |g \circ \sigma|_{\alpha,\tilde{L}} < \varepsilon \lambda. \tag{3.2} \]

In the same way,
\[ |\tilde{f} \circ \sigma|_{\alpha,\tilde{L}} < \varepsilon \mu. \tag{3.3} \]

We will further decompose \( h \) and \( [f]_\omega \) as follows. Without loss of generality, we can assume that \( h(0) = 0 \). Then by Taylor’s formula, we can expand \( h \) at \( I = 0 \) at order 2:
\[ h(\sqrt{\varepsilon} I) = \sqrt{\varepsilon} \varepsilon \cdot I_1 + \varepsilon AI_1 \cdot I_1 + \varepsilon BI_1 \cdot I_2 + \varepsilon CI_1 \cdot I_2 + \varepsilon \varepsilon \varepsilon R_h(I), \]
where
\[ R_h(I) = 2^{-\frac{1}{2}} \int_0^1 (1-t)^2 \varepsilon^3 h(t \sqrt{\varepsilon} I)(I, I, I)dt, \]

\( A \) and \( B \) are square matrices of size \( d \) and \( m \), respectively, and \( C \) is a matrix of size \( m \times d \).

The rectangular term represented by \( CI_1 \cdot I_2 \) is slightly inconvenient, so we will assume:

(A.1) \( C = 0 \).

Then we expand \( [f]_\omega \) at \( I = 0 \) at order 0:
\[ [f]_\omega(\theta_2, \sqrt{\varepsilon} I) = [f]_\omega(\theta_2, 0) + \sqrt{\varepsilon} \int_0^1 \nabla I [f]_\omega(\theta_2, t\sqrt{\varepsilon} I)Idt = \tilde{V}(\theta_2) + \sqrt{\varepsilon} R_{[f]_\omega}(\theta_2, I). \]

and so
\[ \varepsilon^{-1} [f]_\omega \circ \sigma = \varepsilon^{-1} \tilde{V} + \sqrt{\varepsilon}^{-1} R_{[f]_\omega} = \tilde{V} + \sqrt{\varepsilon}^{-1} R_{[f]_\omega}. \]

Thus we have
\[ H(\theta, I) = \sqrt{\varepsilon}^{-1} \varepsilon \cdot I_1 + AI_1 \cdot I_1 + BI_2 \cdot I_2 + V(\theta_2) + (\sqrt{\varepsilon} R_h(I) + \sqrt{\varepsilon}^{-1} R_{[f]_\omega}(\theta_2, I) + \varepsilon^{-1} f(\theta_2, \sqrt{\varepsilon} I)) + \varepsilon^{-1} f(\theta, \sqrt{\varepsilon} I). \]

Applying Lemma 7 (respectively, with \( s = 3 \) and \( s = 1 \)), we obtain \( |R_h|_{\alpha, L'} < 1 \) and \( |R_{[f]_\omega}|_{\alpha, L'} < \varepsilon \), for \( L' = \tilde{L}/2 \), so that the first two terms in the parenthesis above (which are the Hamiltonians independent of \( \theta_1 \)) are of order \( \sqrt{\varepsilon} \). However, the last term in this parenthesis is of order \( \lambda \gg \varepsilon \) by (3.2), so we define
\[ R = \lambda^{-1} \sqrt{\varepsilon} R_h + \lambda^{-1} \sqrt{\varepsilon}^{-1} R_{[f]_\omega} + \lambda^{-1} \varepsilon^{-1} g \circ \sigma, \]

so that the Hamiltonian in the parenthesis above is \( \lambda R \), with \( |R|_{\alpha, L'} < 1 \). Let us also write
\[ F = \varepsilon^{-1} \mu^{-1} \tilde{f} \circ \sigma \]
so that the last term in the expression of \( H \) is \( \mu F \), with \( |F|_{\alpha, L'} \leq |F|_{\alpha, \tilde{L}} < 1 \) by (3.3). Finally, we obviously have \( |V|_{\alpha, L'} \leq 1 \).

Now recalling the dependence of \( \lambda \) and \( \mu \) in the notation, we have obtained a Hamiltonian \( H = H_{\lambda, \mu} \), defined on \( D_1 \), of the form
\[
\begin{align*}
H_{\lambda, \mu}(\theta, I) &= H_\lambda(\theta_2, I) + \mu F(\theta, I), \quad |F|_{\alpha, L'} < 1 \\
H_\lambda(\theta_2, I) &= H^{\text{av}}(\theta_2, I) + \lambda R(\theta_2, I), \quad |R|_{\alpha, L'} < 1 \\
H^{\text{av}}(\theta_2, I) &= \sqrt{\varepsilon}^{-1} \varepsilon \cdot I_1 + AI_1 \cdot I_1 + BI_2 \cdot I_2 + V(\theta_2), \quad |V|_{\alpha, L'} \leq 1.
\end{align*}
\]
they both depend on $\varepsilon$ as in (3.1). One has to consider $\mathcal{H}_{\lambda, \mu}$ as an arbitrary $\mu$-perturbation of $\mathcal{H}_{\lambda} = \mathcal{H}_{\lambda, 0}$, and $\mathcal{H}_{\lambda}$ as a special $\lambda$-perturbation of the “averaged” system $\mathcal{H}^{av}$ (special because the perturbation $R$ is independent of $\theta_1$). Note that the averaged system can be further decomposed as a sum of two Hamiltonians

$$\mathcal{H}^{av}(\theta_2, I) = K(I_1) + P(\theta_2, I_2),$$

where $K(I_1) = \sqrt{\varepsilon^{-1}} \omega \cdot I_1 + AI_1 \cdot I_1$ is a completely integrable system on $D^d_1 = \mathbb{T}^d \times B^d_1$ and $P(\theta_2, I_2) = BI_2 \cdot I_2 + V(\theta_2)$ is a mechanical system (or a “multidimensional pendulum”) on $D^m_1 = \mathbb{T}^m \times \mathbb{R}^m$. We make the following assumption on the mechanical system:

(A.2) The matrix $B$ is positive definite (or negative definite), and the function $V: \mathbb{T}^m \rightarrow \mathbb{R}$ has a non-degenerate maximum (or minimum).

Without loss of generality, we may assume that $V$ reaches its maximum at $\theta_2 = 0$, so that $O = (0, 0) \in D^m_1$ is a hyperbolic fixed point for the Hamiltonian flow generated by $P$. This, in turn, implies that, for any $I^* \in B^d_1$, the set $T(I^*) = \{I_1 = I^*_1\} \times O$ is a $d$-dimensional torus invariant for the averaged system, and it is hyperbolic in the sense that it has $C^1$ stable and unstable manifolds

$$W^{\pm}(T(I^*)) = \{I_1 = I^*_1\} \times W^{\pm}(O),$$

where $W^{\pm}(O)$ are the stable and unstable manifolds of $O$, which are Lagrangian. In particular, the torus $T(0)$ is quasi-periodic with frequency $\sqrt{\varepsilon^{-1}} \omega$.

Now this picture is easily seen to persist if we move from $\mathcal{H}^{av}$ to $\mathcal{H}_{\lambda}$. Indeed, since $\mathcal{H}_{\lambda}$ is still independent of $\theta_1$, the level sets of $I_1$ are still invariant, hence for a given $I^*_1 \in B^d_1$, the Hamiltonian flow generated by the restriction of $\mathcal{H}_{\lambda}$ to $\{I_1 = I^*_1\} \times D^m_1$ (considered as a flow on $D^m_1$ depending on $I^*_1$) is a $\lambda$-perturbation of the Hamiltonian flow generated by $P$; as a consequence, it has a hyperbolic fixed point $O_{\lambda}(I^*_1) \in D^m_1$ which is $\lambda$-close to $O$, for $\lambda$ small enough. Hence $T_{\lambda}(I^*) = \{I_1 = I^*_1\} \times O_{\lambda}(I^*_1)$ is invariant under the Hamiltonian flow of $\mathcal{H}_{\lambda}$, and it is hyperbolic with Lagrangian stable and unstable manifolds

$$W^{\pm}(T_{\lambda}(I^*)) = \{I_1 = I^*_1\} \times W^{\pm}(O_{\lambda}(I^*_1)),$$

The torus $T_{\lambda}(0)$ is still quasi-periodic with frequency $\sqrt{\varepsilon^{-1}} \omega$.

**Remark 1.** Let us point out here that in [12], it is stated incorrectly that the size of the Hamiltonian we called $AR$ is of order $\sqrt{\varepsilon}$ (as we mentioned above, this Hamiltonian is composed of three terms and only two of them are of order $\sqrt{\varepsilon}$, the last one being of order $\lambda(\varepsilon)$). Now in [12], normal forms taken in [18] were used, which contains estimates that are less accurate than ours (as we have already explained, $\lambda(\varepsilon)$ is only of order one in [18]) and, consequently, these estimates do not allow one to show the existence of the invariant torus for the system we called $\mathcal{H}_{\lambda}$.

Our next assumption concerns the persistence of the torus $T_{\lambda}(0)$, as well as its stable and unstable manifolds, when we move from $\mathcal{H}_{\lambda}$ to $\mathcal{H}_{\lambda, \mu}$:

(A.3) For any $0 \leq \lambda < 1$ and $0 \leq \mu < \lambda$, the system $\mathcal{H}_{\lambda, \mu}$ has an invariant torus $T_{\lambda, \mu}$, with $T_{\lambda, 0} = T_{\lambda} = T_{\lambda}(0)$, of frequency $\sqrt{\varepsilon^{-1}} \omega$, with $C^1$ stable and unstable manifolds $W^{\pm}(T_{\lambda, \mu})$ which are exact Lagrangian graphs over fixed relatively compact domains $U^{\pm} \subseteq \mathbb{T}^n$. Moreover, $W^{\pm}(T_{\lambda, \mu})$ are $\mu$-close to $W^{\pm}(T_{\lambda})$ for the $C^1$-topology.

Let us denote by $S^{\pm}_{\lambda, \mu}$ generating functions for $W^{\pm}(T_{\lambda, \mu})$ over $U^{\pm}$, that is if $V^{\pm} = U^{\pm} \times B_1$, then

$$W^{\pm}(T_{\lambda, \mu}) \cap V^{\pm} = \{(\theta_s, I_s) \in V^{\pm} \mid I_s = \partial_\theta S^\pm_{\lambda, \mu}(\theta_s)\},$$

where $S^\pm_{\lambda, \mu}: U^{\pm} \rightarrow \mathbb{R}$ are $C^2$ functions. Since $W^{\pm}(T_{\lambda, \mu})$ are $\mu$-close to $W^{\pm}(T_{\lambda})$ for the $C^1$-topology, the first derivatives of the functions $S^\pm_{\lambda, \mu}$ are $\mu$-close to the first derivatives of $S^{\pm}_{\lambda} = S^{\pm}_{\lambda, 0}$ for the $C^1$-topology.
Then, in order to evaluate the splitting, we need the existence of orbits which are homoclinic to $T_{λ,μ}$:

(A.4) For any $0 ≤ λ < 1$ and $0 ≤ μ < λ$, the set $W^+(T_{λ,μ}) \cap W^-(T_{λ,μ}) \setminus T_{λ,μ}$ is non-empty.

11. We can finally define the notions of the splitting matrix and splitting angles, and state our results.

The set $W^+(T_{λ,μ}) \cap W^-(T_{λ,μ}) \setminus T_{λ,μ}$ is invariant, so it consists of orbits of the Hamiltonian system defined by $H$. Let $γ_{λ,μ}$ be one of these homoclinic orbits, and $p_{λ,μ} = γ_{λ,μ}(0) = (θ_{λ,μ}, I_{λ,μ})$ a homoclinic point. Since $p_{λ,μ}$ is a homoclinic point, $θ_{λ,μ} ∈ U^+ \cap U^−$ and $∂θS^+_{λ,μ}(θ_{λ,μ}) = ∂θS^−_{λ,μ}(θ_{λ,μ})$.

Then we can define the splitting matrix $M(T_{λ,μ}, p_{λ,μ})$ of $T_{λ,μ}$ at the point $p_{λ,μ}$ as the symmetric square matrix of size $n$

$$M(T_{λ,μ}, p_{λ,μ}) = ∂^2_θ(S^+_{λ,μ} - S^-_{λ,μ})(θ_{λ,μ}).$$

Moreover, for $1 ≤ i ≤ n$, we define the splitting angles $α_i(T_{λ,μ}, p_{λ,μ})$ as the eigenvalues of the matrix $M(T_{λ,μ}, p_{λ,μ})$. Note that since the homoclinic point belongs to a homoclinic orbit, at least one of these angles is necessarily zero.

Theorem 6. Let $H_{λ,μ}$ be as in (3.4), and assume that (A.2), (A.3) and (A.4) are satisfied. Then, with the previous notations, we have the estimates

$$|α_i(T_{λ,μ}, p_{λ,μ})| < μ, \quad 1 ≤ i ≤ d.$$  

We have used the fact that the stable and unstable manifolds are exact Lagrangian graphs over some domains in $T^n$ in order to define the splitting matrix and splitting angles, but in fact only the Lagrangian property (and not the exactness or the graph property) is necessary to make those definitions (see [12]).

Now the solutions of the Hamiltonian system defined by $H \circ Φ$ differ from those of $H_{λ,μ}$ only by a time change, so $T_{λ,μ}$ is still an invariant hyperbolic torus for $H \circ Φ$, with the same stable and unstable manifolds. Coming back to our original system, the torus $T_ε = Φ(T_{λ,μ})$ is hyperbolic for $H$, with stable and unstable manifolds $W^±(T_ε) = Φ(W^±(T_{λ,μ}))$, and for $γ_ε = Φ(γ_{λ,μ})$ and $p_ε = Φ(p_{λ,μ})$ we can define a splitting matrix $M(T_ε, p_ε)$ and splitting angles $α_i(T_ε, p_ε)$ for $1 ≤ i ≤ n$.

Theorem 7. Let $H$ be as in (G2), with $r = 2√ε$ satisfying (1.7). Assume that (A.1) is satisfied, and that (A.2), (A.3) and (A.4) are satisfied for the Hamiltonian $H_{λ,μ}$. Then, with the previous notations, we have the estimates

$$|α_i(T_ε, p_ε)| < α(1 + (Δ^1_ε(·) - 1)^1/2) \exp\left(- (γ - 1)^{-1/α}\right), \quad 1 ≤ i ≤ d.$$  

Corollary 5. Let $H$ be as in (G2), with $ω ∈ Ω_2(γ, τ)$ and $r = 2√ε$ satisfying (1.10). Assume that (A.1) is satisfied, and that (A.2), (A.3) and (A.4) are satisfied for the Hamiltonian $H_{λ,μ}$. Then, with the previous notations, we have the estimates

$$|α_i(T_ε, p_ε)| < α(1 + (γ - 1)^{-1/2}) \exp\left(- (γ - 1)^{-1/α}\right), \quad 1 ≤ i ≤ d.$$  

Let us note that the exponent in the exponential factor in our corollary is not far from being optimal, at least for $α > 1$, $d = n - 1$ and for a badly approximable vector $ω ∈ Ω_{2α(n - 1)}(γ, n - 2)$: indeed, in this case we have an exponentially small upper bound with the exponent $(2α(n - 1))^{-1}$, whereas in [14], a sequence of $ε_j$-perturbations, with $ε_j$ going to zero as $j$ goes to infinity, is constructed such that the perturbed system has an invariant hyperbolic torus of dimension $d = n - 1$, with $d - 1 = n - 2$ splitting angles which have an exponential small lower bound with the exponent $(2α(n - 1))^{-1}$.

Let us now briefly discuss the validity of our assumptions (A.1), (A.2), (A.3) and (A.4), referring to [12] for more details. Concerning (A.1), in principle it is just a simplifying assumption and it

REGULAR AND CHAOTIC DYNAMICS Vol. 18 No. 3 2013
can be removed, though we shall not try to do it here. The assumption (A.2) is crucial as it ensures that the averaged system has an invariant hyperbolic torus. Then, using classical KAM theory, (A.3) follows from (A.2) under usual assumptions (there are many references in the analytic case, but we do not know any for non-analytic but sufficiently regular systems). Now concerning the existence of homoclinic orbits (A.4), this is a general assumption: using a variational argument one can prove that the mechanical system has orbits homoclinic to the hyperbolic fixed point and so the averaged system has orbits homoclinic to the hyperbolic torus $T(0)$, then assuming that the stable and unstable manifolds of $T(0)$ intersect transversely along one of these homoclinic orbits inside the energy level, the assumption (A.4) is satisfied, that is this homoclinic orbit for the averaged system can be continued to a homoclinic orbit for the full system.

Note that the results in Theorem 6 and Theorem 7 are just upper bounds on $d$ splitting angles. These splitting angles can actually be equal to zero, as we did not assume that the stable and unstable manifolds intersect transversely inside the energy level. As we have already mentioned, this transversality assumption implies (A.4) provided (A.1), (A.2) and (A.3) are satisfied, and, moreover, under this assumption, one can prove that they are exactly $d$ splitting angles which are non-zero and exponentially small as the other angles can only be polynomially small (we recall that at least one of them is zero, but we could have avoided this situation by taking a Poincaré section and by studying the associated discrete system).

12. Let us now give some details concerning the proof of Theorem 6 and Theorem 7, which follow from simple lemmas (the proof of which, or references, can be found in [12]).

Consider first the Hamiltonian $H_\lambda$, with its hyperbolic invariant torus $T_\lambda = T_{\lambda,0} = T_\lambda(0)$ with stable and unstable manifolds

$$W^\pm(T_\lambda) = \{I_1 = 0\} \times W^\pm(O_\lambda(0)).$$

Let us denote by $S^\pm_{\lambda} = S^\pm_{\lambda,0}$ the corresponding generating functions, and by $p_\lambda = p_{\lambda,0}$ the homoclinic point. From the expression of $W^\pm(T_\lambda)$, one immediately has $\partial_\theta S^\pm_{\lambda} = 0$. Let us write $M_\lambda = M(T_\lambda, p_\lambda)$. The following lemma is then obvious.

**Lemma 1.** The matrix $M_\lambda$ admits the following block decomposition:

$$M_\lambda = \begin{pmatrix} 0 & 0 \\ 0 & M^\perp_\lambda \end{pmatrix},$$

where $M^\perp_\lambda = \partial^2_\theta (S^+_{\lambda} - S^-_{\lambda})(\theta_\lambda)$ is a square symmetric matrix of size $m$.

Note that the matrix $M^\perp_\lambda$ always has zero as an eigenvalue (because we are in a continuous setting), and that the fact that $W^\pm(T_\lambda)$ intersect transversely along the homoclinic orbit inside the energy level is equivalent to zero being a simple eigenvalue of $M^\perp_\lambda$.

Now let us come back to the Hamiltonian $H_{\lambda,\mu}$, and let us write $M_{\lambda,\mu} = M(T_{\lambda,\mu}, p_{\lambda,\mu})$. From assumption (A.3) the second derivatives of the functions $S^\pm_{\lambda,\mu}$ are $\mu$-close to the second derivatives of $S^\pm_{\lambda}$ for the $C^0$-topology. This immediately implies the following lemma, where we also denote by $|.|$ the norm induced by the supremum norm on the space of matrices.

**Lemma 2.** We have the estimate

$$|M_{\lambda,\mu} - M_\lambda| < \mu.$$

Then, in order to use the previous lemma, we need to know how the eigenvalues of a symmetric matrix vary under perturbation. This is the content of the next lemma, where we denote by $d$ the distance on $\mathbb{R}^n$ induced by the supremum norm.

**Lemma 3.** Let $A$ and $A'$ be two symmetric matrices, with spectrum $\text{Spec}(A)$ and $\text{Spec}(A')$. Then we have the estimate

$$d(\text{Spec}(A), \text{Spec}(A')) \leq |A - A'|.$$
Note that the above lemma is quite specific to symmetric matrices, and the fact that the splitting matrix is indeed symmetric in the general case (in our restricted case, this is obvious by definition) ultimately comes from the Lagrangian character of the stable and unstable manifolds.

The proof of Theorem 6 is now a straightforward consequence of Lemma 1, Lemma 2 and Lemma 2.

Let us now come back to the Hamiltonian $H$ as in (G2). It is related to the Hamiltonian $\mathcal{H}_{\lambda(\epsilon),\mu(\epsilon)}$ by a scaling transformation and by a symplectic transformation $\Phi$. The effect of the scaling transformation is simply to multiply the splitting matrix by $\sqrt{\epsilon}$. The effect of the symplectic transformation is more complicated to describe, but the overall effect is contained in the following lemma, where we denote $M_\epsilon = M(T_\epsilon, p_\epsilon)$.

**Lemma 4.** There exist two invertible square matrices $B$ and $C$ of size $n$ such that
$$
M_\epsilon = \sqrt{\epsilon}BM_{\lambda(\epsilon),\mu(\epsilon)}C
$$
where $B$ and $C$ satisfy
$$
|B| < 1 + (\Delta^*_\epsilon (\cdot \sqrt{\epsilon}^{-1}))^{-1}, \quad |C^{-1}| < 1 + (\Delta^*_\epsilon (\cdot \sqrt{\epsilon}^{-1}))^{-1}.
$$

The matrices $B$ and $C$ depend on the differential of the transformation $\Phi$ (we refer to [12], Proposition 1.5.4 for an explicit expression), and the estimates on $B$ and $C$ follow from the estimates on $\partial_\Phi \Phi_0$ and $\partial_\Phi \Phi_I$ contained in (1.8) in Theorem 2.

Finally, we need yet another lemma from linear algebra.

**Lemma 5.** Let $A$ and $A'$ be two symmetric matrices, with $A' = BAC$ for two invertible square matrices $B$ and $C$ of size $n$. Assume that $A$ has $d$ eigenvalues in the interval $(-\delta, \delta)$. Then $A'$ has $d$ eigenvalues in the interval $(-\delta |B||C^{-1}|, \delta |B||C^{-1}|)$.

The proof of Theorem 7 is now an obvious consequence of Theorem 6, Lemma 4 and Lemma 5.

### 4. PROOF OF THE MAIN RESULTS

This section is devoted to the proofs of Theorem 1 and Theorem 2.

**13.** Let us first state an approximation lemma of an arbitrary vector by linearly independent periodic vectors, which was proved in [2]. Recall that a vector $v \in \mathbb{R}^n \setminus \{0\}$ is called periodic if there exists a real number $t > 0$ such that $tv \in \mathbb{Z}^n$. In this case, $T = \inf\{t > 0 \mid tv \in \mathbb{Z}^n\}$ is called the period of $v$, and a periodic vector with period $T$ will be simply called $T$-periodic. It is easy to see that a vector is periodic if and only if its minimal rational subspace is one-dimensional.

Now consider an arbitrary vector $\omega \in \mathbb{R}^n \setminus \{0\}$, and let $d$ be the dimension of its minimal rational subspace $F = F_\omega$. For a given $Q \geq 1$, it is always possible to find a $T$-periodic vector $v \in F$, which is a $Q$-approximation in the sense that $|T\omega - Tv| \leq Q^{-1}$, and for which the period $T$ satisfies the upper bound $T < Q^{d-1}$; this is essentially the content of Dirichlet’s theorem. Then it is not hard to see that there exist not only one, but $d$ linearly independent periodic vectors in $F$ which are $Q$-approximations. Moreover, one can obtain not only linearly independent vectors, but periodic vectors $v_1, \ldots, v_d$ of periods $T_1, \ldots, T_d$ such that the integer vectors $T_1v_1, \ldots, T_dv_d$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^n \cap F$. However, the upper bound on the associated periods $T_1, \ldots, T_d$ is necessarily larger than $Q^{d-1}$, and is given by a function that we call here $\Psi_\omega'$ (once again, see [2] for more precise and general information, but note that in [2] $\Psi_\omega'$ was denoted by $\Psi_\omega$ and $\Psi_\omega$, which we defined in (1.1), was denoted by $\Psi_\omega'$). The main Diophantine result of [2] is that this function $\Psi_\omega'$ is in fact equivalent to the function $\Psi_\omega$, up to constants and for $Q$ large enough. This gives the following result.

**Proposition 1.** Let $\omega \in \mathbb{R}^n \setminus \{0\}$. For any $Q > 1$, there exist $d$ periodic vectors $v_1, \ldots, v_d$, of periods $T_1, \ldots, T_d$, such that $T_1v_1, \ldots, T_dv_d$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^n \cap F$ and for $j \in \{1, \ldots, d\}$,
$$
|\omega - v_j| < (T_j Q)^{-1}, \quad 1 < T_j < \Psi_\omega(Q).
$$
For the proof, we refer to [2], Proposition 2.3. The implicit constants depend only on \( d \) and \( \omega \) (the dependence on \( \omega \) is through its norm \(|\omega|\) and the discriminant of the lattice \( \mathbb{Z}^n \cap F \)).

Now a consequence of the fact that the vectors \( T_1v_1, \ldots, T_dv_d \) form a \( \mathbb{Z} \)-basis of \( \mathbb{Z}^n \cap F \) is contained in the following corollary. For simplicity, we shall write \([\cdot]_{v_1, \ldots, v_d} = [\cdots [\cdot]_{v_1} \cdots]_{v_d} \), where \([\cdot]_w \) has been defined for an arbitrary vector \( w \) in (1.3).

**Corollary 6.** Under the assumptions of Proposition 1, let \( l, \omega(I) = \omega \cdot I + l, v_j(I) = v_j \cdot I \) for \( j \in \{1, \ldots, d\} \). For any \( g \in C^1(D_R) \), we have \([g]_\omega = [g]_{v_1, \ldots, v_d} \) and therefore \([g, l, \omega] = 0 \) if and only if \([g, l, v_j] = 0 \) for any \( j \in \{1, \ldots, d\} \).

**Proof.** Let \( \Lambda = \mathbb{Z}^n \cap F \) and \( dv \) be the Haar measure on the compact quotient group \( F/\Lambda \). The flow \( \chi_t \) leaves invariant the foliation on \( \mathbb{T}^n \) induced by the trivial affine foliation on \( \mathbb{R}^n \) defined by \( F \). Each leaf of this foliation on \( \mathbb{T}^n \) is diffeomorphic to \( F/\Lambda \), and the restriction of \( \chi_t \) to each leaf is uniquely ergodic. By Birkhoff’s ergodic theorem, we have

\[
[g]_\omega = \lim_{s \to +\infty} \frac{1}{s} \int_0^s g \circ \chi_t dt = \int_{\theta \in F/\Lambda} g \circ \chi_t d\theta.
\]

Also, as \( T_jv_j \) is an integer vector, we have

\[
[g]_{v_j} = \lim_{s \to +\infty} \frac{1}{s} \int_0^s g \circ \chi_t dt = \frac{1}{T} \int_0^T g \circ \chi_t dt = \int_0^1 g \circ \chi_t_{T_jv_j} dt.
\]

Using the fact that the vectors \( T_1v_1, \ldots, T_dv_d \) form a \( \mathbb{Z} \)-basis of \( \Lambda \), the first assertion follows easily from these expressions by a change of variables. Now if \([g, l, v_j] = 0 \) for any \( j \in \{1, \ldots, d\} \), then \( g = [g]_{v_j} \) and so by the first assertion, \( g = [g]_\omega \), which means that \([g, l, \omega] = 0 \). Conversely, if \([g, l, \omega] = 0 \), then \( g = [g]_\omega \) and therefore \( g = [g]_{v_1, \ldots, v_d} \) by the first assertion. Since the maps \([\cdot]_{v_j} \) are projectors (that is \([\cdot]_{v_j} = [\cdot]_{v_j} \)), this implies that

\[
[g]_{v_j} = [g]_{v_1, \ldots, v_d, v_j} = [g]_{v_1, \ldots, v_d} = g,
\]

and since they commute, we eventually find \([g]_{v_j} = g \) for any \( j \in \{1, \ldots, d\} \) and therefore \([g, l, v_j] = 0 \) for any \( j \in \{1, \ldots, d\} \).

**14.** Then we shall make use of the statement of Theorem 1 (or Corollary 1) in the particular case \( d = 1 \), that is when \( F \) is one-dimensional, which was proved in [16]. As previously stated, in this situation the vector is in fact periodic so we shall denote it by \( v \), and for any non-zero integer vector \( k \in F \), we have the lower bound \(|k \cdot v| \geq T^{-1} \) and so we will have \( \Delta \geq (T\varepsilon)^{-1} \) in the statement of Theorem 1 (or \( \tau = 0 \) and \( \gamma = T^{-1} \) in the statement of Corollary 1) for this particular case.

For subsequent use, we introduce another parameter \( \nu > 0 \) and consider the Hamiltonian

\[
\begin{aligned}
H(\theta, I) &= l_\nu(I) + s(I) + u(\theta, I), \quad (\theta, I) \in D_R, \\
T &\in \mathbb{Z}, \quad |s|_{\alpha, \nu} \leq \nu, \quad |u|_{\alpha, \nu} \leq \varepsilon.
\end{aligned}
\]

Let us define \( \delta = (2d)^{-1}R \). Then we have the following result.

**Proposition 2.** Let \( H \) be as in (4.1), and assume that

\[
\varepsilon < \nu, \quad T \nu < 1.
\]

Then there exists a symplectic map \( \Phi \in G^{\alpha, L'}(D_R - \delta, D_R) \), with \( C = 1 \) and \( L' = CL \), such that

\[
H \circ \Phi = l_\nu + s + |u|_{\nu} + u' + \tilde{u}, \quad \{u', l_\nu\} = 0
\]

with the estimates

\[
|\Phi - \text{Id}|_{\alpha, L'} < T\nu, \quad |u'|_{\alpha, L'} < \varepsilon T\nu, \quad |\tilde{u}|_{\alpha, L'} < \varepsilon \exp(- \cdot (T\nu)^{-1/\alpha}).
\]
Let us further define $L$ as the perturbation. Since we have assumed that $\varepsilon < \mu$, the size of the perturbation is of order $\mu$, and thus $L$ is integrable, the size of the non-resonant part of the perturbation is of order $\varepsilon$. Note that if we are only interested in the periodic case, then we may take $s = 0$ in (4.1), $\varepsilon = \nu$ and write $u = f$ in the statement of Proposition 2, and this gives exactly the statement of Theorem 1.

For a proof of Proposition 2, we refer to [16], Proposition 3.2 (note that in [16], the size of the perturbation is called $\varepsilon$ instead of $\mu$, and the size of the non-resonant part of the perturbation is called $\varepsilon'$ instead of $\varepsilon$). The implicit constants in the above statement depend only on $n, R, \alpha$ and $L$.

Let us now state a simple algebraic property (in our restricted setting), which complements Proposition 2.

**Proposition 3.** Under the assumptions of Proposition 2, suppose that $l_w$ is a linear integrable Hamiltonian such that $\{u, l_w\} = 0$. Then, in the conclusions of Proposition 2, we have $\{u', l_w\} = 0$.

This property has been used several times (this was first used in [1] and [19]) and is valid under much more general assumptions, we refer to [4] for a proof in the analytic case (see also [6]). Of course, this is a purely algebraic property, and is valid regardless of the regularity of the system.

**15.** Now we can finally prove Theorem 1, which is a straightforward consequence of the more flexible proposition below.

**Proposition 4.** Let $H$ be as in (G1), and $Q \geq 1$. If

$$Q \cdot 1, \quad \varepsilon < \Delta_\omega(Q)^{-1},$$

then there exists a symplectic map $\Phi \in G^{\alpha, \tilde{L}}(\mathcal{D}_{R/2}, \mathcal{D}_R)$, with $\tilde{L} = \cdot L$, such that

$$H \circ \Phi = l_\omega + |f|_\omega + g + \tilde{f}, \quad \{g, l_\omega\} = 0$$

with the estimates

$$|\Phi - \text{Id}|_{\alpha, \tilde{L}} < Q^{-1}, \quad |g|_{\alpha, \tilde{L}} < \varepsilon Q^{-1}, \quad |\tilde{f}|_{\alpha, \tilde{L}} < \varepsilon \exp \left( - \cdot Q^{-1/\alpha} \right).$$

**Proof (Proof of Theorem 1).** We choose

$$Q = \Delta_{\omega}(\varepsilon^{-1})$$

with a well-chosen implicit constant so that the second part of (4.3) is satisfied. Proposition 4 with this value of $Q$ implies Theorem 1, as the first part of (4.3) is satisfied by the threshold (1.4).

It remains to prove Proposition 4.

**Proof (Proof of Proposition 4).** Recall that we are considering $H$ as in (G1). Since $Q \cdot 1$ by the first part of (4.3), we can apply Proposition 1: there exist $d$ periodic vectors $v_1, \ldots, v_d$, of periods $T_1, \ldots, T_d$, such that $T_1 v_1, \ldots, T_d v_d$ form a $Z$-basis of $\mathbb{Z}^n \cap F$ and for $j \in \{1, \ldots, d\}$,

$$|\omega - v_j| < (T_j Q)^{-1}, \quad 1 < T_j < \Psi_\omega(Q).$$

For $j \in \{1, \ldots, d\}$, let us define

$$s_j = l_\omega - l_{v_j}, \quad \nu_j = (T_j Q)^{-1}$$

with a suitable implicit constant so that $|s_j|_{\alpha, L} \leq \nu_j$. Note that $l_\omega = l_{v_j} + s_j$, and that $T_j \nu_j = \cdot Q^{-1}$. Let us further define $L_j = C^j L$ and $R_j = R - j \delta$ so that in particular $R_d = R/2$.

Then we claim that for all $1 \leq j \leq d$, there exists a symplectic map $\Phi_j \in G^{\alpha, L_j}(\mathcal{D}_{R_j}, \mathcal{D}_R)$ such that

$$H \circ \Phi_j = l_\omega + |f|_{v_1, \ldots, v_j} + g_j + f_j, \quad \{g_j, l_{v_i}\} = 0$$

for any $1 \leq i \leq j$, with the estimates

$$|\Phi_j - \text{Id}|_{\alpha, L_j} < Q^{-1}, \quad |g_j|_{\alpha, L_j} < \varepsilon Q^{-1}, \quad |f_j|_{\alpha, L_j} < \varepsilon \exp(- \cdot Q^{-1/\alpha}).$$
The proof of the proposition follows from this claim: it is sufficient to let \( \Phi = \Phi_d \), \( g = g_d \), \( \tilde{f} = f_d \)
and \( \tilde{L} = L_d \), since from Corollary 6, \( [f]_{v_1, \ldots, v_d} = [f]_\omega \) and \( \{g, l_{v_i}\} = 0 \) for \( 1 \leq i \leq d \) is equivalent to
\( \{g, l_{\omega}\} = 0 \).

Now let us prove the claim by induction. For \( j = 1 \), writing \( l_{\omega} = l_{v_1} + s_1 \), this is nothing but
Proposition 2 (up to a change of notations, namely \( s_1 \) instead of \( s \), \( f \) instead of \( u, g_1 \) instead of \( u' \) and \( f_1 \) instead of \( \tilde{u} \)). So assume the statement holds true for some \( 1 \leq j \leq d - 1 \), and let us prove that it
is true for \( j + 1 \). By the inductive assumption, there exists a symplectic map \( \Phi_j \in G^{\alpha, L_j}_{\omega}(D_{R_j}, D_R) \)
such that
\[
H \circ \Phi_j = l_{\omega} + [f]_{v_1, \ldots, v_j} + g_j + f_j, \quad \{g_j, l_{v_i}\} = 0,
\]
for any \( 1 \leq i \leq j \), with the estimates
\[
|\Phi_j - \text{Id}|_{\alpha, L_j} < Q^{-1}, \quad |g_j|_{\alpha, L_j} < \varepsilon Q^{-1}, \quad |f_j|_{\alpha, L_j} < \varepsilon \exp(-\cdot Q^{-1/\alpha}).
\]
Now let us consider the Hamiltonian
\[
H \circ \Phi_j - f_j = l_{\omega} + [f]_{v_1, \ldots, v_j} + g_j = l_{v_{j+1}} + s_{j+1} + [f]_{v_1, \ldots, v_j} + g_j = l_{v_{j+1}} + s_{j+1} + u_j
\]
with \( u_j = [f]_{v_1, \ldots, v_j} + g_j \), so \( |u_j|_{\alpha, L_j} < (\varepsilon + \varepsilon Q^{-1}) < \varepsilon \). We want to apply Proposition 2 to this
Hamiltonian, and we observe that (4.3) implies (4.2): indeed, \( Q > 1 \) implies \( T_j v_j \cdot Q^{-1} < 1 \), whereas
\[
\varepsilon < \Delta_\omega(Q)^{-1} = (Q \Psi_\omega(Q))^{-1} < (T_j Q)^{-1} = \nu_j.
\]
So we can apply Proposition 2: there exists a symplectic map \( \Phi^{j+1} \in G^{\alpha, L_{j+1}}_{\omega}(D_{R_{j+1}}, D_{R_j}) \) such that
\[
(H \circ \Phi_j - f_j) \circ \Phi^{j+1} = l_{v_{j+1}} + s_{j+1} + [u_j]_{v_{j+1}} + u_j' + \tilde{u}_j
\]
\[
= l_{\omega} + [f]_{v_1, \ldots, v_j} + g_j]_{v_{j+1}} + u_j' + \tilde{u}_j
\]
\[
= l_{\omega} + [f]_{v_1, \ldots, v_{j+1}} + [g_j]_{v_{j+1}} + u_j' + \tilde{u}_j
\]
with \( \{u_j', l_{v_{j+1}}\} = 0 \) and the estimates
\[
|\Phi^{j+1} - \text{Id}|_{\alpha, L_{j+1}} < T_j v_{j+1} \cdot Q^{-1}
\]
and
\[
|u_j'|_{\alpha, L_{j+1}} < \varepsilon Q^{-1}, \quad |\tilde{u}_j|_{\alpha, L_{j+1}} < \varepsilon \exp(-\cdot (T_{j+1} v_{j+1})^{-1/\alpha}) = \varepsilon \exp(-\cdot Q^{-1/\alpha}).
\]
Obviously, we have \( |[g_j]_{v_{j+1}}|_{\alpha, L_j} < \varepsilon Q^{-1} \), so we set
\[
\Phi_{j+1} = \Phi_j \circ \Phi^{j+1}, \quad g_{j+1} = [g_j]_{v_{j+1}} + u_j', \quad f_{j+1} = \tilde{u}_j + f_j \circ \Phi^{j+1}.
\]
We have \( \Phi_{j+1} \in G^{\alpha, L_{j+1}}_{\omega}(D_{R_{j+1}}, D_{R_j}) \), and since
\[
|\Phi^{j+1} - \text{Id}|_{\alpha, L_{j+1}} < Q^{-1} < 1,
\]
using Corollary 7 we have the estimate
\[
|\Phi_{j+1} - \text{Id}|_{\alpha, L_{j+1}} \leq |\Phi_j - \text{Id}|_{\alpha, L_j} + |\Phi^{j+1} - \text{Id}|_{\alpha, L_{j+1}} < Q^{-1}.
\]
Now \( \{g_j, l_{v_i}\} = 0 \) for \( 1 \leq i \leq j \) from the induction hypothesis, and obviously we have
\( \{[f]_{v_1, \ldots, v_j}, l_{v_i}\} = 0 \) for \( 1 \leq i \leq j \), hence \( \{u_j, l_{v_i}\} = 0 \) for \( 1 \leq i \leq j \). Then if we apply Proposition 3
with \( l_{\omega} = l_{v_i} \) for \( 1 \leq i \leq j \), we obtain that \( \{u_j', l_{v_i}\} = 0 \) for \( 1 \leq i \leq j \). But \( \{u_j', l_{v_{j+1}}\} = 0 \) and
hence \( \{u_j', l_{v_i}\} = 0 \) for \( 1 \leq i \leq j + 1 \). Moreover, we claim that \( \{g_j, l_{v_i}\} = 0 \) for \( 1 \leq i \leq j \) implies
\( \{g_j, l_{v_{j+1}}\} = 0 \) for \( 1 \leq i \leq j \), and as \( \{[g_j]_{v_{j+1}}, l_{v_{j+1}}\} = 0 \), this gives \( \{[g_j]_{v_{j+1}}, l_{v_i}\} = 0 \) for \( 1 \leq i \leq j + 1 \), and this eventually gives \( \{g_{j+1}, l_{v_i}\} = 0 \) for \( 1 \leq i \leq j + 1 \). To prove the claim, note that since
\( \{l_{v_i}, l_{v_{j+1}}\} = 0 \) for \( 1 \leq i \leq j \), \( l_{v_i} = l_{v_i} \circ X^0_{v_{j+1}} \) and so
\[
\{[g_j]_{v_{j+1}}, l_{v_i}\} = \lim_{s \rightarrow +\infty} \frac{1}{s} \int_0^s \{g_j \circ X^t_{v_{j+1}}, l_{v_i}\} dt
\]
REGULAR AND CHAOTIC DYNAMICS Vol. 18 No. 3 2013
which sends the domain $D$ neighborhood of size $r$ around the origin to obtain

Without loss of generality, we assume that $h$.

Proof (Proof of Theorem 2).

This proves the claim and completes the proof of the proposition.

Finally, from Lemma 6 and (4.4), we have

Let $\Phi = \partial h/\partial l_0$.

Moreover, for any $l = (l_1, l_2) \in \mathbb{N}^{2n}$, we have the estimates

and therefore

This proves the claim and completes the proof of the proposition.

16. Let us now prove Theorem 2, which is an easy consequence of Theorem 1. As before, this will be deduced from the more flexible proposition below.

Proposition 5. Let $H$ be as in (G2), and $Q \geq 1$. If

\[ Q > 1, \quad \varepsilon \leq r^2, \quad r < \Delta_\omega(Q)^{-1}, \quad r \leq R, \tag{4.5} \]

then there exists a symplectic map $\Phi \in G^s(D_{r/2}, D_r)$, such that

\[ H \circ \Phi = h + [f]_\omega + g + \tilde{f}, \quad \{g, l_\omega\} = 0. \]

Moreover, for any $l = (l_1, l_2) \in \mathbb{N}^{2n}$, we have the estimates

\[ |\Phi_I - \text{Id}_I|_{\alpha, \tilde{L}, l} < r^{-|l_2|}Q^{-1}, \quad |\Phi_\theta - \text{Id}_\theta|_{\alpha, \tilde{L}, l} < r^{-|l_2|}Q^{-1} \]

and

\[ |g|_{\alpha, \tilde{L}, l} < r^2 r^{-|l_2|}Q^{-1}, \quad |f|_{\alpha, \tilde{L}, l} < r^2 r^{-|l_2|} \exp \left( - \cdot Q^{-1/\alpha} \right). \]

Proof (Proof of Theorem 2). We choose

\[ Q = \Delta_\omega^*(r^{-1}) \]

with a well-chosen implicit constant so that the third part of (4.5) is satisfied. Proposition 5 with this value of $Q$ implies Theorem 2, as the other conditions of (4.5) are satisfied by (1.7).

Now let us prove Proposition 5.

Proof (Proof of Proposition 5). To analyze our Hamiltonian $H$ in the domain $D_r$, which is a neighborhood of size $r$ around the origin, we rescale the action variables using the map

\[ \sigma : (\theta, I) \longrightarrow (\theta, rI) \]

which sends the domain $D_1$ onto $D_r$, the latter being included in $D_R$ by the last part of (4.5). Let

\[ H' = r^{-1}(H \circ \sigma) \]

be the rescaled Hamiltonian, so $H'$ is defined on $D_1$ and reads

\[ H'((\theta, I) = r^{-1}H(\theta, rI) = r^{-1}h(rI) + r^{-1}f(\theta, rI), \quad (\theta, I) \in D_1. \]

Without loss of generality, we assume that $h(0) = 0$. Now, using Taylor’s formula, we can expand $h$ around the origin to obtain

\[ h(rI) = r\omega \cdot I + r^2 \int_0^1 (1 - t)\nabla^2 h(trI)(I, I)dt = r\omega \cdot I + r^2 h'(I) \]
and so we can write
\[ H' = l_\omega + f' \]
with
\[ f' = rh' + r^{-1}(f \circ \sigma). \]
Now we know that \(|f|_{\alpha, L} \leq \varepsilon\), so that \(|r^{-1}(f \circ \sigma)|_{\alpha, L} \leq r^{-1}\varepsilon\), and using the second part of (4.5), \(|r^{-1}(f \circ \sigma)|_{\alpha, L} \leq r\). Moreover, applying Lemma 7 (with \(s = 2\)) we have \(|h'|_{\alpha, L/2} < |h|_{\alpha, L} < 1\), so that \(|rh'|_{\alpha, L/2} < r\). This eventually gives \(|f'|_{\alpha, L/2} < r\). Now we will apply Proposition 4 to the Hamiltonian \(H' = l_\omega + f'\), defined on the domain \(D_1\) and such that \(|f'|_{\alpha, L/2} < r\). The first and third parts of condition (4.5) imply condition (4.3) with 1 instead of \(R\), \(L/2\) instead of \(L\) and \(r\) instead of \(\varepsilon\), so that there exists a symplectic map \(\Phi' \in G^{\alpha, L}(D_1/2, D_1)\), with \(\tilde{L} = L\), such that
\[ H' \circ \Phi' = l_\omega + [f']_\omega + g' + \tilde{f}', \quad \{g', l_\omega\} = 0 \]
with the estimates
\[ |\Phi' - \text{Id}|_{\alpha, \tilde{L}} < Q^{-1}, \quad |g'|_{\alpha, \tilde{L}} < rQ^{-1}, \quad |\tilde{f}'|_{\alpha, \tilde{L}} < r \exp \left(-\cdot Q^{-1/\alpha}\right). \]
Note that \([f']_\omega = [rh' + r^{-1}(f \circ \sigma)]_\omega = rh' + r^{-1}[f]_\omega \circ \sigma\), so the transformed Hamiltonian can be written as
\[ H' \circ \Phi' = l_\omega + rh' + r^{-1}[f]_\omega \circ \sigma + g' + \tilde{f}'. \]
Now, scaling back to our original coordinates, we define \(\Phi = \sigma \circ \Phi' \circ \sigma^{-1}\), therefore \(\Phi : D_{r/2} \rightarrow D_r\) and
\[ H \circ \Phi = rH' \circ \Phi' \circ \sigma^{-1} \]
\[ = r(l_\omega + rh' + r^{-1}[f]_\omega \circ \sigma + g' + \tilde{f}') \circ \sigma^{-1} \]
\[ = (rl_\omega + r^2h') \circ \sigma^{-1} + [f]_\omega + ry' \circ \sigma^{-1} + r\tilde{f}' \circ \sigma^{-1}. \]
Observe that \((rl_\omega + r^2h') \circ \sigma^{-1} = h\), so we may set
\[ g = ry' \circ \sigma^{-1}, \quad \tilde{f} = r\tilde{f}' \circ \sigma^{-1}, \]
and write
\[ H \circ \Phi = h + [f]_\omega + g + \tilde{f}. \]
The equality \(\{g, l_\omega\} = 0\) is obvious. Now let \(l = (l_1, l_2) \in \mathbb{N}^{2n}\). Observe that from the definition of \(\sigma\) and \(g\),
\[ \partial^I g = \partial_\theta^I \partial_\alpha^I g = rr^{-|l_2|}(\partial^I g') \circ \sigma^{-1} \]
so
\[ |\partial^I g|_{C^0(D_{r/2})} \leq rr^{-|l_2|}|\partial^I g'|_{C^0(D_{r/2})} \]
and therefore
\[ |g|_{\alpha, \tilde{L}, l} \leq rr^{-|l_2|}|g'|_{\alpha, \tilde{L}, l} < r^2r^{-|l_2|}Q^{-1}. \]
Replacing \(g'\) by \(\tilde{f}'\) in the above argument, we obtain
\[ |\tilde{f}'|_{\alpha, \tilde{L}, l} \leq rr^{-|l_2|}|\tilde{f}'|_{\alpha, \tilde{L}, l} < r^2r^{-|l_2|} \exp \left(-\cdot Q^{-1/\alpha}\right). \]
Finally, writing \(\Phi = (\Phi_\theta, \Phi_I)\) and \(\Phi' = (\Phi_\theta, \Phi'_I)\), we observe that \(\Phi_\theta = \Phi_\theta' \circ \sigma^{-1}\) and \(\Phi_I = r\Phi'_I \circ \sigma^{-1}\), which immediately gives the estimates for \(\Phi_\theta\) and \(\Phi_I\). This concludes the proof.
In this section we recall some technical estimates concerning Gevrey functions, taken from [16], which are used in the proofs.

Note that our Gevrey norm differs from the one in [16], in which the authors used

\[ \|H\|_{G^{\alpha,L}(\mathcal{D}_R)} = \sum_{l \in \mathbb{N}^{2n}} L^{l[\alpha]}(1)^{-\alpha} |\partial^l H|_{C^0(\mathcal{D}_R)} < \infty \]

instead of (1.2). However, these norms are “almost equivalent” in the sense that obviously we have \( \|H\|_{G^{\alpha,L}(\mathcal{D}_R)} \leq \|H\|_{G^{\alpha,L}(\mathcal{D}_R)} \) while, for instance, \( \|H\|_{G^{\alpha,L}(\mathcal{D}_R)} \leq 2^\alpha (2^\alpha - 1)^{-1} \|H\|_{G^{\alpha,L}(\mathcal{D}_R)} \). So this change only affects the implicit constants.

First recall that our proof of Theorem 1 proceeds by induction on \( 1 \leq j \leq d \), starting with the case \( d = 1 \), which is exactly Proposition 2, and the transformation in Theorem 1 is a composition of transformations given by Proposition 2. Therefore we shall have to estimate the Gevrey norm of a composition of functions, and unlike classical \( C^0 \) norms used in the analytic case, these estimates are not completely trivial.

**Lemma 6.** There exists a constant \( C \) depending on \( n, \alpha \) and \( L \) such that for any \( 0 < \delta < R \), if \( F \in G^{\alpha,L}(\mathcal{D}_R, \mathbb{R}^{2n}) \) and \( \Phi \in G^{\alpha,L'}(\mathcal{D}_{R'}, \mathcal{D}_R) \), with \( L' = CL \) and \( R' = R - \delta \), and if \( |\Phi - \text{Id}|_{\alpha,L'} < 1 \), then \( F \circ \Phi \in G^{\alpha,L'}(\mathcal{D}_{R'}, \mathbb{R}^{2n}) \) and we have the estimate

\[ |F \circ \Phi|_{\alpha,L'} \leq |F|_{\alpha,L}. \]

This lemma is contained in the statement of Corollary A.1, Appendix A.2, in [16], in the case where \( F \) is a real-valued function, but it can be extended immediately to vector-valued functions, by applying it component by component. The implicit constant also depends on \( \delta \), but the statement will be used for a value of \( \delta \) depending only on \( d \) and \( R \).

Here is a direct consequence, which we will also use.

**Corollary 7.** There exists a constant \( C \) depending on \( n, \alpha \) and \( L \) such that for any \( 0 < \delta < R \), if \( \Psi \in G^{\alpha,L}(\mathcal{D}_R, \mathbb{R}^{2n}) \) and \( \Phi \in G^{\alpha,L'}(\mathcal{D}_{R'}, \mathcal{D}_R) \), with \( L' = CL \) and \( R' = R - \delta \), and if \( |\Phi - \text{Id}|_{\alpha,L'} < 1 \), then \( \Psi \circ \Phi \in G^{\alpha,L'}(\mathcal{D}_{R'}, \mathbb{R}^{2n}) \) and we have the estimate

\[ |\Psi \circ \Phi - \text{Id}|_{\alpha,L'} \leq |\Psi - \text{Id}|_{\alpha,L} + |\Phi - \text{Id}|_{\alpha,L'}. \]

Let us also state a straightforward lemma which says that the Gevrey norm of a function controls the Gevrey norm of its derivatives, provided we restrict the parameter \( L \) (in the statement below, we simply choose \( L/2 \)).

**Lemma 7.** Let \( f \in G^{\alpha,L}(\mathcal{D}_R) \). For any \( s \in \mathbb{N} \), there exists a constant \( c \) depending on \( s, \alpha \) and \( L \) such that

\[ \sup_{l \in \mathbb{N}^{2n}, |l|=s} |\partial^l f|_{\alpha,L/2} \leq c |f|_{\alpha,L}. \]

REFERENCES

1. Bambusi, D., Nekhoroshev Theorem for Small Amplitude Solutions in Nonlinear Schrödinger Equations, *Math. Z.*, 1999, vol. 230, no. 2, pp. 345–387.
2. Bounemoura, A. and Fischler, S., A Diophantine Duality Applied to the KAM and Nekhoroshev Theorems, *Math. Z.*, 2013 (to appear).
3. Baldomá, I., Fontich, E., Guardia, M., and Seara, Tere M., Exponentially Small Splitting of Separatrices beyond Melnikov Analysis: Rigorous Results, *J. Differential Equations*, 2012, vol. 253, no. 12, pp. 3304–3439.
4. Bounemoura, A. and Niederman, L., Generic Nekhoroshev Theory without Small Divisors, *Ann. Inst. Fourier (Grenoble)*, 2012, vol. 62, no. 1, pp. 277–324.
5. Bounemoura, A., Nekhoroshev Estimates for Finitely Differentiable Quasi-Convex Hamiltonians, *J. Differential Equations*, 2010, vol. 249, no. 11, pp. 2905–2920.
6. Bounemoura, A., Effective Stability for Gevrey and Finitely Differentiable Prevalent Hamiltonians, *Comm. Math. Phys.*, 2011, vol. 307, no. 1, pp. 157–183.
7. Bounemoura, A., Optimal Stability and Instability for Near-Linear Hamiltonians, *Ann. Henri Poincaré*, 2012, vol. 13, no. 4, pp. 857–868.
8. Bounemoura, A., Normal Forms, Stability and Splitting of Invariant Manifolds: II. Finitely Differentiable Hamiltonians, *Regul. Chaotic Dyn.*, 2013, vol. 18, no. 3, pp. 261–276.
9. Bolotin, S. V. and Treschev, D. V., Remarks on the Definition of Hyperbolic Tori of Hamiltonian Systems, *Regul. Chaotic Dyn.*, 2000, vol. 5, no. 4, pp. 401–412.
10. Delshams, A. and Gutiérrez, P., Effective Stability and KAM Theory, *J. Differential Equations*, 1996, vol. 128, no. 2, pp. 415–490.
11. Fassò, F., Lie Series Method for Vector Fields and Hamiltonian Perturbation Theory, *Z. Angew. Math. Phys.*, 1990, vol. 41, no. 6, pp. 843–864.
12. Lochak, P., Marco, J.-P., and Sauzin, D., On the Splitting of Invariant Manifolds in Multidimensional Near-Integrable Hamiltonian Systems, *Mem. Amer. Math. Soc.*, 2003, vol. 163, no. 775, 145 pp.
13. Loshak, P., Canonical Perturbation Theory: An Approach Based on Joint Approximations, *Uspekhi Mat. Nauk*, 1992, vol. 47, no. 6(288), pp. 59–140 [Russian Math. Surveys, 1992, vol. 47, no. 6, pp. 57–133].
14. Marco, J.-P., Sur les bornes inférieures de l’écart des variétés invariantes, *C. R. Math. Acad. Sci. Paris*, 2005, vol. 340, no. 11, pp. 839–842.
15. Mitev, T. and Popov, G., Gevrey Normal Form and Effective Stability of Lagrangian Tori, *Discrete Contin. Dyn. Syst. Ser. S*, 2010, vol. 3, no. 4, pp. 643–666.
16. Marco, J.-P. and Sauzin, D., Stability and Instability for Gevrey Quasi-Convex Near-Integrable Hamiltonian Systems, *Publ. Math. Inst. Hautes Études Sci.*, 2002, no. 96, pp. 199–275.
17. Nekhoroshev, N. N., An Exponential Estimate of the Time of Stability of Nearly-Integrable Hamiltonian Systems, *Uspekhi Mat. Nauk*, 1977, vol. 32, no. 6(198), pp. 5–66 [Russian Math. Surveys, 1977, vol. 32, no. 6, pp. 1–65].
18. Pöschel, J., Nekhoroshev Estimates for Quasi-Convex Hamiltonian Systems, *Math. Z.*, 1993, vol. 213, no. 2, pp. 187–216.
19. Pöschel, J., On Nekhoroshev Estimates for a Nonlinear Schrödinger Equation and a Theorem by Bambusi, *Nonlinearity*, 1999, vol. 12, no. 6, pp. 1587–1600.
20. Ramis, J.-P. and Schäfke, R., Gevrey Separation of Fast and Slow Variables, *Nonlinearity*, 1996, vol. 9, no. 2, pp. 353–384.
21. Simó, C., Averaging under Fast Quasiperiodic Forcing, in *Hamiltonian mechanics (Toruń, 1993)*, NATO Adv. Sci. Inst. Ser. B Phys., vol. 331, New York: Plenum, 1994, pp. 13–34.