Umbilical torus bifurcations in Hamiltonian systems

Henk W. Broer\textsuperscript{a}, Heinz Hanßmann\textsuperscript{b,1}, Jiangong You\textsuperscript{c}

\textsuperscript{a}Instituut voor Wiskunde en Informatica, Rijksuniversiteit Groningen, Blauwborgje 3, 9747 AC Groningen, The Netherlands
\textsuperscript{b}Institut für Reine und Angewandte Mathematik der RWTH Aachen, 52056 Aachen, Germany
\textsuperscript{c}Department of Mathematics, Nanjing University, Nanjing 210093, PR China

Received 9 December 2004
Available online 19 September 2005

Abstract

We consider perturbations of integrable Hamiltonian systems in the neighbourhood of normally umbilic invariant tori. These lower dimensional tori do not satisfy the usual non-degeneracy conditions that would yield persistence by an adaption of KAM theory, and there are indeed regions in parameter space with no surviving torus. We assume appropriate transversality conditions to hold so that the tori in the unperturbed system bifurcate according to a (generalised) umbilical catastrophe. Combining techniques of KAM theory and singularity theory we show that such bifurcation scenarios of invariant tori survive the perturbation on large Cantor sets. Applications to gyrostat dynamics are pointed out.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Hamiltonian dynamical system; KAM theory; Lower dimensional torus; Multiparameter bifurcations; Singularity theory; Umbilical catastrophe; Gyrostat

1. Introduction

The classical Kolmogorov–Arnol’d–Moser (KAM) theory deals with the persistence of Lagrangian invariant tori in nearly integrable Hamiltonian systems, see e.g. [35].

\textsuperscript{1}Present address: Mathematisch Instituut, Universiteit Utrecht, Postbus 80010, 3508 TA Utrecht, The Netherlands.
Also persistence of normally hyperbolic and normally elliptic tori has been studied, cf. e.g. [36,14,39]. In all cases the persistent tori constitute subsets of the phase space that have a Cantor like structure and a relatively large Hausdorff measure of twice the torus dimension. For an overview of these and related results see [13,39].

Given a non-degenerate integrable Hamiltonian system, the maximal tori are the regular fibres of the ramified torus bundle defined by the dynamics of this system. The singular fibres of this bundle, i.e. the lower dimensional tori together with their stable and unstable manifolds, determine how the maximal tori are distributed in phase space. Invariant $n$-tori form $n$-parameter families, parametrised by the actions conjugate to the toral angles. Generically one therefore expects to encounter bifurcations of co-dimension one at $(n-1)$-parameter subfamilies of lower dimensional tori, bifurcations of co-dimension two at $(n-2)$-parameter subfamilies and so on. To show persistence of such degenerate tori, embedded in the full bifurcation scenario, becomes more elaborate as the co-dimension increases. In the extreme case of co-dimension $n$ the bifurcating tori are isolated and may disappear in resonance gaps.

In this paper we consider lower dimensional tori with a vanishing Floquet exponent. The ensuing bifurcations with co-dimension one and two have a corresponding normal linear part $(0_{10})$, see [23,11]. As shown in the latter reference such normally parabolic tori may undergo quasi-periodic bifurcations of any co-dimension. Our aim is to similarly treat a number of bifurcations of invariant tori with vanishing normal linear part $(00)$. Under appropriate transversality conditions on the nonlinear terms, see (2), the co-dimension of such bifurcations may be as low as three, cf. [8,9]. Hence, Theorem 1 implies that in Hamiltonian systems with five (or more) degrees of freedom the lower dimensional tori can persistently undergo these bifurcations. Let us briefly sketch the setting of the present problem.

The phase space is $T^n \times \mathbb{R}^n \times \mathbb{R}^2$, with co-ordinates $(x, y, (p, q))$ and symplectic form

$$
\sigma = \sum_{i=1}^{n} dx_i \wedge dy_i + dq \wedge dp.
$$

We are looking at perturbations of a Hamiltonian system for which the torus $T^n \times \{0\} \times \{0\}$ is invariant and the normal linear part vanishes. This means that the Hamiltonian function has no linear or quadratic terms in $p$ and $q$. Mimicking the theory of bifurcations for equilibria and periodic solutions, cf. [30,31] or [8,9], we add the following assumption on the higher order terms. For some integer $d \geq 3$, the expansion of the unperturbed Hamiltonian in the $(p, q)$-direction has the principal part $a p^2 q + b q^d$, with $a, b \neq 0$. In the terminology of [1,3] this is the singularity $D^\pm_{k}$, with $k = d + 1$ and $\pm = \text{sgn}(ab)$. We shall call such invariant tori normally umbilic. To capture all possible bifurcations from the torus with this degenerate normal behaviour we include parameters $\lambda_1, \ldots, \lambda_d$ for a universal unfolding, cf. [7,37], according to the (generalised) umbilic catastrophes. See also Section 2 for more details. Moreover we include parameters for the frequencies $\omega_1, \ldots, \omega_n$. Indeed, disregarding some co-ordinate changes and reparametrisations, we shall assume that the unperturbed family has the “integrable”
form
\[ N(x, y, p, q, \lambda, \omega) = (\omega \mid y) + \frac{a(\omega)}{2} p^2 q + \frac{b(\omega)}{d!} q^d + \sum_{j=1}^{d-1} \lambda_j q^j + \lambda_d p, \quad (2) \]

where \((\cdot \mid \cdot)\) denotes the standard inner product on \(\mathbb{R}^n\). This family first of all has a continuum of normally umbilic invariant tori
\[ T^n \times \{0\} \times \{0\} \times \{0\} \times \mathcal{O} \subseteq T^n \times \mathbb{R}^n \times \mathbb{R}^2 \times \Lambda \times \mathcal{O}, \]
i.e. for every frequency vector \(\omega \in \mathcal{O}\) there is one such \(n\)-torus, given by the equations \(y = 0, (p, q) = 0, \lambda = 0\). Next, for \(\lambda \neq 0\) we find continuous branches of invariant tori of various types, normally hyperbolic, elliptic, parabolic and umbilic corresponding to the hierarchy of singularity theory, cf. [7,37,1,3]. Moreover there are Lagrangian invariant \((n + 1)\)-tori, foliating open pieces of the phase space. The general question of this paper is what remains of this global picture when we perturb to \(H = N + P\) where \(P\) is an arbitrary (not necessarily integrable) perturbation, small in an appropriate sense. Throughout, for simplicity, we assume real analyticity of \(H\) in all variables and parameters, observing however that immediate adaptations exist for \(H \in C^j\), for \(j\) sufficiently large, including \(j = \infty\). See [35] or [14, Appendix].

This perturbation problem is not expected to have an affirmative answer for all parameters \(\omega\), but again only on a set of Cantor like structure. Indeed, on the vector \(\omega \in \mathcal{O}\) we impose Diophantine conditions, saying that
\[ |(\omega \mid k)| \gtrsim \frac{\gamma}{|k|^\tau}, \quad \forall k \in \mathbb{Z}^n \setminus \{0\}, \quad (3) \]
where \(\gamma > 0\) is to be chosen appropriately small later on and where \(\tau > nL - 1\) is fixed with \(L\) given in Theorem 8—in the present bifurcational context we can no longer restrict to \(L = 1\), but \(L = 2\) will usually do. The first result of this paper roughly says the following. For values of \(\omega\) in a Cantor set given by the above restriction, the family \(H = N + P\), with \(P\) sufficiently small in the compact-open topology on holomorphic extensions, again has such normally umbilic invariant \(n\)-tori near \(y = 0, (p, q) = 0, \lambda = 0\). These perturbed tori, moreover, form a Whitney-\(C^\infty\)-family, implying that their union has a large Hausdorff measure. In the next section we shall give a precise formulation of the corresponding theorem. We remark that our conditions are global with respect to \(\mathcal{O}\), i.e. not restricted to a small neighbourhood of some fixed frequency vector \(\omega_0\) satisfying (3).

The perturbed tori just mentioned are the most degenerate ones corresponding to the central singularity at \(\lambda = 0\) and the remaining part of our perturbation problem asks what happens to the invariant tori of \(N\) that occur in the unfolding for \(\lambda \neq 0\). In a second result we approach this problem recursively with respect to \(d\). It turns out that the hierarchy of singularity theory carries over to the KAM-setting—similar to the parabolic case and familiar from catastrophe theory.
Summarising we give a rough all-over description of the invariant tori found by this approach. The key lies already in the behaviour of the unperturbed integrable normal form. The smooth parametrisations of the various families of invariant tori found there will then be subject to Diophantine restrictions, meaning that the final result deals with a Cantor stratification in the product of phase space and parameter space.

The behaviour of the normal form $N$ is best explained noticing that the invariant tori give the product of phase space and parameter space the structure of a ramified torus bundle. An open and dense part is filled by the union of Lagrangian invariant $(n+1)$-tori, these define the regular fibres of this bundle. The complement consists of invariant $n$-tori, defining singular fibres of various degrees according to occurring bifurcations. In the space of external parameters $\lambda$ and frequencies $\omega$ this yields a stratification—each stratum of co-dimension $k$ parametrising invariant $n$-tori that undergo a bifurcation of that same co-dimension. We will focus on the Cantor subsets of this stratification defined by the Diophantine condition (3). However, our results can also be used as starting point for a better understanding of the various resonances (in the gaps of the Cantor sets), as has been done in [27–29] for normally parabolic tori undergoing a Hamiltonian pitchfork bifurcation.

To fix thoughts let us concentrate on the case $d = 3$ with $a \cdot b > 0$, see Fig. 1 for the bifurcation set of the lower dimensional tori defined by $N$. The point $\lambda_1 = \lambda_2 = \lambda_3 = 0$ where the upper and the three lower surfaces meet corresponds to the invariant torus with normal linear part \((000)\). The left and right lower surfaces stand for parabolic invariant tori that undergo (quasi-periodic) centre-saddle bifurcations cf. [30,31,8,9,23]; this is related to a subordinate fold catastrophe. Along the cusp line the bifurcation of the normally parabolic tori becomes degenerate and is related to a (dual) cusp catastrophe, see [8,9,11].

For parameter values “below” the central point there are two normally elliptic and two normally hyperbolic invariant tori, while for parameter values “between” the upper and the three lower surfaces only one of each are left. On the plane emanating from the cusp line the two hyperbolic tori have the same energy and become connected by heteroclinic orbits. This connection bifurcation is an example of a global bifurcation subordinate to the local bifurcations defined by (2). The upper surface parametrises again parabolic tori where the remaining two families of elliptic and hyperbolic tori meet and vanish in a quasi-periodic centre-saddle bifurcation.

For the dynamics defined by $N$ there is one bifurcation diagram for each frequency vector $\omega$. Using a Kolmogorov-type non-degeneracy condition, cf. (6) below, we may switch to the phase space where the actions $y$ conjugate to the toral angles $x$ play the rôle of the frequencies. In the product of phase space and parameter space the union of all lower dimensional tori is a stratified set of co-dimension 2, the complement of which is filled by $(n + d + 1)$-parameter families of invariant $(n + 1)$-tori. In this paper we show that, under a small generic Hamiltonian perturbation, this stratification becomes a Cantor stratification, with all parametrisations getting restricted to Cantor sets defined by Diophantine conditions (while the actual invariant tori remain analytic tori).

Occurrence of the type of bifurcation at hand most often follows from a normalising or averaging procedure. Indeed, in an integrable approximation we may detect
the unperturbed dynamics by finding the most degenerate singularity and checking the parameter dependence. In other examples, like in the gyrostat problem considered in Section 3, the physical model already has enough symmetries to render integrability of the system. Here KAM theory can provide the justification of such symmetry assumptions by showing that small imperfections (inherent to all real-life mechanical systems) do not completely invalidate the analysis of the idealised model, but that in fact the perturbed dynamics adhere rather closely to the unperturbed dynamics.

This paper is organised as follows. In the next section we state our main result, the proof of which occupies the final Section 4. In Section 3 we treat the gyrostat as an application.

2. Formulation of the results

Generally speaking, when proving a persistence theorem the difficult part is to keep track of the most degenerate “object” in the perturbed system. Our first step is therefore to look for the “bifurcating” normally umbilic invariant $n$-tori of $X_H$. 

Fig. 1. Bifurcation set of $N$ for $d = 3$ with $a \cdot b > 0$. There are two reflectional symmetries $\lambda_2 \mapsto -\lambda_2$ and $\lambda_3 \mapsto -\lambda_3$. 
Let $T^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$ be the standard $n$-torus and $\mathcal{Y} \subseteq \mathbb{R}^n$, $\mathcal{S} \subseteq \mathbb{R}^2$, $\Lambda \subseteq \mathbb{R}^d$ be neighbourhoods of the respective origins. By $O$, we denote the set of those frequency vectors $\omega \in O$ that satisfy the Diophantine condition (3). We also need $O_\gamma := \{\omega \in O : d(\omega, \partial O) \geq \gamma\}$. Furthermore $|.|$ stands for the supremum norm on the set $A$.

**Theorem 1.** Let the functions $a, b : O \to \mathbb{R}$ in the normal form (2) satisfy $|a|_O$, $|b|_O$, $\left|\frac{1}{a}\right|_O$, $\left|\frac{1}{b}\right|_O$, $|Da|_O$, $|Db|_O < C$ for some constant $C > 0$. Then there exists a small positive constant $\varepsilon$, independent of $O$, with the following property. For any analytic perturbation $H = N + P$ of (2) with

$$|P|_{T^n \times \mathbb{R}^n \times \mathbb{R}^2 \times \mathbb{R}^d \times O} < \varepsilon$$

there exists a $C^\infty$-diffeomorphism $\Phi$ on $T^n \times \mathbb{R}^n \times \mathbb{R}^2 \times \mathbb{R}^d \times O$ such that

1. $\Phi$ is real analytic for fixed $\omega$.
2. $\Phi$ is symplectic for fixed $(\lambda, \omega)$.
3. $\Phi$ is $C^\infty$-close to the identity.
4. On $T^n \times \mathbb{R}^n \times \mathbb{R}^2 \times \mathbb{R}^d \times O_\gamma \cap \Phi^{-1}(T^n \times \mathbb{Y} \times \mathcal{S} \times \Lambda \times O)$ one can split $H \circ \Phi = N_\infty + P_\infty$ into an integrable part $N_\infty$ and higher order terms $P_\infty$. Here $N_\infty$ has the same form as $N$, see (2). The $x$-dependence is pushed into the higher order terms, i.e. $\partial^{[l]} p^i \partial q^j \partial \lambda^h (x, 0, 0, 0, \omega) = 0$ for all $(x, \omega) \in T^n \times O_\gamma$ and all $l, i, j, h$ satisfying $2d|l| + (d - 1)i + 2j + (2d - 2)h_1 + \cdots + 2h_{d-1} + (d + 1)h_d \leq 2d$.

**Remark 2.** A closer inspection of the proof, cf. [41], reveals that $\Phi$ is not only $C^\infty$, but can in fact be chosen to be Gevrey regular.

We prove Theorem 1 in Section 4, using a KAM iteration scheme. Here let us first elaborate its implications.

An immediate consequence is the persistence of normally umbilic $n$-tori at the “origin” $y = p = q = \lambda = 0$. These are parametrized by the Diophantine frequency vectors $\omega \in O_\gamma$, i.e. they form a Cantor family. This Cantor family at $\lambda = 0$ corresponds to the most degenerate invariant tori. We claim that the whole bifurcation scenario of the integrable family $N$ persists the perturbation by $P$ on Cantor sets. For a precise formulation we need the concept of a Cantor stratification, cf. [11].

The polynomial normal forms from singularity and catastrophe theory all have semi-algebraic catastrophe and bifurcation sets. The further complications in the definition of such stratifications largely arise from the fact that singularity theory allows analytic or smooth transformations and reparametrisations, that need not be algebraic. The ensuing problem is to characterise the analytic or smooth stratifications thus obtained, cf. e.g. [43, 21]. Without stressing this subject too much, we just extend the above class of smooth transformations a bit further. Indeed, inside the semi-algebraic stratification we single out a Cantor set and consider Whitney-$C^\infty$-smooth transformations with respect to this. The corresponding Whitney extensions also are smooth on the whole semi-
algebraic set. The stratification thus obtained will colloquially be referred to as Cantor stratification.

We use Theorem 1 to obtain a Cantor stratification in an inductive manner. Near the above Cantor family of most degenerate umbilic tori we expect bifurcating tori of lower co-dimensions to occur—in exactly the same way as the normal form has a bifurcation set that is stratified into the various subordinate bifurcations. We always have subordinate bifurcations of normally parabolic tori, and for $d \geq 4$ there are also normally umbilic tori of lower co-dimension. For the latter, we invoke Theorem 1 using a normal form like (2) with $d$ replaced by $d-1$, then by $d-2$ and so on until we reach $d = 3$. At the same time we apply Theorem 2.1 of [11] to deal with the occurring normally parabolic tori. Here we use the hierarchical adjacency

$$
\begin{align*}
D_4^\pm & \leftarrow D_5 \leftarrow D_6^\pm \leftarrow \cdots \\
A_1^\pm & \leftarrow A_2 \leftarrow A_3^\pm \leftarrow A_4 \leftarrow A_5^\pm \leftarrow A_6 \leftarrow \cdots
\end{align*}
$$

of singularities of type $D_k$ and $A_k$, cf. [1,3]. In this way we obtain the following result.

**Theorem 3.** Under the conditions of Theorem 1, there is a Cantor stratification of a Cantor subset of $\Lambda \times \mathcal{O}$ of large measure into $(n+d-k)$-dimensional Cantor sets $C_k$, $k = 0, \ldots, d$, such that $C_0$ parametrises Cantor families of elliptic and hyperbolic tori and $C_k$, $k = 1, \ldots, d$ parametrise Cantor families of invariant tori of co-dimension $k$.

For the proof one uses the co-ordinates provided by the transformation $\Phi$ of Theorem 1. The quasi-periodic flow induced by the term $(\omega \mid y)$ of (2) is superposed by the one-degree-of-freedom system with Hamiltonian

$$
H_d(p, q, \lambda) = \frac{a(\omega)}{2} p^2 q + \frac{b(\omega)}{d!} q^d + \sum_{j=1}^{d-1} \frac{\lambda_j}{j!} q^j + \lambda_d p.
$$

Applying singularity theory to $H_d$ one obtains the hierarchy (4) of adjacencies. Thus, the unfolding (5) contains all singularities of type $D_{d+1}$ and $A_{k+1}$ with $k \leq d-1$ in a subordinate way (in our context $A_{k+1}^+$ stands for normally elliptic and $A_{k+1}^-$ for hyperbolic tori).

This allows us to lead the situation around some higher stratum $S_k$ back to [11] or Theorem 1, depending on whether the corresponding tori with normal linear behaviour $(\omega \mid 0)$ satisfy $\alpha \neq 0$ or $\alpha = 0$. The explicit computations (which we do not repeat here, but see [11] for the adjacencies $A_k \leftarrow A_{k+1}$) show that after a translation that puts tori of lower co-dimension at the origin $(p, q) = (0, 0)$ one recovers (5) with $d$ replaced by $k$, while additional (higher order) terms may be treated as a perturbation. Where an adjacency $A_{k+1} \leftarrow D_{d+1}$ is concerned one does not recover (5) as lowest order terms, but the normal form

$$
(\omega \mid y) + \frac{\alpha(\omega)}{2} p^2 + \frac{\beta(\omega)}{(k+2)!} q^{k+2} + \sum_{j=1}^{k} \frac{\Lambda_j}{j!} q^j
$$
Remark 4. The normal form (2) has many homoclinic orbits to hyperbolic \( n \)-tori, where stable and unstable manifolds coincide. In the present Hamiltonian context homoclinic orbits are a typical phenomenon. However, one expects the stable and unstable manifolds to split, cf. [2]. For a generic perturbation \( P \) this leads to transversal homoclinic orbits. Under variation of parameters, also homo- and heteroclinic bifurcations are involved. The angle between the stable manifold and the unstable manifold, at a transversal homoclinic orbit, in this analytic setting is expected to be exponentially small in \( \varepsilon \), and also exponentially small in \( \lambda \) for the ‘newlyborn’ homoclinic loops generated by the unfolding (2). Subordinate to a “primary” homoclinic orbit, variation of the parameters \( \lambda \) may lead to homoclinic bifurcations, involving tangencies between the stable and unstable manifolds, cf. [25,33]. Similar observations apply mutatis mutandi to homoclinic orbits of parabolic and umbilic \( n \)-tori.

Remark 5. Whenever two unstable \( n \)-tori have the same energy they may be connected by heteroclinic orbits. Let us again concentrate on the case of hyperbolic tori, though almost no modifications are needed if one or both tori are parabolic or umbilic. For the integrable normal form there is a set of co-dimension one in parameter space for which connection bifurcations occur. Under variation of a further, transversal, parameter, the energy difference of the two hyperbolic tori changes from a positive to a negative value.

The circumstances of the formation of heteroclinic orbits change drastically under perturbation. In the generic case the stable and unstable manifolds that coincide for the unperturbed system have transversal intersections, which, however, are expected to be exponentially small. As a result the region in parameter space where heteroclinic orbits exist becomes an (exponentially small) “open horn”, cf. [17,16], at the boundary of which one has primary heteroclinic tangencies.

Remark 6. In applications the Hamiltonian is often invariant under some compact symmetry group. This strongly influences the bifurcations occurring in that the co-dimension within the corresponding “symmetric universe” is typically much lower. Correspondingly, one can use equivariant singularity theory, see [34], to derive adapted unfoldings. As the proof of Theorem 1 is of Lie algebra type and hence structure-preserving, cf. [32,14], the result carries over.

Remark 7. An important case occurs when the Hamiltonian is invariant under an involution, e.g. \( R : (x, y, p, q) \mapsto (x, -y, -p, q) \). Then \( R \) maps phase curves to phase curves, reversing the time, and the system is called reversible, cf. e.g. [40,10[a,b]]. For instance, the normal form (2) is reversible with respect to \( (x, y, p, q) \mapsto (-x, y, -p, q) \) if one drops the unfolding term \( \lambda_d p \).

Thus, in this reversible setting the co-dimension of normally umbilic tori with principal nonlinear terms \( \frac{a}{2} p^2 q + \frac{b}{d!} q^d \) drops from \( d \) to \( d - 1 \). In particular, with \( d = 3 \) there are already 3-parameter families of lower dimensional tori in four degrees of freedom persistently displaying quasi-periodic reversible umbilic bifurcations, cf. [15,22].
The normal form (2) depends on the parameters $\omega_i, i = 1, \ldots, n$ and $\lambda_j, j = 1, \ldots, d$. This seemingly special situation is in fact very general. Given a (single) Hamiltonian system with (unperturbed) Hamiltonian function $H_0$, the Kolmogorov-type non-degeneracy condition

$$\det (D^2 y H_0) \neq 0$$

(6)

enforces the frequency mapping $y \mapsto \omega(y) := D_y H_0$ to be a local diffeomorphism. In this way one can always replace the parameter vector $\omega \in \mathcal{O}$ by the variables $y \in \mathcal{Y}$. We are interested in furthermore also replacing the multiparameter $\lambda$ by $y$ in (2). To let $y$ compensate for all parameters $(\lambda, \omega)$ we use Diophantine approximation of dependent quantities, cf. [13, § 2.5] and references therein.

Let us explain how to recover the normal form (2) from a given integrable system with Hamiltonian function

$$H_0(y, p, q) = h(y) + \frac{a(y)}{2} p^2 q + \frac{b(y)}{d!} q^d$$

$$+ \sum_{j=1}^{d-1} \frac{c_j(y)}{j!} q^j + c_d(y)p + \text{higher order terms.}$$

(7)

As always the frequency vector $\omega$ is given by $\omega(y) = D_y h(y)$. While $a(y)$ and $b(y)$ are bounded from below for $y \in \mathcal{Y}$, the most degenerate bifurcation occurs at $y = 0$ as $c(0) \in \mathbb{R}^d$ vanishes.

The number of parameters $\lambda_j$ depends on the degeneracy $d$ as the universal unfolding of the singularity $D_{d+1}$ requires $d$ parameters, see [7,37]. In order that the corresponding bifurcation diagram be faithfully represented by means of the $y_i$, we require the map

$$c : \mathbb{R}^n \longrightarrow \mathbb{R}^d \\
y \mapsto c(y)$$

(8)

in (7) to be a submersion. This implies $n \geq d$, which is in agreement with the following genericity consideration.

In the present setting of $n + 1$ degrees of freedom, a non-degenerate integrable Hamiltonian system will have $n$-parameter families of invariant $n$-tori. Within these, normally elliptic and normally hyperbolic tori are parametrised over open subsets, while normally umbilic tori with dominant terms $p^2 q$ and $q^k$ are expected to form subfamilies of co-dimension $k \in \mathbb{N}$. In this way, bifurcating tori of degeneracy $d > n$ are not encountered and those of degeneracy $d = n$ are isolated.

The non-degeneracy condition (6) expresses that the partial derivatives $\frac{\partial |\ell| \omega_i}{\partial y^j}$ span $\mathbb{R}^n$, where $\ell \in \mathbb{N}^n$ with $|\ell| = 1$ ($|\ell| \leq 1$ in case of iso-energetic non-degeneracy, cf. [12]). This allows to control the frequency (the frequency ratio) of the perturbed tori. In the present case the proper requirement is that the image of

$$(c, \omega) : \mathbb{R}^n \longrightarrow \mathbb{R}^d \times \mathbb{R}^n$$

is “sufficiently curved” and does not lie in any linear hyperplane in $\mathbb{R}^{n+d}$ passing through the origin, see [38,4,39].
In this way we use (8) to pull back the bifurcation diagram to the space of actions. The remaining first derivatives together with the higher derivatives of \((c, \omega)\) then ensure that most frequencies perturbed from the \(\omega(y)\) are Diophantine and, hence, yield invariant tori in the perturbed system. Let us explicitly formulate our findings.

**Theorem 8.** Let \(T^n\) be the standard \(n\)-torus, \(Y\) a neighbourhood of the origin in \(\mathbb{R}^n\) and \(S\) a neighbourhood of the origin in \(\mathbb{R}^2\). Supply \(T^n \times Y \times S\) with the symplectic structure (1). Consider a perturbed Hamiltonian

\[
H = H_0(y, p, q) + \varepsilon H_1(x, y, p, q)
\]

with \(H_0\) given by (7) satisfying \(c(0) = 0\) and \(a(0), b(0) \neq 0\). Furthermore the mapping \(c : Y \rightarrow \mathbb{R}^d\) is a submersion and the \((n+L)\) vectors

\[
\left. \frac{\partial |\ell|}{\partial y^\ell} \right|_{y^\ell} (c, \omega), \quad |\ell| \leq L
\]

span \(\mathbb{R}^d \times \mathbb{R}^n\), where \(\omega(y) = Dh(y)\). Then the Cantor stratification of \(\Lambda \times O\) described in Theorem 3 induces a similar Cantor stratification of the phase space \(T^n \times Y \times S\), with all invariant tori \(C^\infty\)-close to some \(T^n \times \{\text{const.}\}\).

**Remark 9.** Since \(L\) enters the Diophantine condition (3) through \(\tau > nL - 1\) it is preferable to keep \(L\) as small as possible, i.e. to work with \(L = 2\).

### 3. Application to the gyrostat problem

A gyrostat is a rigid body to which one or several flywheels are attached, see [26,18,19]. The flywheels are assumed to be axially symmetric, in particular the mass distribution of the whole gyrostat system is not influenced by the individual rotation of the flywheels. We concentrate on a gyrostat with three flywheels rotating about the three principal axes of inertia. Furthermore we consider a “free” gyrostat, without any external forces or torques.

In the above description the gyrostat is a Hamiltonian system with nine degrees of freedom—the free rigid body has six degrees and each flywheel contributes an additional degree. The absence of external forces or torques yields conservation laws deriving from symmetries, the reduction of which allows to reduce to four degrees of freedom. In a first step we get rid of the three translational degrees and concentrate on the rotational aspect of the motion by fixing the body at one point. This point is the common intersection of the axes of the three flywheels (and not necessarily the centre of mass).

Without external forces or torques not only the linear momentum is constant (and put to zero in the above inertial frame), the angular momentum is a conserved quantity as well. Thus, two more degrees of freedom are reduced by fixing the three components
Table 1
Poisson bracket relations on the reduced phase space

<table>
<thead>
<tr>
<th>{ }, x_1</th>
<th>x_2</th>
<th>x_3</th>
<th>y_1</th>
<th>y_2</th>
<th>y_3</th>
<th>z_1</th>
<th>z_2</th>
<th>z_3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x_1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>x_2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>x_3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>y_1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>y_2</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>y_3</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>z_1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-z_3</td>
<td>z_2</td>
</tr>
<tr>
<td>z_2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>z_3</td>
<td>0</td>
</tr>
<tr>
<td>z_3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-z_1</td>
</tr>
</tbody>
</table>

of the angular momentum vector with respect to the inertial frame and dividing out the angle about the axis along the direction of the angular momentum. Co-ordinates on the resulting reduced phase space are provided by the moments \(y_1, y_2, y_3\) of the (relative) rotations of the three flywheels, together with the flywheel-angles \(x_1, x_2, x_3\) conjugate to these actions, and the three components \(z_1, z_2, z_3\) of the angular momentum vector with respect to the body frame of axes. The latter satisfy the relation

\[ z_1^2 + z_2^2 + z_3^2 = \mu^2, \]

where \(\mu\) denotes the (constant) length of the angular momentum, which we assume to be nonzero. The Poisson structure on the reduced phase space is given in Table 1.

Following [18,19,22] we express the Hamiltonian of the reduced system as

\[ H(x, y, z) = \sum_{i=1}^{3} \frac{y_i^2}{2J_i} + \frac{z_i^2 - 2y_iz_i}{2I_i}, \]

where \(I_1 < I_2 < I_3\) are the three principal moments of inertia of the whole gyrostat system and \(J_1, J_2, J_3\) denote the individual moments of inertia of the three flywheels. Because of the idealising assumption that the flywheels are (perfectly) axially symmetric the three angles \(x_1, x_2, x_3\) do not enter in (9) and one may immediately further reduce to one degree of freedom. Thereby the actions \(y_1, y_2, y_3\) become (internal, or distinguished) parameters of the system—in the same way as this happened for our model equations (2) and (7).

It is shown in [22] that four hyperbolic umbilic bifurcations take place. For

\[ (\hat{y}_1, \hat{y}_2, \hat{y}_3) = \left( \varepsilon \sqrt{\frac{I_3(I_2 - I_1)^3}{I_2(I_3 - I_1)^3}} \mu, 0, \delta \sqrt{\frac{I_1(I_3 - I_2)^3}{I_2(I_3 - I_1)^3}} \mu \right) \]

the one-degree-of-freedom Hamiltonian has a singularity of type \(D^+_4\) at the equilibria

\[ (\hat{z}_1, \hat{z}_2, \hat{z}_3) = \left( -\varepsilon \sqrt{\frac{I_3(I_2 - I_1)}{I_2(I_3 - I_1)}} \mu, 0, \delta \sqrt{\frac{I_1(I_3 - I_2)}{I_2(I_3 - I_1)}} \mu \right), \]
where $\varepsilon, \delta = \pm 1$ and $\mu = \sqrt{2^2_1 + 2^2_3}$ is the total angular momentum. Furthermore $y - \hat{y}$ provides a universal unfolding. Thus, in four degrees of freedom the Hamiltonian system defined by (9) has a 3-parameter family of invariant 3-tori that undergo four hyperbolic umbilic bifurcations. The corresponding motion is the conditionally periodic superposition

$$x(t) = x(0) + t \cdot \left( J^{-1} \hat{y} + I^{-1} \hat{z} \right)$$

(10)

of the three periodic rotations of the flywheels, where $J^{-1} = \text{diag}(J_1^{-1}, J_2^{-1}, J_3^{-1})$ and $I^{-1} = \text{diag}(I_1^{-1}, I_2^{-1}, I_3^{-1})$, while $y(t) \equiv \hat{y}$ and $z(t) \equiv \hat{z}$.

We now consider a small perturbation to model possible imperfections of the flywheels. As there are still no external forces or torques the reduction to four degrees of freedom remains valid. To Hamiltonian (9) we have to add a small $x$-dependent term.

Unfortunately we cannot apply Theorem 8 to this situation. The actions $y_1, y_2, y_3$ can easily play the rôle of unfolding parameters as $c : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is an invertible affine mapping, see [22], and hence a diffeomorphism. The frequency mapping

$$\omega : \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

$$y_i \mapsto \frac{y_i}{I_i} - \frac{\hat{z}_i}{I_i}$$

is a diffeomorphism as well. But the three actions $y_1, y_2, y_3$ cannot simultaneously replace six parameters. Indeed, the four normally umbilic invariant 3-tori are isolated and may fall into resonance gaps.

We can reconstruct the dynamics in five degrees of freedom by superposing the rotational motion about the angular momentum axis and allowing the value $y_4 := \mu$ of the length of the angular momentum to vary. In this way the lower dimensional tori have dimension 4 and form 4-parameter families, yielding four 1-parameter subfamilies of normally umbilic tori. The corresponding conditionally periodic motion consists of (10) superposed with a periodic rotation of the gyrostat about the (fixed) angular momentum axis, while $y_4(t) \equiv y_4(0) = \sqrt{2^2_1 + 2^2_3}$ is fixed in time as well. Again the four actions $y_1, y_2, y_3, y_4$ can easily play the rôle of unfolding parameters, the mapping $c : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ is a submersion. However, the image of $(c, \omega) : \mathbb{R}^4 \rightarrow \mathbb{R}^6$ is not sufficiently curved as the necessary second derivatives of $\omega$ vanish. In this sense the gyrostat problem is not a well-posed physical problem.

To remedy the situation we consider the moments of inertia as (external) parameters, thereby overcoming our “lack of parameter” problem. Indeed, not all of the above normally umbilic tori can disappear in resonance gaps as these gaps are separated by a Cantor set of large relative measure. Thus, for “most” gyrostats we can conclude that
even for not perfectly axially symmetric flywheels (quasi-periodic) hyperbolic umbilic bifurcations take place.

We remark that reconstructing the remaining four degrees of freedom does not help us as both the direction of the angular momentum vector in the inertial frame and the three translational degrees of freedom do not carry any dynamics. In particular, these quantities do not enter the expressions for $c$ or $\omega$ and therefore cannot be used to make $(c, \omega) : \mathbb{R}^9 \rightarrow \mathbb{R}^6$ a submersion. Indeed, the gyrostat is a superintegrable system. Therefore, small perturbing forces or torques may lead to much more complicated situations.

Let us restrict to a gyrostat with a fixed point (the common intersection of the axes of the three flywheels). Then the perturbation analysis takes place in six degrees of freedom since the perturbing external forces or torques cannot lead to translational motion. In this constellation the “free gyrostat with a fixed point” is a minimally superintegrable system and it is generic for the perturbation to remove the degeneracy, cf. [2]: there is an “intermediate” non-degenerate integrable system that is a better approximation of the non-integrable dynamics than is the superintegrable system.

The Lagrangian tori of the intermediate system have 5 fast and 1 slow frequency. The fast “free” motion of the gyrostat is a superposition of a periodic motion $z(t)$ of the angular momentum in the body frame of axes that keeps $\|z(t)\| \equiv y_4(0)$ fixed, a periodic rotation of the gyrostat about the angular momentum axis, the flywheels’ rotations and a slow periodic motion of the angular momentum in space. Superposition of this slow periodic motion with the lower dimensional fast tori leads to invariant 5-tori in six degrees of freedom that undergo hyperbolic umbilic bifurcations along 2-parameter subfamilies. Equilibria of the slow dynamics lead to invariant 4-tori in six degrees of freedom that have normal dynamics with two degrees of freedom. Where this leads to partial hyperbolicity we may immediately reduce to a centre manifold to obtain invariant 4-tori in five degrees of freedom that have normal dynamics with two degrees of freedom. Where this leads to partial hyperbolicity we may immediately reduce to a centre manifold to obtain invariant 4-tori in five degrees of freedom that have normal dynamics with two degrees of freedom. Where this leads to partial hyperbolicity we may immediately reduce to a centre manifold to obtain invariant 4-tori in five degrees of freedom that have normal dynamics with two degrees of freedom. Where this leads to partial hyperbolicity we may immediately reduce to a centre manifold to obtain invariant 4-tori in five degrees of freedom that have normal dynamics with two degrees of freedom. Where this leads to partial hyperbolicity we may immediately reduce to a centre manifold to obtain invariant 4-tori in five degrees of freedom that have normal dynamics with two degrees of freedom. Where this leads to partial hyperbolicity we may immediately reduce to a centre manifold to obtain invariant 4-tori in five degrees of freedom that have normal dynamics with two degrees of freedom. Where this leads to partial hyperbolicity we may immediately reduce to a centre manifold to obtain invariant 4-tori in five degrees of freedom that have normal dynamics with two degrees of freedom. Where this leads to partial hyperbolicity we may immediately reduce to a centre manifold to obtain invariant 4-tori in five degrees of freedom that have normal dynamics with two degrees of freedom. Where this leads to partial hyperbolicity we may immediately reduce to a centre manifold to obtain invariant 4-tori in five degrees of freedom that have normal dynamics with two degrees of freedom.
In fact we first fix the Diophantine constant $\gamma = 1$ when proving Theorem 1. This allows for a more transparent argumentation where the sizes $\beta_v$ of the shrinking neighbourhoods of $O'_1$ are effectively decoupled from the Diophantine constant $\gamma > 0$. A simple scaling argument, cf. [11] or [14, 6], allows to extend the result thus obtained to the $O'_0$ of Theorem 1.

To ensure that the limit is Whitney-$C^\infty$-smooth in $\omega$ we work on domains $D(r_v, s_v, \beta_v)$ that shrink geometrically in the $x$- and $\omega$-directions. Then an exponentially fast decreasing sequence $(\varepsilon_v)_{v \in \mathbb{N}}$ that controls at the $v$th step the (transformed) perturbation $P_v$ allows to use the Inverse Approximation Lemma of [45] for the desired Whitney-$C^\infty$-smoothness. The necessary control of $P_v$ is in turn obtained by letting shrink $D(r_v, s_v, \beta_v)$ exponentially in the $(y, p, q, \lambda)$-directions, described by $s_v = \varepsilon_v^{\frac{1}{d+\sigma}}$ with $\sigma \in \{0, 1\}$. The limit

$$\bigcap_v D(r_v, s_v, \beta_v)$$

consists in the $\omega$-direction exactly of the set $O'_1$ of Diophantine frequency vectors, while it shrinks to $\{0\}$ in the $(y, p, q, \lambda)$-directions. Analyticity in the latter variables then is obtained by interpreting the limit functions as the $(x, \omega)$-dependent coefficients of polynomials like $N_\infty$.

An additional complication is that a mere polynomial truncation of the $\Phi_v$ would cease to preserve the symplectic structure. For this reason we introduce generating functions $S_v$ of the $\Phi_v$, the polynomial truncations $\tilde{S}_v$ of which generate symplectomorphisms as well. The limit $\tilde{S}_\infty$ of these then generates the desired $\Phi_\infty$.

We define the $v$th transformation

$$\Phi_v = \Psi_0 \circ \Psi_1 \circ \cdots \circ \Psi_{v-1}$$

where

$$\Psi_{v-1} : D(r_v, s_v, \beta_v) \rightarrow D(r_{v-1}, s_{v-1}, \beta_{v-1}).$$

At each iteration step we want $\Psi_{v-1}$ to solve two problems. The $x$-dependence has to be confined to the new (and smaller) perturbation $P_v$, and the $x$-independent terms have to be transformed into normal form $N_v$. We can explicitly decouple the solution of these two problems and construct

$$\Psi_{v-1} = \phi_{v-1} \circ \varphi_{v-1}.$$

Thus, $\varphi_{v-1}$ is the solution of the linearised (or “1-bite”) small denominator problem and $\phi_{v-1}$ uses explicit transformations from singularity theory to put the (now $x$-independent) lower order terms again into normal form (2).
In this way the method of proof follows the standard KAM recipe, inspired by the iterative schemes of e.g. [32] and [36]. At the \( \lambda \)th step we have a perturbation form \( H_\lambda = N_\lambda + P_\lambda \) with \( N_\lambda \) in normal form (2) and \( P_\lambda \) sufficiently small. In the limit we obtain \( H_\infty = N_\infty + P_\infty \), where \( P_\infty \) has no longer any influence on the tori at \( \lambda = 0 \) and their normal behaviour, thus yielding the desired persistence result.

The “sufficiently small” term \( P_{\lambda+1} \) is obtained at the \( \lambda \)th iteration step as a “remainder term” and mainly consists of “higher order terms”. The key observation that helps deciding which terms are relegated to \( P_{\lambda+1} \) is that \( N_\lambda \) is a quasi-homogeneous polynomial with weight \((2d, d-1, 2; 2d-2, \ldots, 2, d+1)\) in the variables \((y, p, q, \lambda)\), cf. [1,3,11]. This weight in turn induces the weighted order

\[
\| (l, i, j, h) \| := 2d|l| + (d-1)i + 2j + (2d-2)h_1 + \cdots + 2h_d = (d+1)h_d
\]

on indices

\[
l = (l_1, \ldots, l_n) \in \mathbb{N}_0^n, \quad i, j \in \mathbb{N}_0, \quad h = (h_1, \ldots, h_d) \in \mathbb{N}_0^d
\]

of monomials \( y^l p^i q^j \lambda^h \) whence e.g. \( N_\lambda \) has (weighted) order \( 2d \). The weights \( 2d, d-1, 2d-2j \) and \( d+1 \) also enter the shrinking domains

\[
D(r_\lambda, s_\lambda, \beta_\lambda) = D(r_\lambda, s_\lambda) \times U_{\beta_\lambda}(O'_1),
\]

where

\[
D(r_\lambda, s_\lambda) = \left\{ (x, y, p, q, \lambda) \mid \|x\| \leq r_\lambda, \|y\| \leq s_\lambda^{2d}, \right. \\
\left. |p| \leq s_\lambda^{d-1}, |q| \leq s_\lambda^{2}, |\lambda| \leq s_\lambda^{2d-2j}, |\lambda| \leq s_\lambda^{d+1} \right\}
\]

and the second factor is a complex \( \beta_\lambda \)-neighbourhood

\[
U_{\beta_\lambda}(O'_1) = \left\{ w \in \mathbb{C}^n \mid \exists \omega \in O'_1 | w - \omega | < \beta_\lambda \right\}
\]

of the set \( O'_1 \subseteq O \subseteq \mathbb{R}^n \) of frequency vectors. The reason to consider our (analytic) Hamiltonians on complex domains is that this allows to control derivatives by the supremum norm using Cauchy’s inequality, i.e. the Cauchy integral formula. For the set \( D(r_0, s_0, \beta_0) \) to which we extend the initial Hamiltonian \( H \) we can use \( U_{\beta_0}(O) \) as second factor since the Diophantine conditions have not yet entered. Similar to the smallness condition on the initial perturbation \( P_0 \) we want to achieve

\[
\| P_\lambda \|_{D(r_\lambda, s_\lambda, \beta_\lambda)} \leq \varepsilon_\lambda
\]

on the \( \lambda \)th iteration step. Here \( \varepsilon_\lambda \) is related to \( s_\lambda \) through \( \varepsilon_\lambda = s_\lambda^{2d+\sigma} \), and both converge exponentially fast to 0. For the precise definition of the sequences \( r_\lambda, s_\lambda, \beta_\lambda, \varepsilon_\lambda \) see (14).
We follow [11] very closely and in particular split the presentation into Lemmata 10–20. The proof of Lemma 10 occupies most of the next subsection (and in particular uses Lemmata 12–18), but the proofs of Lemmata 12–20 are very similar to those of Lemmata 6.5–6.15 in [11] and therefore omitted. This allows us to concentrate on the changes that have to be incorporated, while still giving a complete proof of Theorem 1.

4.1. The iteration step

The aim of a single step of the KAM iteration is to find a co-ordinate transformation that turns the given Hamiltonian $H_\nu$ into a “new” Hamiltonian $H_{\nu+1}$ that differs “less” from the “new” normal form $N_{\nu+1}$. To this end we rewrite $H_\nu$ as

$$H_\nu = N_\nu + R_\nu + (P_\nu - R_\nu),$$

where $R_\nu$ is a conveniently chosen higher order truncation of $P_\nu$, see (18). We show below how the Newton-like accelerated convergence implies that $|P_\nu - R_\nu|$ is less than $|P_\nu|^2, \zeta > 1$, on the smaller domain $\mathcal{D}(r_{\nu+1}, s_{\nu+1}, \beta_{\nu+1})$.

Let $F_\nu$ be a function defined in a domain $\mathcal{D} \subseteq \mathcal{D}(r_\nu, s_\nu, \beta_\nu)$ and let $X_{F_\nu}$ be the vector field with Hamiltonian function $F_\nu$. Denote by $\varphi_{F_\nu}^\tau$ the flow of $X_{F_\nu}$ and $\varphi_{F_\nu} := \varphi_{F_\nu}^{\tau=1}$. We then have

$$H_\nu \circ \varphi_{F_\nu} = (N_\nu + R_\nu) \circ \varphi_{F_\nu} + (P_\nu - R_\nu) \circ \varphi_{F_\nu}$$

$$= N_\nu + R_\nu + \{N_\nu, F_\nu\} + \{R_\nu, F_\nu\}$$

$$+ \int_0^1 (1 - t)\{[N_\nu + R_\nu, F_\nu], F_\nu\} \circ \varphi_{F_\nu}^\tau \, dt + (P_\nu - R_\nu) \circ \varphi_{F_\nu}$$

$$= N_\nu + R_\nu + \{N_\nu, F_\nu\} + \hat{P}_\nu,$$  \hspace{1cm} (11)

where we use in the second equality the Taylor formula for the function $g(t) = (N_\nu + R_\nu) \circ \varphi_{F_\nu}^t$ with its derivatives $\dot{g}(0) = \{g(0), F_\nu\}$ and $\ddot{g}(t) = \{\{g(0), F_\nu\}, F_\nu\} \circ \varphi_{F_\nu}^t$.

The philosophy of the KAM method is to find a special $F_\nu$ defined in a shrunken domain which makes the new perturbation $\hat{P}_\nu$ in (11) much smaller and $N_\nu + R_\nu + \{N_\nu, F_\nu\}$ a new normal form $N_{\nu+1}$. In the present context such a normal form not only means a Hamiltonian function that is independent of the angles $x$, i.e. integrable, but that furthermore defines a versal unfolding of the bifurcating tori at $\lambda = 0$. In the case of normally elliptic or hyperbolic tori, we do not need to put the higher order terms of $q$ into the normal form; $F_\nu$ is thus obtained by solving a linear partial differential equation, the so-called homological equation

$$N_\nu + R_\nu + \{N_\nu, F_\nu\} = N_{\nu+1},$$ \hspace{1cm} (12)
where as usual,
\[
\{N, F\} = \frac{\partial N}{\partial x} \frac{\partial F}{\partial y} - \frac{\partial N}{\partial y} \frac{\partial F}{\partial x} + \frac{\partial N}{\partial q} \frac{\partial F}{\partial p} - \frac{\partial N}{\partial p} \frac{\partial F}{\partial q}.
\]

In the present bifurcating case, since \(N_v\) contains higher order terms in \(q\), Eq. (12) cannot be solved completely. Note that the purpose of solving (12) is to find a function \(F_v\) so that (11) becomes the sum of a new normal form and a smaller perturbation. To achieve this, we split the iteration step into two parts.

1. Instead of solving (12), we solve the “intermediate homological equation”
\[
N_v + R_v + \{N_v, F_v\} = \bar{N}_v
\]
up to some order and treat the higher order terms (which are smaller) as a part of the new perturbation. The “intermediate” \(\bar{N}_v\), already independent of \(x\), but not yet normalised in \(p\) and \(q\), is defined later in (22). The solution of (13) leads to small denominators, for which the Diophantine conditions (3) are needed. For the \(v\)th iteration step we only use finitely many of these conditions, up to some “ultraviolet” cut-off for the order \(K_v\) of the Fourier truncation (18) defined in (14).

2. Then we look for a symplectic change of variables which transforms \(\bar{N}_v\) into normal form (2). This passage from \(\bar{N}_v\) to \(N_{v+1}\) does not involve small denominators, but requires methods from singularity theory instead.

4.1.1. The iteration lemma
To formulate the iteration lemma we need several convergent sequences of numbers, and the interplay of geometrically fast and exponentially fast convergence later on yields the desired (Whitney)-smoothness. For any given positive numbers \(r_0, s_0\) we recursively define the sequences
\[
\rho_v = \frac{\rho_{v-1}}{4} = \frac{1}{4^v} \cdot \frac{3r_0}{32},
\]
\[
r_v = r_{v-1} - 4\rho_{v-1} = \frac{r_0}{2} \left(1 + \frac{1}{4^v}\right),
\]
\[
\beta_v = \rho_v^{2\sigma+2},
\]
\[
K_v = [\beta_v^{\frac{1}{\sigma+1}}] = [\rho_v^{-2}],
\]
\[
s_v = s_{v-1}^{\frac{2d+\sigma}{\sigma+1}} = s_{v-1}^{\gamma (1+\frac{2d+\sigma}{\sigma+1})},
\]
\[
\varepsilon_v = s_v^{2d+\sigma}
\]
with \(0 < \kappa < \sigma < 1\). The constants in the estimates below will be absorbed in \(r_0\) and \(s_0\), leading to inequalities of the form
\[
r_0 \leq c, \quad s_0 \leq c, \quad s_0^{\frac{\varepsilon}{r_0}} \leq cr_0
\]
with constants $c > 0$ and exponents $\zeta > 0$. The only exception is the (omitted) proof of Lemma 12 where an inequality

$$r_0 < \frac{1}{c - \zeta \ln(s_0)}$$

occurs, see [11]. Since $\zeta \ln(s_0) \to 0$ for all $\zeta > 0$ it is possible to find small $r_0, s_0$ satisfying all these inequalities.

With these sequences at hand we now can formulate the iteration lemma. We consider a Hamiltonian function

$$H_v = N_v + P_v$$

with

$$N_v = (\omega \mid y) + \frac{a_v}{2} p^2 q + b_v q^d + \sum_{j=1}^{d-1} \frac{\lambda_j}{j!} q^j + \lambda_d p$$

and defined in

$$D_v := D(r_v, s_v, \beta_v) = D(r_v, s_v) \times U_{\beta_v}(O_1').$$

We also use the abbreviation

$$U_v := U_{\beta_v}(O_1')$$

for the $\beta_v$-neighbourhood in the second factor.

**Lemma 10.** Suppose that $H_v = N_v + P_v$ satisfies (16) in $D_v$ and that $P_v$ can be estimated by

$$|P_v|_{D_v} \leq \varepsilon_v.$$  \hspace{1cm} (17)

Then, for sufficiently small $s_0$, there is a symplectic change of variables

$$\Psi_v : D_{v+1} \to D_v$$

such that $H_{v+1} = H_v \circ \Psi_v$, defined on $D_{v+1}$, has the form

$$H_{v+1} = N_{v+1} + P_{v+1},$$

satisfying

$$|P_{v+1}|_{D_{v+1}} \leq \varepsilon_{v+1},$$

$$|a_{v+1} - a_v|_{U_{\beta+1}} \leq s_v,$$

$$|b_{v+1} - b_v|_{U_{\beta+1}} \leq s_v.$$
Moreover,
\[
\left| \frac{\partial^{|l|+i+j+|h|} P_{v+1}}{\partial y^l \partial p^i \partial q^j \partial \lambda^h} \right| D_{v+1} \leq s_{v+1}^{2d+\alpha-m}.
\]

where \( m := \|(l, i, j, h)\| \leq 2d \).

**Remark 11.** Compared to the perturbation, the coefficient functions \( a_v \) and \( b_v \) are of order one, i.e. they satisfy bounds as formulated in Theorem 1. The estimates by \( s_v \) on the differences \( |a_{v+1} - a_v|, |b_{v+1} - b_v| \) imply that the same is true for \( a_{v+1} \) and \( b_{v+1} \) as well, and also for the (existing) limit functions \( a_\infty \) and \( b_\infty \).

### 4.1.2. The “intermediate” homological equation

To prove Lemma 10 we describe a single iteration step in detail. Therefore, we drop the index \( \alpha \) and use the so-called “+”-notation, replacing occurrences of the index \( \alpha \) by an index \( + \). As said earlier, we look for a symplectic co-ordinate transformation such that the transformed Hamiltonian function satisfies (15)–(17) with \( s_+, e_+ \) and so on. This also emphasizes that the constants in our estimates have to be independent of \( \varepsilon \). The generic letter “\( c \)” is used where we do not need to remember the value of such a constant, and we also use the shorthand \( A \leq c \cdot B \).

We expand the perturbation \( P \) into a Fourier–Taylor series
\[
P(x, y, p, q, \lambda) = \sum_{m=0}^{\infty} \sum_{\|(l, i, j, h)\|=m} \sum_{k \in \mathbb{Z}^n} P_{klijh} e^{i(k|x|)} y^l p^i q^j \lambda^h
\]
and define the truncation
\[
R = \sum_{|k| \leq K} P_k e^{i(k|x|)}
\]
(18)
of \( P \) with
\[
P_k(y, p, q, \lambda) = \sum_{m \leq 2d} P_{km} = \sum_{m \leq 2d} \left( \sum_{\|(l, i, j, h)\|=m} P_{klijh} y^l p^i q^j \lambda^h \right).
\]
(19)

Here \( |k| = K \), with \( K = \lfloor \beta^{-1} \rfloor \), is the maximal order \( |k| = |k_1| + \cdots + |k_n| \) of the resonances we have to cope with at this stage. We need bounds on both the truncation \( R \) of \( P \) we use to define the co-ordinate transformation (by solving (13)) and on the remaining term \( P - R \) which will be included in the new (and smaller!) perturbation.
Lemma 12. Under the conditions of Lemma 10 the inequality

\[ |R|_{\mathcal{D}(r-\rho, \frac{1}{2}s, \beta)} \leq \varepsilon \]  

(20)

holds. Moreover, on a smaller domain we have

\[ |P - R|_{\mathcal{D}(r-\rho, 2s, \beta)} \leq \alpha^{1-\sigma} s^{\kappa} \varepsilon, \]  

(21)

where \( \alpha = 9s^{\frac{\sigma}{2d+\sigma}} \).

Our next goal is to solve the “intermediate” homological equation (13). To this end we add the average of (18) to \( N \), i.e. we let

\[ \tilde{N} = N + P_0(y, p, q, \lambda), \]  

(22)

where

\[
\begin{align*}
P_0(y, p, q, \lambda) &= \sum_{j=1}^{d} P_j(\lambda) q^j + \sum_{j=0}^{[\frac{d+1}{2}]} Q_j(\lambda) p q^j + R(\lambda) p^2 \\
&\quad + P_{00210} p^2 q + (P_{01000} | y)
\end{align*}
\]

is given by (19) with

\[
\begin{align*}
P_j(\lambda) &= \sum_{(2d-2)h_1+\cdots+2h_{d-1}+(d+1)h_d \leq 2d-2j} P_{00j1h}^{\lambda h}, \\
Q_j(\lambda) &= \sum_{(2d-2)h_1+\cdots+2h_{d-1}+(d+1)h_d \leq d+1-2j} P_{001j1h}^{\lambda h}, \\
R(\lambda) &= \sum_{(2d-2)h_1+\cdots+2h_{d-1}+(d+1)h_d \leq 2} P_{0020h}^{\lambda h}.
\end{align*}
\]

Here and below we completely suppress the \( \omega \)-dependence, in particular the coefficients \( P_{0ijh} = P_{0ijh}(\omega) \) are functions on \( U = U_\beta(\mathcal{O}_1') \). For future use we also define \( Q_j(\lambda) \equiv 0 \) for \( j > \frac{d+1}{2} \). Since we cannot solve (13) completely, we let

\[
F = \sum_{0 \leq |k| \leq K} F_k e^{i(k|x)},
\]

\[
F_k = \sum_{m \leq 2d} F_{km} = \sum_{m \leq 2d} \left( \sum_{\|l,i,j,h\| = m} F_{l ijh} y^l p^i q^j \lambda^h \right)
\]
be the solution of

\[ N + R + \{N, F\} = \tilde{N} \pmod{\mathcal{F}_{2d}}, \]

i.e. up to weighted order \(2d\). The coefficients of the function \(F\) can be defined inductively by

\[ i(k | \omega)F_{km} = P_{km} + \{N_0, F_{k,m+1-d}\}, \]

where \(N_0 = \frac{a}{2} p^2 q + \frac{b}{d!} q^d + \sum_{j=1}^{d-1} \frac{\lambda_j}{j!} q^j + \lambda_d p\). More precisely,

\[ F_{km} = \Delta P_{km} + \sum_{i=1}^{3} \Delta^{i+1} \left\{ \{N_0, \ldots, \{N_0, P_{k,m-i(d-1)}\} \ldots \} \right\}. \quad (23) \]

Here we define \(P_{km} = 0\) if \(m < 0\) and denote \(\Delta = \frac{1}{i(k | \omega)}\) for simplicity. We stress that, since

\[ m - 3(d - 1) \leq 2d - 3(d - 1) = 3 - d \leq 0 \]

the right hand side of (23) contains at most 4 terms.

**Remark 13.** Since we only solved Eq. (13) up to the weighted order \(2d\), the higher order terms

\[ \left\{ N_0, \sum_{0 < |k| \leq K \atop d+1 < m \leq 2d} F_{km} e^{i(k|x)} \right\} \]

have to be included in the new perturbation.

To estimate the nested Poisson brackets in (23) we work on the nested domains

\[ D^i = D \left( r - \frac{3+i}{4} \rho, \frac{1}{1+i}s, \frac{\beta}{2} \right) \subset D^1 = D \left( r - \rho, \frac{1}{2}s, \frac{\beta}{2} \right), \quad i = 1, \ldots, 5 \]

and later use the four domains

\[ D^{ix} = D \left( r - \frac{11+i}{4} \rho, 2^{1-i} zs, \frac{7-i}{12} \beta \right) \subset D^x = D \left( r - 3 \rho, zs, \frac{\beta}{2} \right), \quad i = 1, \ldots, 4 \]
to define the normalising co-ordinate transformation. For Poisson brackets with $N_0$ we have the inequality

$$\{|N_0, G\}|_{D^i} = \left| \frac{\partial N_0}{\partial q} \frac{\partial G}{\partial p} - \frac{\partial N_0}{\partial p} \frac{\partial G}{\partial q} \right|_{D^i} \leq s^{d-1}|G|_{D^{i-1}}.$$ 

**Lemma 14.** Under the conditions of Lemma 10 we have $|F|_{D^4} \leq \varepsilon$.

By the Cauchy estimate we have

$$\left| \frac{\partial^{l+i+j}|F|}{\partial y^l \partial p^i \partial q^j} \right|_{\mathcal{D}^4} \leq s^{-m} \varepsilon$$

(24)

if $\| (l, i, j, h) \| \leq m$. Denote by

$$\| X_F \|_{\mathcal{D}} := \max \left\{ \left| \frac{\partial F}{\partial y} \right|_{\mathcal{D}} , s^{-2d} \left| \frac{\partial F}{\partial x} \right|_{\mathcal{D}} , s^{-2d+2} \left| \frac{\partial F}{\partial q} \right|_{\mathcal{D}} , s^{-d+1} \left| \frac{\partial F}{\partial p} \right|_{\mathcal{D}} \right\}$$

$$\uparrow D_\mu X_F \uparrow_{\mathcal{D}} := \max_{|l|+i+j \leq \mu} \left\{ \left| \frac{\partial^{l+i+j}|G|}{\partial y^l \partial p^i \partial q^j} \right|_{\mathcal{D}} \right\} \text{ for } \mu \geq 1$$

where $G$ stands for either of $\frac{\partial F}{\partial y}, \frac{\partial F}{\partial x}, \frac{\partial F}{\partial q}, \frac{\partial F}{\partial p}$. From the Cauchy estimates we obtain a bound $c_\sigma^\sigma$ of the Hamiltonian vector field $X_F$. As $F$ is a polynomial in $y, p$ and $q$ with weighted order $2d$, such a bound even holds for the partial derivatives $\frac{\partial^{l+i+j}|F|}{\partial y^l \partial p^i \partial q^j} X_F$.

**Lemma 15.** Under the conditions of Lemma 10 we have

$$\| X_F \|_{\mathcal{D}^5} \leq s^\sigma, \quad \uparrow D_\mu X_F \uparrow_{\mathcal{D}^5} \leq s^\sigma \quad \forall \mu \geq 1.$$ 

Hence, the flow $\phi_t^F$ of $X_F$ satisfies $\| \phi_t^F - \text{id} \|_{\mathcal{D}^5} \leq c|t|s^\sigma$ as well, i.e. the first, second, third and fourth component of $\phi_t^F - \text{id}$ are bounded by $c|t|s^\sigma, c|t|\varepsilon, c|t|s^{-2}\varepsilon$ and $c|t|s^{d-1+\sigma}$, respectively. Therefore, the inequality

$$\frac{2d(\sigma-\varepsilon)}{s^{2d+\sigma}} \leq \frac{1}{2c}$$

implies that, for $-1 \leq t \leq 1$, the flow $\phi_t^F$ not only maps $\mathcal{D}^5$ into $\mathcal{D}^4$, but also maps $\mathcal{D}^{2x}$ into $\mathcal{D}^2$. Here we slightly abuse notation in that the same symbol $\phi_t^F$ is used for the mapping acting as the identity in the fifth and sixth component. Since $\varepsilon = s^{2d+\sigma}$ we
obtain from (24) the following estimate for $\varphi_F = \varphi_{l=1}$. The norm for $\varphi_F$ is defined by

$$
\|\varphi_F\|_{C^{l, i, j}(\mathcal{D})} = \max_{0 \leq l, i, j \leq 1} \left\| \hat{\partial}^{l+i+j} \varphi_F \right\|_{\mathcal{D}}.
$$

**Lemma 16.** For any given $l, i, j$ there is a constant $s_0$, depending only on $n, \varepsilon$ and $|l| + i + j$, such that if $s \leq s_0$

$$
\|\varphi_F - \text{id}\|_{C^{l, i, j}(\mathcal{D}_{2n})} \leq s^\sigma.
$$

4.1.3. Transformation of $\tilde{N}$ into normal form

So far we have solved the small divisor problem (13) to construct a symplectic change of variables $\varphi_F$ that transforms away the $x$-dependence of the lower order terms entering in $\tilde{N}$. The second part of the iteration step consists in finding a symplectic change of variables $\phi_1 \circ \phi_2$ which transforms $\tilde{N}$ of (22) into the normal form (2) up to some small terms, i.e. $\tilde{N} \circ \phi_1 \circ \phi_2 = N_+ + O(\varepsilon_+)$.

Since $\tilde{N}$ and $N_+$ do not depend on the angular variables $x \in \mathbb{T}^n$, their flows leave the conjugate actions $y \in \mathbb{R}^n$ fixed and define two one-degree-of-freedom systems in the remaining variables $p$ and $q$. As shown in [8,9] one can apply the machinery of (planar) singularity theory to solve normalisation problems (like the passage from $\tilde{N}$ to $N_+$) in one degree of freedom. In fact, we do not have to rely on this heavy machinery, but can derive the necessary transformations $\phi_1$ and $\phi_2$ in an explicit way.

First we use the shear transformation

$$
\phi_1 : \begin{cases}
q_1 = q, \\
p_1 = p + (a + 2P_{00210})^{-1} \sum_{j=1}^{[d+1]} Q_j(\lambda) q_j^{-1}
\end{cases}
$$

(25)

to kill the crossing terms $\sum_{j=1}^{[d+1]} Q_j(\lambda)pq^j$ in $\tilde{N}$ (see (22)). Note that the term with $j = 0$ cannot be removed. We arrive at

$$
\tilde{N} \circ \phi_1 = (\omega_+ | y) + \frac{a_+}{2} p_1^2 q_1 + \frac{b_+}{d!} q_1^d + R(\lambda)p_1^2 \\
+ (\lambda_d + Q_0(\lambda)) p_1 + \sum_{j=1}^{d-1} \left( \frac{\lambda j}{j!} + P_j(\lambda) - \frac{\lambda d}{a_+} Q_{j+1}(\lambda) \right) q_1^j \\
- \frac{Q_0(\lambda) + 2R(\lambda)}{a_+} \sum_j Q_j(\lambda) q_1^{j-1}
$$
\[ + \frac{a_+q_1 - 2R(\lambda)}{2a_+^2} \left( \sum_j Q_j(\lambda)q_1^{j-1} \right) \]

\[ =: \tilde{N} - \tilde{P} \quad (26) \]

with \( \omega_+ = \omega + P_{01000}, a_+ = a + 2P_{00210} \) and \( b_+ = b + d!P_{000d0} \). In this definition \( \tilde{N} \) contains the six terms in the first two lines and \( \tilde{P} \) abbreviates the remaining terms. Recall that we defined \( Q_j(\lambda) \equiv 0 \) for \( j > d + \frac{1}{2} \).

**Remark 17.** The higher order terms entering \( \tilde{P} \) in (26) can be estimated by \( \varepsilon^2 s^2 d \leq \frac{s^2 d^2 + 2s}{2} \) and can therefore be included in the new perturbation.

The remaining term \( R(\lambda)p_1^2 \) of weighted order \( 2d - 2 \) in the part \( \tilde{N} \) of (26) can be dealt with by the translation

\[ \phi_2 : \begin{cases} q_2 = q_1 + \frac{2}{a_+} R(\lambda), \\ p_2 = p_1, \end{cases} \]

yielding

\[ \tilde{N} \circ \phi_1 \circ \phi_2 = N_+ - \tilde{P} \circ \phi_2. \]

Since the parameter transformation

\[ \lambda_i^+ = \lambda_i + i! P_i(\lambda) - \frac{\lambda_i}{a_+} Q_{i+1}(\lambda) \]

\[ + \sum_{j=i+1}^{d-1} \frac{(-1)^{j-i}}{(j-i)!a_+^{j-i}} (2R(\lambda))^{j-i} \left( \lambda_j + j! P_j(\lambda) - \frac{\lambda_j}{a_+} Q_{j+1}(\lambda) \right) \]

\[ \lambda_d^+ = \lambda_d + Q_0(\lambda) \]

has a nonsingular Jacobian, we can (locally) replace \( \lambda \) by \( \lambda^+ \) in the next KAM-step. Thus we get the desired new normal form.

**4.1.4. Estimates of the iteration step**

We now compose our map \( \Psi : D_+ \rightarrow D \) using \( D_+ \subseteq D^{4x} \) and \( D^x \subseteq D \). We have already remarked that \( \phi_F(D^{2x}) \subseteq D^x \). The inequalities

\[ |P_{01000}|_U \leq \frac{\varepsilon}{s^2 d} < \frac{\beta}{12}, \]

\[ \left| a_+^{-1} \sum_j Q_j(\lambda)q_1^{j-1} \right| \leq \frac{\varepsilon}{s^d + 1} < \frac{(xs)^{d-1}}{4} \]
imply \( \phi_1 : \mathcal{D}^{3x} \to \mathcal{D}^{2x} \) where we have subsumed \( \omega \mapsto \omega_+ \) into this mapping. Similarly we subsume \( \lambda \mapsto \lambda^+ \) into \( \phi_2 \) and obtain \( \phi_2 : \mathcal{D}^{4x} \to \mathcal{D}^{3x} \) from

\[
|2a_+^{-1} \mathcal{R}(\lambda)| \leq s^{2+\sigma} < \frac{(zs)^2}{8},
\]

\[
|i!P_i(\lambda) - \lambda d a_+^{-1} Q_{i+1}(\lambda)| \leq \frac{\varepsilon}{s^{2i}} < \frac{(zs)^{2d-2i}}{8},
\]

\[
|Q_0(\lambda)| \leq \frac{\varepsilon}{s^{d-1}} < \frac{(zs)^{d+1}}{8}.
\]

Together we have that

\[
\Psi = \phi_p \circ \phi_1 \circ \phi_2 : \mathcal{D}^{4x} \to \mathcal{D}^x.
\]

This defines the desired co-ordinate transformation for one iteration step. Similar to Lemma 16, we have the estimates for \( \Psi \):

**Lemma 18.** For any given \( l, i, j \) there is a constant \( s_0 \), depending only on \( n, \varepsilon \) and \( |l| + i + j \), such that if \( s \leq s_0 \)

\[
||\Psi - \text{id}||_{C^{ij}(D^{4x})} \leq s^{\sigma}.
\]

The new perturbation is

\[
P_+ = \tilde{P} \circ \phi_1 \circ \phi_2 + N_0, \quad \sum_{0 \leq |k| \leq K} \sum_{d+1 < m \leq 2d} F_{km} e^{i(k|x)} \circ \phi_1 \circ \phi_2 - \tilde{P} \circ \phi_2, \tag{28}
\]

where \( \tilde{P} \) and \( \tilde{P} \) were defined in (11) and (26), respectively. Following [11] almost verbatim, we estimate the various terms in (28) and obtain

\[
|P_+|_{D_+} \leq |P_+|_{D^{4x}} \leq c \lambda^{(1-\sigma)(\sigma-\kappa)} s^{K \varepsilon} \varepsilon < \varepsilon_+,
\]

where we used \( \lambda^{(1-\sigma)(\sigma-\kappa)} \) to absorb the “accumulated constant \( c \)”. Moreover, as the domain \( D_+ \) is again smaller than \( D^{4x} \), we have for \( m := ||(l, i, j, h)|| \leq 2d \)

\[
\left| \frac{\partial^{[l]+i+j+[h]} P_+}{\partial \gamma^l \partial p^i \partial q^j \partial \lambda^h} \right|_{D_+} \leq s^{-m} \varepsilon_+
\]

by Cauchy’s inequality. This concludes the proof of Lemma 10. \( \square \)
4.2. Iteration and convergence

In the previous subsection we were concerned with one step of the iteration process. Thus, given a small perturbation \( H_v = N_v + P_v \) of our normal form \( N_v \), we now know how to construct a co-ordinate change \( \Psi_v \) such that \( H_{v+1} := H_v \circ \Psi_v \) is an even smaller perturbation of the adapted normal form \( N_{v+1} \). Our next aim is to show that this process “converges”, leading to a well-defined limit \( H_\infty = N_\infty + P_\infty \) where the perturbing term \( P_\infty \) no longer forms an obstruction for the desired conclusions.

By composition \( \Phi_{v+1} := \Psi_0 \circ \Psi_1 \circ \cdots \circ \Psi_v \) we obtain a co-ordinate transformation that turns the given \( H_0 = N_0 + P_0 \) into \( H_0 \circ \Phi_{v+1} = N_{v+1} + P_{v+1} \). Our aim is to find a “limit” \( \Phi_\infty \) with

\[
H_0 \circ \Phi_\infty = N_\infty + P_\infty.
\]

The occurrence of \( P_\infty \) reflects that \( \lim_{v \to \infty} \Phi_v \) is only defined on

\[
\cap D_v = U_{\frac{n}{2}}(\mathbb{T}^n) \times \{0\} \times \{0\} \times \{0\} \times O_1.
\]

To obtain the desired convergence we will need a bound on the \( C^\mu \)-norm

\[
\| \Phi_v \|_{C^\mu(D_v)} = \max_{|l|+i+j \leq \mu} \left\| \frac{\partial^{[l]+i+j} \Phi_v}{\partial y^l \partial p^i \partial q^j} \right\|_{D_v}.
\]

**Lemma 19.** A constant \( c > 0 \) exists, depending only on \( n, \tau, d \) and \( \mu \), such that

\[
\| \Phi_v \|_{C^\mu(D^2_v)} \leq c \quad \text{for every } v \in \mathbb{N}.
\]

In the case of e.g. normally elliptic tori the transformations one works with form a group, cf. [36]. This allows to concentrate on the coefficient functions and to use the limits of these coefficient functions to define the desired limit transformation. However, in the present situation the co-ordinate changes \( \Psi_v \) do not form a group. Indeed, the bifurcating tori require higher order terms, which in turn have to be dealt with by both the Hamiltonian \( F_v \) that generates the first part \( q_v \) of the co-ordinate transformation \( \Psi_v \) and by its second part defined explicitly in (25) and (27). The problem is now that one cannot restrict to the fixed weighted order \( 2d \) in \((y, p, q, \lambda)\) as the composition already of \( \Psi_v \) and \( \Psi_{v+1} \) would increase this order to \( 4d \). Therefore, we have to pass to a polynomial truncation of fixed degree in order to define \( \Phi_\infty \) by means of limits of coefficient functions. This truncation has to satisfy the following conditions.

1. We do not want to destroy the symplectic structure, i.e. the “truncated transformations” \( \tilde{\Psi}_v \) have to be symplecto-morphisms as well.
(2) The estimates implied by Lemma 10 should remain valid after the transformed Hamiltonian functions $H_0 \circ \Phi_v$ are replaced by the Hamiltonians $H_0 \circ Y_v$.

In view of the first condition we do not simply truncate $\Phi_v$, but truncate a generating function $S_\nu$ to define $Y_v$ as follows. Since $\Phi_v : (x, y, p, q, \lambda, \omega) \mapsto (X, Y, P, Q)$ is a symplecto-morphism for fixed $(\lambda, \omega)$, the 1-form

$$\sum_{i=1}^{n} (y_i - Y_i) \, dx_i + (X_i - x_i) \, dy_i + (Q - q) \, dP + (p - P) \, dq$$  \hspace{1cm} (29)$$

is closed and can therefore be written as $dS_\nu$. Indeed, being composed from finitely many translations $\phi_0^\mu$, shear transformations $\phi_1^\mu$ and time one maps $\phi_\nu^\mu$, the transformation $\Phi_v$ is homotopic to the identity. Thus, the closed one-form (29) is exact, i.e. $S_\nu$ is one-valued. Note that the function $S_\nu = S_\nu(x, Y, P, q)$ itself is only determined up to a constant and that all partial derivatives are $2\pi$-periodic in the toral co-ordinates $x_1, \ldots, x_n$.

Because of the second condition we define the truncation $\tilde{S}_v$ of $S_\nu$ to be of order $d+1$ in $(Y, P, q, \lambda)$. Furthermore we drop all terms that involve more than one derivative with respect to parameters $\lambda_j$. On the other hand we do not truncate in $x$ or $\omega$.

To be precise, we write

$$\Phi_v(x, y, p, q, \lambda, \omega) = ((x, y, p, q) + W_v(x, y, p, q, \lambda, \omega), \lambda + \tilde{\Lambda}_v(\lambda, \omega), \omega + \tilde{\Omega}_v(\lambda, \omega))$$

and let $F_v : D_v \longrightarrow D_0$ denote the transformation of $(x, y, p, q, \lambda, \omega)$ into

$$(x, y + W_v^2(x, y, p, q, \lambda, \omega), p + W_v^3(x, y, p, q, \lambda, \omega), q, \lambda, \omega) \overset{!}{=} (x, Y, P, q, \lambda, \omega)$$

and $G_v := F_v^{-1}$. The truncations $\tilde{S}_v$ are polynomials in $Y, P, q$ and $\lambda$, the coefficients of which are holomorphic functions in $x$ and $\omega$. To truncate we write $S_v$ as a Taylor series at $F_v(x, 0, 0, 0, 0, \omega) =: (x, Y_v, P_v, 0, 0, \omega)$. Therefore,

$$S_v^{ljh}(x, \omega) = \frac{\partial |l|+j+i+|h|}{\partial Y^l \partial P^i \partial q^j \partial \lambda^h} S_v(x, Y_v, P_v, 0, 0, \omega),$$

and we define

$$\tilde{S}_v(x, Y, P, q, \lambda, \omega) := \sum_{|l|+i+j=0}^{d+1} \sum_{|h|=0}^{\min(|l|+i+j, 1)} S_v^{ljh}(x, \omega) \cdot (Y - Y_v)^l (P - P_v)^i q^j \lambda^h.$$
Lemma 20. Under the conditions of Lemma 10 the sequence \((\tilde{S}_v)_{v \in \mathbb{N}}\) of truncations is uniformly convergent on \(\overline{U_{\frac{r}{2}}(\mathbb{T}^n)} \times \mathcal{O}_1\).

Using the Inverse Approximation Lemma, cf. [45], we obtain Whitney-\(C^\infty\)-smooth limit functions \(\tilde{S}^{lijh}_\infty\) on \(\overline{U_{\frac{r}{2}}(\mathbb{T}^n)} \times \mathcal{O}_1\). They constitute the coefficients of a generating function
\[
\tilde{S}_\infty : \overline{U_{\frac{r}{2}}(\mathbb{T}^n)} \times \mathbb{C}^n \times \mathbb{C}^2 \times \mathbb{C}^{d-2} \times \mathcal{O}_1' \longrightarrow \mathbb{C}
\]
which is analytic (since polynomial) in \(x, Y, P, q\) and \(\lambda\). With Whitney’s Extension Theorem, cf. [42], we get \(\Phi_\infty(x, Y, P, q, \lambda, \omega)\) for all \(\omega \in \mathbb{R}^n\). This defines for every \((\lambda, \omega)\) a symplecto-morphism on \(\mathbb{T}^n \times \mathbb{R}^n \times \mathbb{R}^2\). To obtain \(\Phi_\infty\), we have to complete these symplecto-morphisms by \(\text{id} + (\tilde{\Lambda}_\infty, \tilde{\Omega}_\infty) = \text{id} + \lim_{v \to \infty} (\tilde{\Lambda}_v, \tilde{\Omega}_v)\). This latter convergence to Whitney-\(C^\infty\)-smooth functions is an immediate consequence of Lemma 19 and the Inverse Approximation Lemma.

To conclude the proof of Theorem 1 we apply the Inverse Approximation Lemma to the coefficient functions \(a_v, b_v\) of the normal forms \(N_v\) and obtain a Whitney-\(C^\infty\)-smooth Hamiltonian function \(N_\infty\) which is (again) analytic in \(y, p, q\) and \(\lambda\). Letting \(P_\infty := H_0 \circ \Phi_\infty - N_\infty\) we have, according to our choice of the truncations \(S_v\) of \(S_v\) at order \(d + 1\),

\[
P_{\infty,lijh} = \lim_{v \to \infty} P_{v,lijh}
\]
as long as \(|h| \leq 1\) and \(|l| + i + j \leq d\). In particular we can conclude that these all vanish for weighted order \(\|(l, i, j, h)\| \leq 2d\). This concludes the proof of Theorem 1. \(\square\)

Remark 21. Instead of working with a miniversal unfolding (16) of the singularities \(D_k\) we could have used more parameters \(\lambda_{d+1}, \ldots, \lambda_e\) with \(e = d + \lfloor \frac{d+1}{2} \rfloor + 1\) and include in the normal form (16) the terms

\[
+ \sum_{j=1}^{\lfloor \frac{d+1}{2} \rfloor} \lambda_{d+j} pq^j + \lambda_e p^2
\]

with weights \((d - 1, d - 3, \ldots, 2)\) on \((\lambda_{d+1}, \lambda_{d+2}, \ldots, \lambda_e)\). This would keep the singularity-theoretic Section 4.1.3 out of the iteration procedure as explicit transformations of the form (25) and (27) would only be used once, to turn the final versal unfolding \(N_\infty\) into miniversal form (2), thereby getting rid of the additional parameters \((\lambda_{d+1}, \ldots, \lambda_e)\). This possible separation of the KAM-procedure from the adjustments dictated by (planar) singularity theory suggests generalizations to quasi-periodic bifurcations governed not only by the remaining simple singularities, cf. [24], but also to quasi-homogeneous or even more complicated singularities with modal parameters.
Acknowledgements

This research was supported by the Alexander von Humboldt Foundation and the FOM Grant MFGBb1.

References