Bifurcations of Normally Parabolic Tori in Hamiltonian Systems

Henk W. Broer

Instituut voor Wiskunde en Informatica, Rijksuniversiteit Groningen
Blauwborgje 3, 9747 AC Groningen, The Netherlands

Heinz Hanßmann

Institut für Reine und Angewandte Mathematik der RWTH Aachen
52056 Aachen, Germany

Jiangong You

Department of Mathematics, Nanjing University
Nanjing 210093, P. R. China

6 June 2003

Abstract

We consider perturbations of integrable Hamiltonian systems in the neighbourhood of normally parabolic invariant tori. Under appropriate transversality conditions the tori in the unperturbed system bifurcate according to a (generalized) cuspoid catastrophe. Combining techniques of KAM–theory and singularity theory we show that such bifurcation scenarios survive the perturbation on large Cantor sets. Applications to rigid body dynamics and forced oscillators are pointed out.

1 Setting of the problem

KAM–theory is usually committed in the context of nearly integrable Hamiltonian systems. The, by now, classical part of the theory, see *e.g.* [Pöschel;82], establishes the persistence of quasi–periodic invariant tori of Lagrangian type under small perturbations away from integrability. Also persistence of normally hyperbolic and normally elliptic tori has been studied, cf. *e.g.* [Pöschel;89], [Broer,Huitema,Takens;90]. 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 up to date overview of these and related results see [Broer,Huitema,Sevryuk;96].

In the dissipative (general) context KAM–theory is known to need parameters for the persistence of quasi–periodic invariant tori. In particular parameters are needed to keep track of the frequencies of the torus while perturbing. See [Moser;67], [Broer,Huitema,Takens;90], or again, [Broer,Huitema,Sevryuk;96]. In [Braaksma,Broer,Huitema;90], moreover, a systematic study was made of simple bifurcations of quasi–periodic tori, where the normal hyperbolicity is mildly violated. It turns out that persistent models exist in analogy of the standard bifurcation models of equilibria and closed orbits. In this analogy, certain continua in the parameter space have to be replaced by Cantor like sets of large Hausdorff measure. Notably this whole programme needs the assumption of reducibility of the normal linear part to a constant Floquet matrix by a change of variables. For a discussion see Section 7 of [Broer,Huitema,Takens;90], and see [Wagener;98], [Broer,Takens,Wagener;99], [Takens,Wagener;00], [Broer,Wagener;00] concerning a Hopf–like bifurcation of invariant tori where this assumption is not valid.

The present paper aims to similarly treat a number of bifurcations in the Hamiltonian context, generalizing a first such result in [Hanßmann;98] to arbitrary (finite) co–dimension. The normal linear part of the bifurcating invariant tori is supposed to be parabolic, *i.e.* of the form $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. We base our analysis on the techniques used in [Hanßmann;98] and [You;98]. Let us briefly sketch the setting of the present problem.

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

$$\sigma \;=\; \sum_{i=1}^{n} \mathrm{d}x_i \wedge \mathrm{d}y_i \;+\; \mathrm{d}q \wedge \mathrm{d}p \quad .$$

We are looking at perturbations of a Hamiltonian system for which the torus $\mathbb{T}^n \times \{0\} \times \{0\}$ is invariant and the normal linear part has the (nilpotent) parabolic form

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} \;=\; \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}$$

for some $a \neq 0$. Mimicking the theory of bifurcations for equilibria and periodic solutions, cf. [Meyer;70,75] or [Broer,Chow,Kim,Vegter;93,95], we add the following assumption. For some integer $d \geq 3$, the expansion of the unperturbed Hamiltonian in the (p, q)–direction has the principal part $\frac{a}{2}p^2 + \frac{b}{d!}q^d$, with $b \neq 0$. Moreover we include parameters, first for the frequencies $\omega_1, \omega_2, \ldots, \omega_n$, and second for the unfolding of the (cuspoidal) singularity q^d , as dictated by singularity theory, cf. [Bröcker,Lander;75]. Indeed, disregarding some co–ordinate changes and reparametrisations, we shall assume that the unperturbed family has the following "integrable" form :

$$N(x, y, p, q; \lambda, \omega) \;=\; (\omega \mid y) \;+\; \frac{a(\omega)}{2}p^2 \;+\; \frac{b(\omega)}{d!}q^d \;+\; \sum_{j=1}^{d-2} \frac{\lambda_j}{j!}q^j \quad , \tag{1.1}$$

where $(.. \mid ..)$ denotes the standard inner product on $\mathbb{R}^n$. This family first of all has a continuum of normally parabolic invariant tori

$$\mathbb{T}^n \times \{0\} \times \{0\} \times \{0\} \times \mathcal{O} \;\subseteq\; \mathbb{T}^n \times \mathbb{R}^n \times \mathbb{R}^2 \times \Lambda \times \mathcal{O} \quad ,$$

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 and parabolic, corresponding to the cuspoidal hierarchy of singularity theory, cf. [Bröcker,Lander;75], [Arnol'd,Vasil'ev,Goryunov,Lyashko;88]. 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 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 [Pöschel;82] or Section 8 of [Broer,Huitema,Takens;90].

This perturbation problem is not expected to have an affirmative answer for all parameters ω, but again only on a set of Cantor like structure. Indeed, on the vector $\omega \in \mathcal{O}$ we impose Diophantine conditions, saying that

$$|(\omega \mid k)| \geq \frac{\gamma}{|k|^\tau} \qquad \forall_{k \in \mathbb{Z}^n \setminus \{0\}} \quad , \tag{1.2}$$

where $\tau > n - 1$ is fixed and where $\gamma > 0$ is to be chosen later on. The first result of this paper roughly says the following. For values of ω in a Cantor set given by the above restriction, the family $H = N + P$, with P sufficiently small in an appropriate norm, again has such normally parabolic invariant n–tori near $y = 0$, $(p, q) = 0$, $\lambda = 0$. These perturbed tori, moreover, form a Whitney–C^∞–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 ω_0 satisfying (1.2).

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 the cuspoid families carries over to the KAM–setting.

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 λ and frequencies ω this yields a stratification — each stratum of co–dimension k parametrising invariant n–tori that undergo a bifurcation of that same co–dimension.

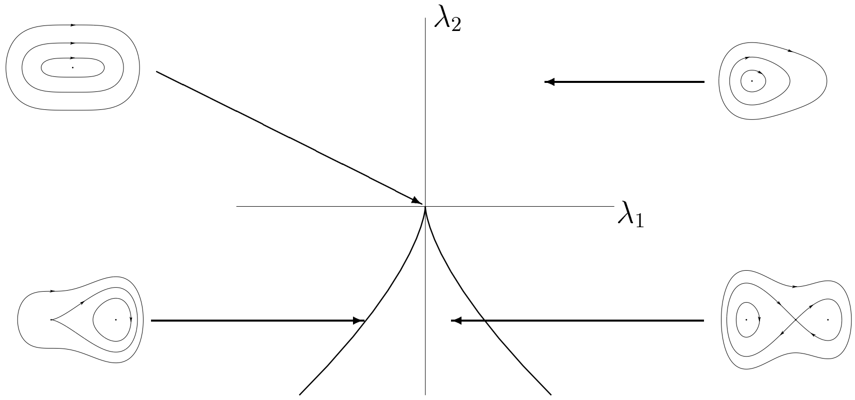

Figure 1: Bifurcation diagram of N for $d = 4$. The phase portraits show the reduced one–degree–of–freedom dynamics of (1.1) for the indicated values of (λ_1, λ_2).

To fix thoughts let us concentrate on the case $d = 4$, see *figure 1* for the bifurcation diagram of the lower dimensional tori defined by N. The point $\lambda_1 = \lambda_2 = 0$ where the two lines $\{ 9\lambda_1^2 + 8\lambda_2^3 = 0 \}$ meet corresponds to the most degenerate parabolic torus. These lines stand for parabolic invariant tori that have some normal form like in (1.1), but with $d = 3$, *i.e.* N undergoes *(quasi–periodic) centre–saddle bifurcations* when these lines are crossed, cf. [Meyer;70,75], [Broer,Chow,Kim,Vegter;93,95], [Hanßmann;98] ; this is related to a subordinate fold catastrophe. Concentrating on the "cusp" case $a \cdot b > 0$, the values of λ with $9\lambda_1^2 > -8\lambda_2^3$ parametrise elliptic tori, while each λ with $9\lambda_1^2 < -8\lambda_2^3$ stands for two elliptic tori and a hyperbolic torus. In particular the system undergoes a *(quasi–perodic) Hamiltonian pitchfork bifurcation* under variation of λ_2 with $\lambda_1 \equiv 0$. When a and b have opposite signs, in the "dual cusp" case, each λ with $9\lambda_1^2 < -8\lambda_2^3$ stands for one elliptic and two hyperbolic tori. This leads to a further bifurcation line $\{ \lambda_1 = 0 , \lambda_2 < 0 \}$ where the two hyperbolic tori have the same energy, have coinciding stable and unstable manifolds and thus get connected by heteroclinic orbits. This *connection bifurcation* is an example of a global bifurcation subordinate to the local bifurcations defined by (1.1).

For the dynamics defined by N there is one bifurcation diagram for each frequency vector ω. Using a Kolmogorov–type non–degeneracy condition, cf. (5.9) 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 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 smooth tori).

Occurrence of the type of bifurcation at hand most often follows from a normalizing 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 the example of the quasi–periodic response problem, sketched below, this situation is clearly illustrated. Sometimes the actual number of parameters is less than the co–dimension of the singularity and we may have to resort to the path formalism. Compare with the periodic case [Meyer;70,75], [Golubitsky,Schaeffer;85], [Golubitsky,Schaeffer,Stewart;88].

2 Results

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 parabolic invariant n–tori of X_H .

Let T be an n–torus and $\mathsf{Y} \subseteq \mathbb{R}^n$, $\mathsf{S} \subseteq \mathbb{R}^2$, $\Lambda \subseteq \mathbb{R}^{d-2}$ be neighbourhoods of the respective origins. By $\mathcal{O}_\gamma$ we denote the set of those frequency vectors $\omega \in \mathcal{O}$ that satisfy the Diophantine condition (1.2). We also need $\mathcal{O}'_\gamma := \{ \omega \in \mathcal{O}_\gamma \mid \mathrm{d}(\omega, \partial\mathcal{O}) \geq \gamma \}$. Furthermore $|..|_A$ stands for the supremum norm on the set A .

Theorem 2.1 *Let the functions* $a, b : \mathcal{O} \longrightarrow \mathbb{R}$ *in the normal form* (1.1) *satisfy* $|a|_\mathcal{O}, |b|_\mathcal{O}, \left|\frac{1}{a}\right|_\mathcal{O}, \left|\frac{1}{b}\right|_\mathcal{O}, |Da|_\mathcal{O}, |Db|_\mathcal{O} < C$ *for some constant* $C > 0$. *Then there exists a small positive constant* ε , *independent of* $\mathcal{O}$, *with the following property. For any analytic perturbation* $H = N + P$ *of* (1.1) *with*

$$|P|_{\mathsf{T}\times\mathsf{Y}\times\mathsf{S}\times\Lambda\times\mathcal{O}} < \varepsilon$$

there exists a C^∞*–diffeomorphism* Φ *on* $\mathsf{T} \times \mathbb{R}^n \times \mathbb{R}^2 \times \mathbb{R}^{d-2} \times \mathcal{O}$ *such that*

1). Φ *is real analytic for fixed* ω .

2). Φ *is symplectic for fixed* (λ, ω) .

3). Φ *is* C^∞*–close to the identity.*

4). *On* $\mathsf{T}\times\mathbb{R}^n\times\mathbb{R}^2\times\mathbb{R}^{d-2}\times\mathcal{O}'_\gamma \cap \Phi^{-1}(\mathsf{T}\times\mathsf{Y}\times\mathsf{S}\times\Lambda\times\mathcal{O})$ *one can split* $H\circ\Phi = N_\infty + P_\infty$ *into an integrable part* N_∞ *and higher order terms* P_∞ . *Here* N_∞ *has the same form as* N , *see* (1.1). *The* x*–dependence is pushed into the higher order terms, i.e.* $\dfrac{\partial^{|l|+i+j+|h|} P_\infty}{\partial y^l \partial p^i \partial q^j \partial \lambda^h}(x, 0, 0, 0, 0, \omega) = 0$ *for all* $(x, \omega) \in \mathsf{T}\times\mathcal{O}'_\gamma$ *and all* l, i, j, h *satisfying* $2d|l| + di + 2j + (2d-2)h_1 + \ldots + 4h_{d-2} \leq 2d$.

We prove this theorem in Section 6, using a KAM iteration scheme. But let us first elaborate its implications.

An immediate consequence is the existence of normally parabolic n–tori at the "origin" $y = p = q = \lambda = 0$. These are parametrised by the Diophantine frequency vectors $\omega \in \mathcal{O}'_\gamma$, *i.e.* they form a *Cantor family*. This Cantor family at $\lambda = 0$ corresponds to the most degenerate normally parabolic 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.

Recall that a subset $S \subseteq M$ of a C^∞–manifold is said to be stratified into finitely many locally closed C^∞–submanifolds (called "strata") $S_k \subseteq M$, $k = 0, \ldots, m$ if $S = \bigcup S_k$ and the topological boundary $\partial S_k \subseteq S$ lies in the union $\bigcup_{l>k} S_l$ for all $k = 0, \ldots, m$. Thus, S_0 is open in S and the complement $S \backslash S_0$ is stratified into $S_1, \ldots, S_m$. It is convenient to choose $m = \dim M$, giving S_k the co–dimension k (within M).

The polynomial normal forms from singularity and catastrophe theory all have semi–algebraic catastrophe and bifurcation sets. This gives the simplest examples of stratified sets. The further complications in the definition of such stratifications largely arise from the fact 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.* [Whitney;65] or [Gibson,Wirthmüller,du Plessis,Looijenga;76]. 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^∞–smooth transformations with respect to this. The corresponding Whitney extensions also are smooth on the whole semi–algebraic set, which brings us in the above setting. The stratification obtained thus will colloquially be referred to as Cantor stratification.

We use Theorem 2.1 to obtain a Cantor stratification in an inductive manner. Near the above Cantor family of most degenerate parabolic 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. Thus, we invoke Theorem 2.1 using a normal form like (1.1) with d replaced by $d-1$, then by $d-2$ and so on until we reach the subordinate quasi–periodic centre–saddle bifurcations. Here we use the cuspoidal hierarchy of singularities of type A_k , cf. [Bröcker,Lander;75], [Arnol'd,Vasil'ev,Goryunov,Lyashko;88]. In Section 4 this is used to prove the following result.

Theorem 2.2 *Under the conditions of Theorem 2.1, there is a Cantor stratification of a Cantor subset of* $\Lambda \times \mathcal{O}$ *of large measure into* $(n+d-k-2)$*–dimensional Cantor sets* C_k , $k = 0, \ldots, d-2$, *such that* C_0 *parametrises Cantor families of elliptic and hyperbolic tori and* C_k , $k = 1, \ldots, d-2$ *parametrise Cantor families of parabolic tori of co–dimension* k .

Remark : The normal form (1.1) 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. For a generic perturbation P this leads to transversal homoclinic orbits. When the system depends on parameters, also homo– and heteroclinic bifurcations are involved. The angle between the stable manifold and the unstable manifold, at such a transversal homoclinic orbit, is expected to be exponentially small. Subordinate to a "primary" homoclinic orbit, variation of the parameters λ may lead to homoclinic bifurcations, involving tangencies between the stable and unstable manifolds, cf. [Kan,Koçak,Yorke;92], [Palis,Takens;93]. Similar observations apply *mutatis mutandi* to homoclinic orbits of parabolic n–tori.

Remark : 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. 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. A model Hamiltonian on $\mathbb{T}^n \times \mathbb{R}^n \times \mathbb{R}^2$ is given by

$$K(x, y, p, q) \;=\; (\omega(y) \mid y) \;+\; \frac{1}{2}p^2 \;+\; \frac{1}{2}q^2 \;-\; \frac{1}{24}q^4 \;-\; y_1 q \quad . \tag{2.3}$$

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) "horn", cf. [Broer,Roussarie,Simó;96], [Broer,Roussarie;01]. In the dual cusp case, for instance, this leads to a bifurcation diagram as sketched in *figure 2a.* The boundary lines of the horn stand for "first" heteroclinic tangencies, as depicted in *figure 2b.* Within the horn secondary heteroclinic tangencies are *abundant.*

To speak of heteroclinic orbits the α– and ω–limit sets, the hyperbolic tori, have to persist. As this happens on Cantor sets of large measure, their intersection is again a large Cantor set (and in particular non–empty). It should be instructive to experiment with concrete perturbations of *e.g.* (2.3).

Remark : 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 [Poénaru;76], to derive adapted unfoldings. As the proof of Theorem 2.1 is of Lie algebra type and hence structure–preserving, cf. [Moser;67], [Broer,Huitema,Sevryuk;96], the result carries over.

On the other hand, every symmetry raises the question : what happens if this symmetry is broken, how does the symmetric system unfold within the space of all systems ? Here the concept of *distinguished parameters* comes into play, see [Golubitsky,Schaeffer;85], [Broer,Chow,Kim,Vegter;93,95]. Within a "complete" unfolding like (1.1) those parameters "μ" that are also present in a symmetric unfolding are distinguished with respect to the parameters "ν" that break the symmetry in that only reparametrisations of the form

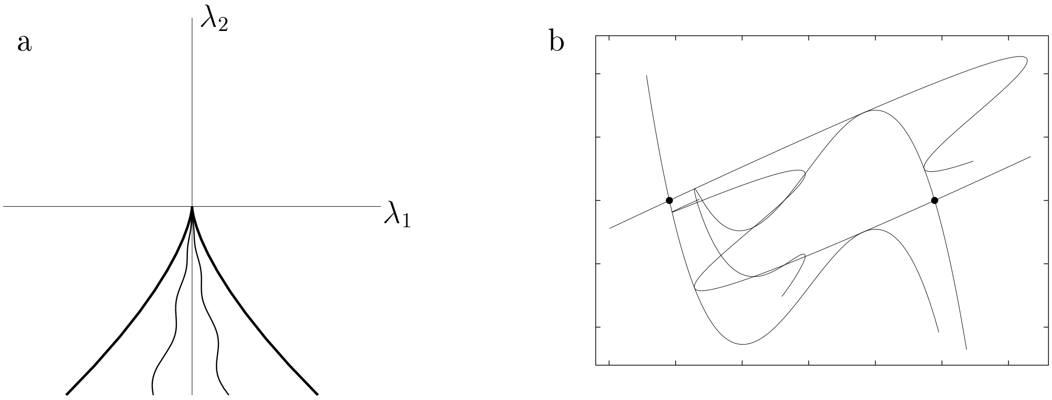

Figure 2: *a).* In the dual cusp case the connection bifurcation at $\{\lambda_1 = 0, \lambda_2 < 0\}$ perturbs to an (exponentially small) open region where heteroclinic orbits occur.
b). A "first" tangency bifurcation of heteroclinic orbits. (Courtesy of Carles Simó.)

$(\mu, \nu) \mapsto (\tilde{\mu}(\mu, \nu), \tilde{\nu}(\nu))$ with $\tilde{\nu}(0) = 0$ are admitted to ensure that the new $\tilde{\mu}$ can still be interpreted as "symmetry parameters" while the new $\tilde{\nu}$ again break the symmetry.

Remark : An important case is if 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.* [Sevryuk;86], [Broer,Huitema;95]. For instance, the normal form (1.1) is reversible with respect to $(x, y, p, q) \mapsto (-x, y, -p, q)$.

The involution $(x, y, p, q) \mapsto (-x, y, p, -q)$ allows to reduce the parameters in the unfolding (1.1), with even d, by half. In this way, *e.g.* the cusp case ($d = 4$) leads to the (quasi–periodic) Hamiltonian pitchfork bifurcation ($\lambda_1 \equiv 0$).

Let $\mathbb{T}^n \times \{0\} \times \{0\}$ be an invariant torus of a Hamiltonian system on $\mathbb{T}^n \times \mathbb{R}^n \times \mathbb{R}^2$ with normal "linearisation" $\frac{1}{2}p^2$. Then any x–independent (whence integrable) perturbation that is invariant under $q \mapsto -q$ leads to an invariant torus close to the origin, and also non–integrable perturbations that are reversible with respect to $(x, y, p, q) \mapsto (-x, y, p, -q)$ lead to invariant tori parametrised by pertinent Cantor sets, see [You;99]. Naturally, the normal behaviour of such tori cannot be controlled and depends on the perturbation at hand.

Remark : A 1–parameter family of periodic orbits that encounters the Floquet multiplier -1 (generically) undergoes a period–doubling bifurcation. Similarly, we expect a frequency–halving bifurcation of invariant tori when one of the normal frequencies enters the lattice

$$\{ \pi \mathrm{i} + 2\pi \mathrm{i}(\omega \mid k) \mid k \in \mathbb{Z}^n \}$$

defined by the (internal) frequency vector ω. Passing to a 2–fold covering space, cf. [Braaksma,Broer,Huitema;90], [Broer,Vegter;92], [Ferrer,Hanßmann,Palacián,Yanguas;02], [Broer,Hanßmann,Jorba,Villanueva,Wagener;02], we obtain a lifted system that is invariant under the π–rotation $(x, y, p, q) \mapsto (x, y, -p, -q)$. Here the normal frequency of the bifurcating torus becomes zero, *i.e.* the normal linear part is either $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 0 & a \\ 0 & 0 \end{smallmatrix}\right)$ with $a \neq 0$. For the latter normally parabolic case we indicate at the end of the next section what our results imply for such frequency–halving bifurcations.

3 Applications

As soon as dynamical systems depend on (external) parameters, bifurcations become a general phenomenon. In the present case of Hamiltonian systems the action variables $y_1, \ldots, y_n$ conjugate to the toral angles on $\mathbb{T}^n$ play the rôle of internal or distinguished parameters, we come back to the resulting possibility of parameter reduction in Section 5 below. Here we illustrate our results with applications to perturbations of the Euler top and forced oscillators and treat the normally parabolic frequency–halving bifurcations.

3.1 Perturbations of the Euler top

The *free rigid body* is a Hamiltonian system that describes the motion of a rigid (*i.e.* non–deformable) body in the absence of external forces or torques. To concentrate on the rotational aspect of the motion one fixes the body at one point (not necessarily the centre of mass), and obtains a Hamiltonian system with three degrees of freedom, cf. [Arnol'd;74]. We speak of an Euler top if the body in addition is *dynamically symmetric*, *i.e.* two moments of inertia are equal.

The Euler top is not only integrable, but even superintegrable, as the energy and the 3 components of the angular momentum define 4 independent integrals of motion. We stress that this is true for the more general free rigid body with a fixed point as well. However, the rotational–precessional motion along the invariant 2–tori is particularly transparent in the dynamically symmetric case of the Euler top where it consists of a pure rotation about its *figure axis*, superposed by a pure precession of the figure axis about the (fixed) direction of the angular momentum.

While the superintegrability simplifies the analysis of the unperturbed motion, the perturbation analysis becomes more involved. The perturbation we have in mind is given by an external force field. For fast motions the kinetic energy of a rigid body is large with respect to the potential energy, and the system may be treated as a perturbation of the Euler top, cf. [Benettin,Fassò;96]. We stress again that there is no conceptual difference between the Euler top and the general free rigid body with a fixed point — the body is not assumed to be symmetric with respect to the figure axis, whence the perturbation analysis takes place in three degrees of freedom.

The 4–parameter family of invariant 2–tori in the unperturbed Euler top system is not structurally stable even with respect to integrable perturbations. In fact, already for the Lagrange top the torque exertet by the perturbing force field causes the angular momentum to move, leading to 3–parameter families of invariant 3–tori, cf. [Arnol'd;74]. Consequently, only 2–dimensional Cantor subfamilies of this 4–parameter family of invariant 2–tori are expected to survive a perturbation, and it is also the perturbation itself that gives rise to the bifurcation scenario, cf. [Hanßmann;95].

In [Hanßmann;96] a perturbation by a $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetric conservative force field is studied. A normal form approach yields a formal 2–torus symmetry, giving rise to an integrable approximation. The 2–torus symmetry of the normal form allows to reduce to a one–degree–of–freedom problem, which turns out to display Hamiltonian pitchfork bifurcations.

The normal form by itself is a perturbation of the Euler top, and the occurring invariant 2–tori may be thought of as surviving this perturbation : the motion (still) consists of a rotation about the *figure axis*, superposed by a precession of the figure axis about the (fixed) direction of the angular momentum. For the majority of initial conditions the perturbation will cause the angular momentum to move, leading to invariant 3–tori of the integrable normal form, cf. [Hanßmann;96]. It is in fact generic for external force fields to have normal forms with 3–parameter families of invariant 3–tori, cf. [Mazzocco;97], where it is also shown that 3–dimensional Cantor families of invariant 3–tori persist the perturbation from the normal form to the original system.

Now the frequency ratio of the invariant 2–tori involved in the Hamiltonian pitchfork bifurcations is determined by the ratio of the actions conjugate to the toral angles, and this same ratio plays the rôle of the bifurcation parameter. Thus, the two distinguished parameters are not sufficient to ensure persistence, the whole bifurcation scenario might fall into a resonance gap. To overcome this "lack of parameter" problem one can consider the inertia tensor as an external parameter, leading to persistence of the quasi–periodic Hamiltonian pitchfork bifurcation for "most" rigid bodies. For a similar treatment of the quasi–periodic centre–saddle bifurcation occurring for "generic" *affine force fields* (without $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry) see [Hanßmann;98].

In fact, one may also consider the coefficients of the external force field as parameters. Within the class of affine force fields

$$G \;=\; -\alpha e_x - \beta e_y - \gamma e_z - axe_x - bye_y - cze_z$$

one can use nonzero values of α and γ to break the discrete symmetry. In this way one obtains the whole "cusp" scenario of (1.1) with $d = 4$, cf. [Nelk;98].

As it is the perturbation itself that gives rise to the bifurcation scenario, the coefficient functions a and b in (1.1) are of the order δ of that perturbation (while the perturbing terms of order δ^2 have to be computed through a first normalizing transformation). We point out in Section 6.3 below that Theorem 2.1 not only does apply to such systems with two time scales, but that the results obtained are in fact even better in this context.

3.2 Applications of the unfolding theory in the response context

One slightly theoretical class of examples is motivated by the oscillator with quasi–periodic forcing

$$\ddot{q} \;=\; -V'_\lambda(q) \;+\; \varepsilon\, f(t) \quad ,$$

with $f(t) = F(t\omega_1, \ldots, t\omega_n)$, for an analytic function $F : \mathbb{T}^n \longrightarrow \mathbb{R}$. We here take the frequency vector $\omega = (\omega_1, \ldots, \omega_n)$ non–resonant or even Diophantine. The family $V = V_\lambda(q)$ of potentials is one of the cuspoid unfoldings described earlier. In this context we may fix ω and only use λ as a multi–parameter. The general aim is to study the possible dynamics near response solutions, quasi–periodic with the same frequency vector ω , see *e.g.* [Stoker;50], [Moser;65], [Broer,Huitema,Sevryuk;96] and the bibliography of the latter reference. First we rewrite the equation in vector field form

$$\begin{aligned} \dot{x} &= \omega \\ \dot{q} &= p \\ \dot{p} &= -V'_\lambda(q) \;+\; \varepsilon\, F(x) \quad , \end{aligned}$$

with phase space $\mathbb{T}^n \times \mathbb{R}^2$. The problem now concerns the dynamical behaviour, in particular quasi–periodic bifurcations of invariant tori close to $\mathbb{T}^n \times \{0\}$. We refer to [Braaksma,Broer;87], [Braaksma,Broer,Huitema;90] for similar bifurcation problems in the general (dissipative) setting. Although these vector fields are volume preserving, they are not necessarily Hamiltonian. To establish the latter property, we modify to

$$\begin{aligned} \dot{x} &= \omega \\ \dot{y} &= \varepsilon\, q\, \nabla F(x) \\ \dot{q} &= p \\ \dot{p} &= -V'_\lambda(q) \;+\; \varepsilon\, F(x) \quad , \end{aligned}$$

which is a Hamiltonian family of vector fields $X = X_\lambda(x, y, p, q)$ on $\mathbb{T}^n \times \mathbb{R}^n \times \mathbb{R}^2$, with corresponding Hamiltonian functions $H_\lambda(x, y, p, q) = \frac{1}{2}p^2 + (\omega \mid y) + V_\lambda(q) - \varepsilon\, qF(x)$. This system clearly fits in our context. To see this consider the family of potentials given by

$$V_\lambda(q) \;=\; q^d \;+\; \sum_{j=1}^{d-2} \lambda_j q^j \quad .$$

Indeed, Theorems 2.1 and 2.2 directy apply to the corresponding systems. For $d = 3$ a quasi–periodic centre–saddle bifurcation occurs, corresponding to a simple fold singularity in the one–parameter family V_λ . This quasi–periodic bifurcation of the invariant tori has an obvious translation to the corresponding families of ω–quasi–periodic response solutions. Similarly for the cusp ($d = 4$) and the higher cuspoids. As it follows from the above theorems, the smooth stratifications for the families of V_λ functions fall apart into Cantor stratifications by the perturbation.

Remark : The general proof in Section 6 below can be easily adapted to this response context, by choosing all conjugacies such that they leave x fixed.

Remark : Compare with the periodic case, where the forcing effects a far milder complication of the unperturbed, smooth stratification. Indeed, in the periodic case only finitely many (strong) resonances have to be excluded, as they would give rise to interruptions of the smooth stratification.

Remark : Symmetry properties like reversibility are optional in this set–up. For example consider the involution $R : (x, y, p, q) \mapsto (-x, y, p, -q)$, and the vector field satisfying $R_*(X) = -X$. This requires that $F(-x) = -F(x)$ and $V_\lambda(-q) = V_\lambda(q)$, which ensues obvious modifications in the normal forms from equivariant singularity theory.

3.3 The frequency–halving bifurcation

Like maximal tori of a Hamiltonian system, normally hyperbolic invariant tori persist under small perturbations provided that the Diophantine conditions (1.2) on the internal frequency vector $\omega = (\omega_1, \ldots, \omega_n)$ are met. For normally elliptic tori the necessary Diophantine conditions involve the normal frequencies $\Omega_1, \ldots, \Omega_m$ and read

$$|k_1\omega_1 + \ldots + k_n\omega_n \; + \; \ell_1\Omega_1 + \ldots + \ell_m\Omega_m| \;\; \geq \;\; \frac{\gamma}{|k|^\tau} \qquad \forall_{k \in \mathbb{Z}^n \setminus \{0\}} \quad \forall_{\ell \in \mathbb{Z}^m, |\ell| \leq 2} \quad , \tag{3.4}$$

where $\tau > n - 1$ is again fixed and $\gamma > 0$. While the conditions with $\ell = 0$ merely rephrase (1.2), the four types of extra conditions exclude possible bifurcations.

For two coefficients of ℓ to be non–zero we need at least two normal degrees of freedom. The violation

$$\Omega_1 \; - \; \Omega_2 \;\; = \;\; \sum_{i=1}^n k_i \omega_i$$

of (3.4) is related to a (quasi–periodic) Hamiltonian Hopf bifurcation. On the other hand, tori with

$$\Omega_1 \; + \; \Omega_2 \;\; = \;\; \sum_{i=1}^n k_i \omega_i$$

persist and do not bifurcate, see [de Jong;99]. In both cases the condition on the normal frequencies is equivalent to

$$\Omega_1 \quad = \quad \pm\Omega_2$$

by a symplectic change of variables, see [Xu,You;01]. The remaining two cases only require one normal degree of freedom. The resonance

$$\Omega \quad = \quad \sum_{\nu=1}^{n} k_i \omega_i \tag{3.5}$$

leads to

$$\Omega \quad = \quad 0$$

after a better choice of the toral angles, cf. [Broer,Hanßmann,Jorba,Villanueva,Wagener;02]. Then either the normal linear part is $\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, or the bifurcating torus is normally parabolic in which case our theory applies. The last possibility to violate (3.4) is

$$2\,\Omega \quad = \quad \sum_{\nu=1}^{n} k_i \omega_i$$

with at least one of the k_i odd. Here we pass to a 2–fold covering space to recover (3.5), cf. [Braaksma,Broer,Huitema;90], [Broer,Hanßmann,Jorba,Villanueva,Wagener;02]. Again there are two cases. We concentrate on the normally parabolic case as the other case (in which the normal linear part vanishes completely) lies outside the scope of the present paper.

The lifted system on the covering space is invariant under the π–rotation $(p, q) \mapsto (-p, -q)$. When passing to the integrable normal form N we aquire an additional reflectional symmetry $(p, q) \mapsto (-p, q)$ since p enters (1.1) only as p^2 . While this symmetry usually has no consequences, in combination with the above π–rotation it now leads to a second reflectional symmetry $(p, q) \mapsto (p, -q)$ whence q only enters squared in (1.1) as well. Thus d is even in the principal part $\frac{a}{2}p^2 + \frac{b}{d!}q^d$ of N and the unfolding terms are even powers of q as well, *i.e.* the normal form reads

$$N(x, y, p, q; \lambda, \omega) \quad = \quad (\omega \mid y) \; + \; \frac{a(\omega)}{2} p^2 \; + \; \frac{b(\omega)}{d!} q^d \; + \; \sum_{j=1}^{\frac{d}{2}-1} \frac{\lambda_j}{j!} q^{2j} \quad . \tag{3.6}$$

This ensures in particular that $(p, q) = (0, 0)$ consists of invariant n–tori, as it is dictated by the system being a 2:1 lift.

The whole perturbation analysis may now be carried out on the covering space. Indeed, the general proof in Section 6 below can as well be carried out with all mappings equivariant with respect to the π–rotation $(p, q) \mapsto (-p, -q)$.

The implications for the bifurcation scenario of the invariant tori are easily obtained from the equivariant form (3.6) of N . For instance, when $d = 4$ we obtain the usual "period–doubling" scenario. Thus, for $a \cdot b > 0$, the invariant torus $(p, q) = (0, 0)$ turns from normally elliptic to normally hyperbolic as λ_1 passes through 0 (recall that λ_1 is here the coefficient of q^2) and a normally elliptic invariant torus with one frequency halved detaches. In the "dual cusp" case $a \cdot b < 0$ the detaching torus with one frequency halved is normally hyperbolic and the invariant torus $(p, q) = (0, 0)$ turns from normally hyperbolic to normally elliptic. See also [Broer,Hanßmann,Jorba,Villanueva,Wagener;02].

4 Proof of Theorem 2.2

Let us first consider the integrable Hamiltonian system with Hamiltonian function (1.1). The quasi–periodic flow induced by the term $(\omega \mid y)$ is superposed by the one–degree–of–freedom system with Hamiltonian

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

We have already seen in Section 3.2 that the latter is an oscillatory system describing the motion of a 1–dimensional particle in the potential

$$V_d(q, \lambda) \;\; := \;\; \frac{b(\omega)}{d!} q^d \; + \; \sum_{j=1}^{d-2} \frac{\lambda_j}{j!} q^j \quad . \tag{4.7}$$

Applying singularity theory to V_d , one obtains the cuspoidal hierarchy of singularities of type A_{k+1} , $k = 0, \ldots, d-2$, cf. [Bröcker,Lander;75], [Arnol'd,Vasil'ev,Goryunov,Lyashko;88]. In our context this means that the parameter space Λ is stratified according to the degeneracy of the occurring parabolic tori (the singularities of type A_1 in the open stratum $S_0 \subseteq \Lambda$ correspond to elliptic and hyperbolic tori).

This allows us to lead the situation around some higher stratum S_k back to Theorem 2.1. For completeness' sake the necessary computations are re–done here. With (4.7) the normal form (1.1) reads

$$N \;\; = \;\; (\omega \mid y) \; + \; \frac{a(\omega)}{2} p^2 \; + \; V_d(q, \lambda)$$

and by a translation $(Y, P, Q) = (y, p, q - c)$ transforms into

$$N \;\; = \;\; (\omega \mid Y) \; + \; \frac{a(\omega)}{2} P^2 \; + \; \frac{b(\omega)}{d!} Q^d \; + \; \frac{c\, b(\omega)}{(d-1)!} Q^{d-1} \; + \; \sum_{j=1}^{d-2} \frac{V_d^{(j)}(c, \lambda)}{j!} Q^j \tag{4.8}$$

where we dropped the constant term $V_d(c, \lambda)$. Here $c = c(\lambda_{k+1}, \lambda_{k+2}, \ldots, \lambda_{d-2})$ only depends on the last $d - k - 2$ parameters (for some k to be chosen) and $V_d^{(j)}$ denotes the

jth derivative of V_d with respect to q. The idea is to choose c on the stratum of parabolic tori of co–dimension k, which can be parametrised by the last $d-k-2$ parameters.

Let us illustrate the situation in the cusp case, $d=4$, $k=1$, for definiteness with positive a and b. Choosing $c := \pm\sqrt{\frac{-2\lambda_2}{b(\omega)}}$ equation (4.8) reads

$$N \;=\; (\omega \mid Y) \;+\; \frac{a(\omega)}{2}P^2 \;+\; \frac{b(\omega)}{24}Q^4 \;\pm\; \sqrt{\frac{-2\lambda_2}{b(\omega)}}\frac{b(\omega)}{6}Q^3 \;+\; (\lambda_1 \pm \frac{2\lambda_2}{3}\sqrt{\frac{-2\lambda_2}{b(\omega)}})Q \quad .$$

Fixing $\lambda_2 = -\hat{\lambda}_2 < 0$ and varying λ_1 around $\hat{\lambda}_1 := \pm\frac{2}{3}\hat{\lambda}_2\sqrt{\frac{\pm 2\hat{\lambda}_2}{b(\omega)}}$ we get an unfolding of the parabolic torus at $(p,q,\lambda) = (0,c,\hat{\lambda}_1,-\hat{\lambda}_2)$. Let us make this explicit writing

$$\Lambda_1 \;:=\; \lambda_1 - \hat{\lambda}_1 \quad .$$

To apply Theorem 2.1 we have to show that the additional term $\frac{b(\omega)}{24}Q^4$ may be treated as a perturbation. This can be achieved by a scaling

$$(y,p,q,\lambda_1,\Omega) \;=\; (\hat{\lambda}_2^{-\frac{11}{4}}Y, \hat{\lambda}_2^{-\frac{7}{4}}P, \hat{\lambda}_2^{-1}Q, \hat{\lambda}_2^{-\frac{5}{2}}\Lambda_1, \hat{\lambda}_2^{-\frac{3}{4}}\omega)$$

and then dividing N by $\hat{\lambda}_2^{7/2}$. Note that this scaling amounts to rescaling time by $\hat{\lambda}_2^{3/4}$ and to a global multiplication of the symplectic structure by $\hat{\lambda}_2^{-11/4}$. Indeed, we cannot scale $x \in \mathsf{T}$, and therefore had also to rescale the frequency ; even if ω varies in a compact domain the "new" Ω grows unboundedly for $\hat{\lambda}_2 \to 0$. This is the reason why we formulate our theorems for possibly unbounded frequency domains $\mathcal{O}$.

We have now transformed N into the normal form of the quasi–periodic centre–saddle bifurcation, perturbed by $\sqrt{\hat{\lambda}_2}\frac{b(\omega)}{24}q^4$ — which becomes arbitrarily small for $\hat{\lambda}_2 \to 0$, *i.e.* the closer we get to the central bifurcation at $(\lambda_1,\lambda_2) = (0,0)$. In fact, *figure 1* also lies at the basis of the stratification to subordinate bifurcations for general d as the parabolic tori involved in subordinate bifurcations of co–dimension $d-1$ occur at points $(0,c,\hat{\lambda}) \neq (0,0,0)$ satisfying

$$V_d'(c,\hat{\lambda}) \;=\; V_d''(c,\hat{\lambda}) \;=\; \ldots \;=\; V_d^{(d-3)}(c,\hat{\lambda}) \;=\; 0 \quad ,$$

whence $c = \pm\sqrt{\frac{-2\hat{\lambda}_{d-2}}{b(\omega)}}$ and recursively

$$\hat{\lambda}_j \;=\; -\frac{b(\omega)}{(d-j)!}c^{d-j} \;-\; \frac{\hat{\lambda}_{d-2}}{(d-j-2)!}c^{d-j-2} \;-\; \ldots \;-\; \hat{\lambda}_{j+1}c \quad , \quad j = 1,\ldots,d-3 \quad .$$

The projection to the $(\lambda_{d-3},\lambda_{d-2})$–space again is the cusp.

For general $k < d-2$, the values $\hat{\lambda}_1,\ldots,\hat{\lambda}_k$ and c are determined according to

$$V_d'(c,\hat{\lambda}) \;=\; V_d''(c,\hat{\lambda}) \;=\; \ldots \;=\; V_d^{(k+1)}(c,\hat{\lambda}) \;=\; 0 \;\neq\; V_d^{(k+2)}(c,\hat{\lambda}) \quad .$$

Fixing the last $d-k-2$ parameters and varying the first k around $(\hat{\lambda}_1, \ldots, \hat{\lambda}_k)$ as thus obtained we get an unfolding of the parabolic torus at $(p, q, \lambda) = (0, c, \hat{\lambda})$. Let us again write explicitly

$$\Lambda_j \quad := \quad \lambda_j - \hat{\lambda}_j \quad , \quad j = 1, \ldots, k \quad .$$

To apply Theorem 2.1 we scale with respect to $\delta := |V_d^{(k+2)}(c, \hat{\lambda})|$. Note that in the cusp case $d = 4$, $k = 1$ this amounts to $\delta = \sqrt{\hat{\lambda}_2}$. While this quantity "directly" only depends on $\hat{\lambda}_{k+2}, \hat{\lambda}_{k+3}, \ldots, \hat{\lambda}_{d-2}$, it also depends on $\hat{\lambda}_{k+1}$ through $c = c(\hat{\lambda}_{k+1}, \ldots, \hat{\lambda}_{d-2})$; and it is in particular possible to let $\delta \to 0$ by letting $\hat{\lambda}$ tend to the stratum of co–dimension $k+1$, changing only $\hat{\lambda}_{k+1}$. Then the scaling

$$\begin{aligned}
y &= \delta^{-k-\frac{9}{2}} Y \\
p &= \delta^{-k-\frac{5}{2}} P \\
q &= \delta^{-2} Q \\
\lambda_1 &= \delta^{-2k-3} \Lambda_1 \\
\lambda_2 &= \delta^{-2k-1} \Lambda_2 \\
&\vdots \\
\lambda_k &= \delta^{-5} \Lambda_k \\
\Omega &= \delta^{-k-\frac{1}{2}} \omega \\
T &= \delta^{k+\frac{1}{2}} t
\end{aligned}$$

together with a division of N by δ^{2k+5} allows to factor a common δ from the higher order terms and thus treat these as a perturbation ; **q.e.d.**

5 Reduction of parameters

The normal form (1.1) depends on the parameters ω_i , $i = 1, \ldots, n$ and λ_j , $j = 1, \ldots, d-2$. Below we discuss why this seemingly special situation is in fact very general. Then we show how the (Cantor)–strata of lower dimensions fit within the (Cantor)–strata of higher dimensions.

5.1 Replacing parameters by phase space variables

Given a (single) Hamiltonian system with (unperturbed) Hamiltonian function H_0 , the Kolmogorov–type nondegeneracy condition

$$\det(D_y^2 H_0) \;\neq\; 0 \tag{5.9}$$

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 $y \in \mathsf{Y}$. We are interested in furthermore also replacing the multiparameter λ by y in (1.1). In general one can gain one more parameter by time reparametrisation, see § 7c of [Broer,Huitema,Takens;90], or [Broer,Huitema;91]. In this way the phase space variables $y \in \mathsf{Y}$ take care of all necessary parameters in the case $d = 3$ of the centre–saddle bifurcation, see [Hanßmann;98].

Using Diophantine approximation of dependent quantities, cf. § 2.5 of [Broer,Huitema, Sevryuk;96] and references therein, we generally can obtain that y compensates for all parameters (λ, ω). Given certain nondegeneracy conditions, see Lemma 2.13 of [Broer,Huitema,Sevryuk;96], this implies that in a given (perturbed) system with Hamiltonian function

$$H \;=\; H_0(y,p,q) \;+\; \varepsilon H_1(x,y,p,q) \quad , \tag{5.10}$$

i.e. without any external parameters, all phenomena of the present paper may occur.

Let us explain how to recover the normal form (1.1) from (5.10). Assume that the Taylor series of H_0 near $(p,q) = (0,0)$ has the form

$$\begin{aligned} H_0(y,p,q) \;=\; & (\omega(y) \mid y) \;+\; \frac{a(y)}{2}p^2 \;+\; \frac{b(y)}{d!}q^d \\ & + \sum_{j=1}^{d-2} \frac{c_j(y)}{j!}q^j \;+\; \textit{higher order terms} \quad . \end{aligned} \tag{5.11}$$

Here the term with q^{d-1} has been cancelled by a suitable translation $q \mapsto q + \tau(y)$; to obtain a symplectic transformation the angular variable has to be simultaneously translated according to $x \mapsto x + D\tau(y)\,p$. As always the frequency vector ω is given by $\omega(y) = D_y H_0(y,0,0)$. While $a(y)$ and $b(y)$ are bounded from below for $y \in \mathsf{Y}$, the most degenerate bifurcation occurs at $y = 0$ as $c(0) \in \mathbb{R}^{d-2}$ vanishes.

The number of parameters λ_j depends on the degeneracy d as the universal unfolding of the singularity q^d requires $d-2$ parameters, see [Bröcker,Lander;75]. In order that the corresponding bifurcation diagram be faithfully represented by means of the y_i , we require the map

$$\begin{array}{rccc} c : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^{d-2} \\ & y & \mapsto & c(y) \end{array}$$

in (5.11) to be a submersion. This implies $n \geq d-2$, 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 parabolic tori with dominant terms p^2 and q^{k+2} are expected to form subfamilies of co–dimension $k \in \mathbb{N}$. Indeed, unless forced by additional symmetries, subfamilies of dimension $< n-k$ can avoid these by passing to another integrable system that is a slight perturbation of the former. In this way, bifurcating tori of degeneracy $d > n+2$ are not encountered.

The nondegeneracy condition (5.9) expresses that the partial derivatives $\frac{\partial^{|\ell|}\omega}{\partial y^\ell}$ span $\mathbb{R}^n$, where $\ell \in \mathbb{N}^n$ with $|\ell| = 1$ ($|\ell| \leq 1$ in case of iso–energetic nondegeneracy, cf. [Broer, Huitema;91]). This allows to control the frequency (the frequency ratio) of the perturbed tori. In the present case of normally parabolic tori the crucial map is

$$\begin{array}{rccc} \nu : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^{d-2} \times \mathbb{R}^n \\ & y & \mapsto & (c(y), \omega(y)) \end{array} \quad , \tag{5.12}$$

and since we want the first factor c to be a submersion, we miss $d-2$ directions if we restrict to $|\ell| = 1$. Therefore, we relax the control on the perturbed tori and only aim to show that a perturbed system does have Cantor families of invariant n–tori (according to the hierarchy of co–dimensions of normally parabolic tori addressed in Theorem 2.2) — without trying to further connect the perturbed tori to the unperturbed ones. The proper requirement is that the collection of $\binom{n+L}{n}$ vectors

$$\frac{\partial^{|\ell|}\nu}{\partial y^\ell} \quad , \quad |\ell| \leq L$$

spans $\mathbb{R}^{d-2} \times \mathbb{R}^n$. This implies that the image of ν is "sufficiently curved" and does not lie in any linear hyperplane in $\mathbb{R}^{n+d-2}$ passing through the origin. Since L enters the Diophantine condition (1.2) through $\tau > nL - 1$ it is preferable to keep L as small as possible, *i.e.* to work with $L = 2$.

In this way we use $c : \mathbb{R}^n \longrightarrow \mathbb{R}^{d-2}$ to pull back the bifurcation diagram to the space of actions. The remaining first derivatives together with the higher derivatives of $\nu = (c, \omega)$ then ensure that most frequencies perturbed from the $\omega(y)$ are Diophantine and, hence, yield invariant tori in the perturbed system. Note, however, that for $d = n+2$ the bifurcating torus is isolated and may disappear in a resonance gap.

5.2 Hierarchy of density strata

In the setting of Theorem 2.1, *i.e.* with ω and λ treated as (external) parameters, the most degenerate normally parabolic tori are parametrised by $\mathcal{O}'_\gamma$; recall that the order of degeneracy is $d-2$. This set has locally the product structure $\mathbb{R} \times$ *Cantor dust*. Indeed, if ω satisfies (1.2), then this holds true for $\beta\omega$, $\beta \geq 1$ as well. Accordingly, the Cantor strata C_k, $k = 1, \ldots, d-2$ of Theorem 2.2 locally have the procuct structure

$$\mathbb{R}^{d-k-1} \times \text{Cantor dust} \quad . \tag{5.13}$$

We colloquially refer to such sets as Cantor sets and reserve the name Cantor dust to those Cantor sets that indeed are totally discontinuous.

While these all parametrise normally parabolic tori, the local structure of C_0 depends on whether normally hyperbolic or normally elliptic tori are parametrised. In the former case one may simply put $k = 0$ in (5.13). In the latter case the normal frequency Ω enters the Diophantine condition. In the system defined by (1.1), normally elliptic invariant n–tori are given by $y = 0 = p$ and $(\hat{q}, \hat{\lambda})$ satisfying

$$\frac{b(\omega)}{(d-1)!}\,\hat{q}^{d-1} \;+\; \sum_{j=0}^{d-3} \frac{\hat{\lambda}_{j+1}}{j!}\,\hat{q}^{j} \;=\; 0 \qquad \text{and}$$

$$\operatorname{sgn}(a(\omega)) \left(\frac{b(\omega)}{(d-2)!}\,\hat{q}^{d-2} \;+\; \sum_{j=0}^{d-4} \frac{\hat{\lambda}_{j+2}}{j!}\,\hat{q}^{j} \right) \;>\; 0 \quad ,$$

whence

$$\Omega \;=\; \sqrt{a(\omega)\Big(\tfrac{b(\omega)}{(d-2)!}\,\hat{q}^{d-2} \;+\; \sum_{j=0}^{d-4} \tfrac{\hat{\lambda}_{j+2}}{j!}\,\hat{q}^{j}\Big)}$$

is the normal frequency. To obtain persistence, the Diophantine condition (1.2) has to be replaced by

$$|(\omega \mid k) + \Omega\ell\,| \;\geq\; \frac{\gamma}{|k|^{\tau}} \qquad \forall_{k \in \mathbb{Z}^n \setminus \{0\}} \quad \forall_{\ell \in \{0, \pm 1, \pm 2\}} \quad ,$$

see [Pöschel;89], [Broer,Huitema,Takens;90] or [Broer,Huitema,Sevryuk;96]. Consequently, one has to put $k = 1$ in (5.13) to obtain the local product structure of the Cantor set parametrising the persistent normally elliptic tori. We remark that the “additional” Diophantine conditions, with $\ell \neq 0$, exclude smaller and smaller subsets as $\Omega \to 0$. This implies that normally parabolic tori consist of Lebesgue density points of normally elliptic tori.

When using the mapping $\nu : \mathsf{Y} \longrightarrow \Lambda \times \mathcal{O}$ defined in (5.12) to let the phase space variables y account for the parameters (λ, ω), the strata C_k of Theorem 2.2 get replaced by the Cantor sets $\nu^{-1}(C_k)$ of respective (Hausdorff)–dimension $n-k$. Therefore, the local structure of the Cantor set parametrising normally hyperbolic tori becomes $\mathbb{R} \times$ Cantor dust, while the Cantor set parametrising normally elliptic tori and all strata parametrising normally parabolic tori become Cantor dust. Note that “higher strata” $\nu^{-1}(C_{k+1})$ consist of Lebesgue density points of “lower strata” $\nu^{-1}(C_k)$. Indeed, this is inherited from the Cantor stratification of $\Lambda \times \mathcal{O}$ into $\bigcup C_k$.

6 Proof of Theorem 2.1

In this section we give a detailed proof of Theorem 2.1, following the quite universal set–up of *e.g.* [Moser;67], [Pöschel;89], [Broer,Huitema,Takens;90]. Our aim is to obtain a co–ordinate transformation Φ that pushes the perturbation P of the normal form N into higher order terms and thus allows to recover the bifurcation scenario imposed by N in the perturbed system. We only expect those invariant tori to survive that satisfy a strong form of quasi–periodicity and concentrate on Diophantine frequency vectors. Therefore, we construct Φ as a limit of a sequence of transformations $(\Phi_\nu)_{\nu \in \mathbb{N}}$ defined on shrinking open neighbourhoods of our set $\mathcal{O}'_\gamma$ of Diophantine frequency vectors.

In fact we first fix the Diophantine constant $\gamma = 1$ when proving Theorem 2.1. This allows for a more transparent argumentation where the sizes β_ν of the shrinking neighbourhoods of $\mathcal{O}'_1$ are effectively decoupled from the Diophantine constant $\gamma > 0$. A simple scaling argument given in Section 6.3 allows to extend the result thus obtained to the $\mathcal{O}'_\gamma$ of Theorem 2.1.

To ensure that the limit is Whitney–C^∞–smooth in ω we work on domains $\mathcal{D}(r_\nu, s_\nu, \beta_\nu)$ that shrink geometrically in the x– and ω–directions. Then an exponentially fast decreasing sequence $(\varepsilon_\nu)_{\nu \in \mathbb{N}}$ that controls at the νth step the (transformed) perturbation P_ν allows to use the Inverse Approximation Lemma of [Zehnder;75] for the desired Whitney–C^∞–smoothness. The necessary control of P_ν is in turn obtained by letting shrink $\mathcal{D}(r_\nu, s_\nu, \beta_\nu)$ exponentially, described by $s_\nu = \varepsilon_\nu^{\frac{1}{2d+\sigma}}$ with $\sigma \in \,]0,1[$, in the (y,p,q,λ)–directions. The limit

$$\bigcap_\nu \mathcal{D}(r_\nu, s_\nu, \beta_\nu)$$

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

An additional complication is that a mere polynomial truncation of the Φ_ν would cease to preserve the symplectic structure. For this reason we introduce generating functions S_ν of the Φ_ν , the polynomial truncations $\widetilde{S}_\nu$ of which generate symplecto–morphisms as well. The limit $\widetilde{S}_\infty$ of these then generates the desired Φ_∞ .

We define the νth transformation

$$\Phi_\nu \quad = \quad \Psi_0 \circ \Psi_1 \circ \ldots \circ \Psi_{\nu-1}$$

where

$$\Psi_{\nu-1} : \;\; \mathcal{D}(r_\nu, s_\nu, \beta_\nu) \;\; \longrightarrow \;\; \mathcal{D}(r_{\nu-1}, s_{\nu-1}, \beta_{\nu-1}) \quad .$$

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

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

This is in sharp contrast with the (otherwise similar) approach in [Braaksma,Broer;87], [Braaksma,Broer,Huitema;90], [Hanßmann;98] where these two problems were addressed simultaneously. Thus, $\varphi_{\nu-1}$ is the solution of the linearized (or "1–bite") small denominator problem and $\phi_{\nu-1}$ uses explicit transformations from singularity theory to put the (now x–independent) lower order terms again into normal form (1.1).

In this way the method of proof follows the standard KAM recipe, inspired by the iterative schemes of *e.g.* [Moser;67] and [Pöschel;89]. At the νth step we have a perturbation form $H_\nu = N_\nu + P_\nu$ with N_ν in normal form (1.1) and P_ν sufficiently small. Although not of dynamical importance, one should in fact introduce a constant term $e_\nu = e_\nu(\lambda, \omega)$ to compensate for constant terms introduced by P_ν . Since it is *a priori* clear that these do not add up to ∞ we do not burden our presentation with these constant terms and suppress them on every stage.

In the limit we obtain $H_\infty = N_\infty + P_\infty$, where P_∞ has no longer any influence on the tori at $\lambda = 0$ and their normal behaviour, thus yielding the desired persistence result.

One of the basic tools of the KAM method is Cauchy's inequality. Here the analyticity of our Hamiltonian functions comes into play, allowing to control derivatives by the supremum norm on complex domains. Therefore, we extend the Hamiltonian H into a complex domain

$$\mathcal{D}(r_0, s_0, \beta_0) \quad = \quad D(r_0, s_0) \times U_{\beta_0}(\mathcal{O})$$

where

$$D(r_0, s_0) \quad = \quad \Big\{ \, (x, y, p, q, \lambda) \;\; | \;\; |\mathrm{Im}\, x| \le r_0 \, , \; |y| \le s_0^{2d} \, , \; |p| \le s_0^{d} \, , \; |q| \le s_0^{2} \, , \; |\lambda_j| \le s_0^{2d-2j} \, \Big\}$$

and the second factor is a complex β_0–neighbourhood

$$U_{\beta_0}(\mathcal{O}) \quad = \quad \{ \; w \in \mathbb{C}^n \;\; | \;\; \exists\, \omega \in \mathcal{O} \;\; |w - \omega| < \beta_0 \; \}$$

of the set $\mathcal{O} \subseteq \mathbb{R}^n$ of all frequency vectors, and assume that $|P| \le \varepsilon_0$ in $\mathcal{D}(r_0, s_0, \beta_0)$. Relating ε_ν to s_ν through $\varepsilon_\nu = s_\nu^{2d+\sigma}$ we require the perturbation part P_ν of H_ν to satisfy

$$|P_\nu| \quad \le \quad \varepsilon_\nu$$

in the shrinking domain

$$\mathcal{D}(r_\nu, s_\nu, \beta_\nu) \quad = \quad D(r_\nu, s_\nu) \times U_{\beta_\nu}(\mathcal{O}'_1)$$

where

$$D(r_\nu, s_\nu) \quad = \quad \Big\{ \, (x, y, p, q, \lambda) \;\; | \;\; |\mathrm{Im}\, x| \le r_\nu \, , \; |y| \le s_\nu^{2d} \, , \; |p| \le s_\nu^{d} \, , \; |q| \le s_\nu^{2} \, , \; |\lambda_j| \le s_\nu^{2d-2j} \, \Big\}$$

and the β_ν–neighbourhood is only taken of the Diophantine frequency vectors in $\mathcal{O}'_1$. Here s_ν (and ε_ν) converge exponentially to 0 , while r_ν converges geometrically to $\frac{1}{2}r_0$ and β_ν is related to r_ν , converging (geometrically) to 0 . For the precise definition of all these sequences see (6.17) below. We motivate the choice of the weights $2d, d, 2$ and $2d-2j$ in Definition 6.2, see Section 6.1.2.

6.1 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_ν into a "new" Hamiltonian $H_{\nu+1}$ that differs "less" from the normal form $N_{\nu+1}$. To this end we rewrite H_ν as

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

where R_ν is a conveniently chosen higher order truncation of P_ν , see (6.22). We show below how the Newton–like accelerated convergence implies that $|P_\nu - R_\nu|$ is less than $|P_\nu|^\kappa_{\mathcal{D}}$, $\kappa > 1$, on the smaller domain $\mathcal{D}(r_{\nu+1}, s_{\nu+1}, \beta_{\nu+1})$.

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

$$\begin{aligned} H_\nu \circ \varphi_{F_\nu} \;&=\; (N_\nu + R_\nu) \circ \varphi_{F_\nu} \;+\; (P_\nu - R_\nu) \circ \varphi_{F_\nu} \\ &=\; N_\nu \;+\; R_\nu \;+\; \{N_\nu, F_\nu\} \;+\; \{R_\nu, F_\nu\} \\ &\quad +\; \int_0^1 (1-t)\{\{N_\nu + R_\nu, F_\nu\}, F_\nu\} \circ \varphi_t^{F_\nu}\, \mathrm{d}t \;+\; (P_\nu - R_\nu) \circ \varphi_{F_\nu} \\ &=\; N_\nu \;+\; R_\nu \;+\; \{N_\nu, F_\nu\} \;+\; \bar{P}_\nu \quad . \end{aligned} \tag{6.14}$$

The philosophy of the KAM method is to find a special F_ν defined in a shrunken domain which makes the new perturbation $\bar{P}_\nu$ in (6.14) much smaller and $N_\nu + R_\nu + \{N_\nu, F_\nu\}$ a new normal form $N_{\nu+1}$. In the present context by this notion we not only mean 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_ν is thus obtained by solving a linear partial differential equation, the so–called homological equation

$$N_\nu \;+\; R_\nu \;+\; \{N_\nu, F_\nu\} \;=\; N_{\nu+1} \tag{6.15}$$

where

$$\{N, F\} \;=\; \frac{\partial N}{\partial x}\frac{\partial F}{\partial y} \;-\; \frac{\partial N}{\partial y}\frac{\partial F}{\partial x} \;+\; \frac{\partial N}{\partial q}\frac{\partial F}{\partial p} \;-\; \frac{\partial N}{\partial p}\frac{\partial F}{\partial q} \quad .$$

In the present bifurcating case, since N_ν contains higher order terms in q, equation (6.15) cannot be solved completely. Note that the purpose of solving (6.15) is to find a function F_ν so that (6.14) 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 (6.15), we solve the "intermediate homological equation"
$$N_\nu \;+\; R_\nu \;+\; \{N_\nu, F_\nu\} \;=\; \bar{N}_\nu \tag{6.16}$$
up to some order and treat the higher order terms (which are smaller) as a part of the new perturbation. The "intermediate" $\bar{N}_\nu$, already independent of x , but not yet normalized in p and q , is defined later in (6.26). The solution of (6.16) leads to small denominators, whence the Diophantine conditions (1.2) are needed. For the νth iteration step we only use finitely many of these conditions, up to some "ultraviolet" cut–off for the order K_ν of the Fourier truncation (6.22) defined in (6.17) below.

2. Then we look for a symplectic change of variables which transforms $\bar{N}_\nu$ into normal form (1.1). This passage from $\bar{N}_\nu$ to $N_{\nu+1}$ does not involve small denominators, but requires methods from singularity theory instead.

6.1.1 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 following sequences :

$$\begin{aligned}
\rho_\nu &= \frac{\rho_{\nu-1}}{4} = \frac{1}{4^\nu}\cdot\frac{3r_0}{32}\\
r_\nu &= r_{\nu-1} - 4\rho_{\nu-1} = \frac{r_0}{2}\,(1+\frac{1}{4^\nu})\\
\beta_\nu &= \rho_\nu^{2\tau+2}\\
K_\nu &= [\beta_\nu^{\frac{-1}{\tau+1}}] = [\rho_\nu^{-2}]\\
s_\nu &= s_{\nu-1}^{\frac{\kappa}{2d+\sigma}}\, s_{\nu-1} = s_0^{(1+\frac{\kappa}{2d+\sigma})^\nu}\\
\varepsilon_\nu &= s_\nu^{2d+\sigma}
\end{aligned}\tag{6.17}$$

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

$$r_0 \;\le\; c \quad , \qquad s_0 \;\le\; c \quad , \qquad s_0^\varsigma \;\le\; c\,r_0$$

with constants $c > 0$ and exponents $\varsigma > 0$. The only exception is Lemma 6.3 below where an inequality

$$r_0 \quad < \quad \frac{1}{c - \varsigma \ln(s_0)} \tag{6.18}$$

occurs. Since $s_0^{\varsigma} \ln(s_0) \stackrel{s_0 \to 0}{\longrightarrow} 0$ for all $\varsigma > 0$ 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

$$H_\nu \quad = \quad N_\nu \; + \; P_\nu \tag{6.19}$$

with

$$N_\nu \quad = \quad (\omega \mid y) \; + \; \frac{a_\nu}{2} p^2 \; + \; \frac{b_\nu}{d!} q^d \; + \; \sum_{j=1}^{d-2} \frac{\lambda_j}{j!} q^j \tag{6.20}$$

and defined in

$$\mathcal{D}_\nu \quad := \quad \mathcal{D}(r_\nu, s_\nu, \beta_\nu) \quad = \quad D(r_\nu, s_\nu) \times U_{\beta_\nu}(\mathcal{O}_1') \quad .$$

We also use the abbreviation

$$U_\nu \quad := \quad U_{\beta_\nu}(\mathcal{O}_1')$$

for the β_ν–neighbourhood in the second factor.

Lemma 6.1 *Suppose that* $H_\nu = N_\nu + P_\nu$ *satisfies* (6.20) *in* $\mathcal{D}_\nu$ *and that* P_ν *can be estimated by*

$$|P_\nu|_{\mathcal{D}_\nu} \quad \leq \quad \varepsilon_\nu \quad . \tag{6.21}$$

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

$$\Psi_\nu : \mathcal{D}_{\nu+1} \longrightarrow \mathcal{D}_\nu$$

such that $H_{\nu+1} = H_\nu \circ \Psi_\nu$ *, defined on* $\mathcal{D}_{\nu+1}$ *, has the form*

$$H_{\nu+1} \quad = \quad N_{\nu+1} \; + \; P_{\nu+1} \quad ,$$

satisfying

$$\begin{aligned} |P_{\nu+1}|_{\mathcal{D}_{\nu+1}} \quad &\leq \quad \varepsilon_{\nu+1} \\ |a_{\nu+1} - a_\nu|_{U_{\nu+1}} \quad &\leq \quad s_\nu \\ |b_{\nu+1} - b_\nu|_{U_{\nu+1}} \quad &\leq \quad s_\nu \quad . \end{aligned}$$

Moreover,

$$\left| \frac{\partial^{|l|+i+j+|h|} P_{\nu+1}}{\partial y^l \partial p^i \partial q^j \partial \lambda^h} \right|_{\mathcal{D}_{\nu+1}} \quad \leq \quad s_{\nu+1}^{2d+\sigma-m}$$

where $m := 2d|l| + di + 2j + (2d-2)h_1 + \ldots + 4h_{d-2} \leq 2d$ *.*

Remark : Compared to the perturbation, the coefficient functions a_ν and b_ν are of order one, *i.e.* they satisfy bounds as formulated in Theorem 2.1. The estimates by s_ν on the differences $|a_{\nu+1} - a_\nu|$, $|b_{\nu+1} - b_\nu|$ imply that the same is true for $a_{\nu+1}$ and $b_{\nu+1}$ as well, and also for the (existing) limit functions a_∞ and b_∞.

6.1.2 The "intermediate" homological equation

To prove Lemma 6.1 we describe a single iteration step in detail. Therefore, we drop the index ν and use the so–called "+"–notation, replacing occurrences of the index $\nu + 1$ by an index $+$. As said earlier, we look for a symplectic co–ordinate transformation such that the transformed Hamiltonian function satisfies (6.19)—(6.21) with s_+, ε_+ and so on. This also emphasizes that the constants in our estimates have to be independent of ν. 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 $A \leq. B$ for $A \leq c \cdot B$. Adapted to the normal form (1.1) we introduce the concept of quasi–homogeneous polynomials with weight, cf. [Arnol'd,Vasil'ev,Goryunov,Lyashko;88].

Definition 6.2 *A polynomial $F(y, p, q, \lambda_1, \ldots, \lambda_k)$ is said to be quasi–homogeneous of order m with weight $(\alpha_y, \alpha_p, \alpha_q; \alpha_1, \ldots, \alpha_k)$ if*

$$F(e^{\alpha_y \varsigma} y, e^{\alpha_p \varsigma} p, e^{\alpha_q \varsigma} q, e^{\alpha_1 \varsigma} \lambda_1, \ldots, e^{\alpha_k \varsigma} \lambda_k) \;\; \equiv \;\; e^{m\varsigma} F(y, p, q, \lambda_1, \ldots, \lambda_k) \quad ,$$

where m is a positive integer.

Remark : In this way

$$N \;\; = \;\; (\omega \mid y) \; + \; \frac{a}{2} p^2 \; + \; \frac{b}{d!} q^d \; + \; \sum_{j=1}^{d-2} \frac{\lambda_j}{j!} q^j$$

is a $(2d)$th order quasi–homogeneous polynomial with weight $(2d, d, 2; 2d-2, \ldots, 4)$. This weight in turn induces the weighted order

$$\|(l, i, j, h)\| \;\; := \;\; 2d|l| \; + \; di \; + \; 2j \; + \; (2d-2)h_1 \; + \; \ldots \; + \; 4h_{d-2}$$

on indices

$$l = (l_1, \ldots, l_n) \in \mathbb{N}_0^n \;\; , \;\; i, j \in \mathbb{N}_0 \;\; , \;\; h = (h_1, \ldots, h_{d-2}) \in \mathbb{N}_0^{d-2} \quad .$$

On the ring of all (=formal) power series in y, p, q and λ this defines in the terminology of [Bourbaki;85] a *gradation*

$$\mathcal{A}_m \;\; := \;\; \left\{ \; F \in \mathbb{C}[y, p, q, \lambda] \;\; \Big| \; \substack{F \text{ quasi–homogeneous of order } m \\ \text{with weight } (2d, d, 2; 2d-2, \ldots, 4)} \right\}$$

together with the *filtration* $\mathcal{F}_n := \prod_{m\geq n} \mathcal{A}_m$. With this terminology we may write

$$F \;=\; G \quad (\operatorname{mod} \mathcal{F}_{2d})$$

if all monomials up to weighted order $2d$ in the power series F and G have equal coefficients.

We expand the perturbation P into a Fourier–Taylor series

$$P(x,y,q,p,\lambda) \;=\; \sum_{m=0}^{\infty} \sum_{\|(l,i,j,h)\|=m} \sum_{k\in\mathbb{Z}^n} P_{klijh}\, \mathrm{e}^{\mathrm{i}(k|x)} y^l p^i q^j \lambda^h$$

and define the truncation

$$R \;=\; \sum_{|k|\leq K} P_k\, \mathrm{e}^{\mathrm{i}(k|x)} \tag{6.22}$$

of P with

$$P_k(y,p,q,\lambda) \;=\; \sum_{m\leq 2d} P_{km} \;=\; \sum_{m\leq 2d} \Big(\sum_{\|(l,i,j,h)\|=m} P_{klijh}\, y^l p^i q^j \lambda^h \Big) \quad . \tag{6.23}$$

Here $|k| \;=\; K$, with $K \;=\; [\beta^{\frac{-1}{\tau+1}}]$, is the maximal order $|k| \;=\; k_1 + \ldots + k_n$ 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 (6.16)) and on the remaining term $P - R$ which will be included in the new (and smaller!) perturbation.

Lemma 6.3 *Under the conditions of Lemma 6.1 the inequality*

$$|R|_{\mathcal{D}(r-\rho,\frac{1}{2}s,\beta)} \;\leq\!\cdot\; \varepsilon \tag{6.24}$$

holds. Moreover, on a smaller domain we have

$$|P-R|_{\mathcal{D}(r-\rho,\alpha s,\beta)} \;\leq\!\cdot\; \alpha^{1-\sigma} s^{\kappa} \varepsilon \tag{6.25}$$

where $\alpha = 9s^{\frac{\kappa}{2d+\sigma}}$.

Proof. We only prove (6.25), since (6.24) can be obtained from the following proof by taking $\alpha = \frac{1}{2}$. Note that

$$P - R \;=\; \sum_{\substack{m\leq 2d\\ |k|>K}} P_{km} \mathrm{e}^{\mathrm{i}(k|x)} \;+\; \sum_{\substack{m>2d\\ k\in\mathbb{Z}^n}} P_{km} \mathrm{e}^{\mathrm{i}(k|x)} \quad .$$

To estimate the first term we use the Paley–Wiener estimate and an upper bound of the number of vectors $k \in \mathbb{Z}^n$ with $|k| = \mu$ to obtain

$$\begin{aligned} \Big|\sum_{\substack{m\leq 2d\\ |k|>K}} P_{km} \mathrm{e}^{\mathrm{i}(k|x)}\Big| \;&\leq\; \sum_{|k|>K} |P|_{\mathcal{D}}\, \mathrm{e}^{-|k|r} \mathrm{e}^{|k|(r-\rho)} \;\leq\; \varepsilon \sum_{\mu=K+1}^{\infty} \mu^n \mathrm{e}^{-\mu\rho} \\ &\leq\; \varepsilon \int_K^{\infty} x^n \mathrm{e}^{-\rho x}\, \mathrm{d}x \;\leq\!\cdot\; K^n\, \mathrm{e}^{-K\rho}\, \varepsilon \quad . \end{aligned}$$

Using $K=[\beta^{\frac{-1}{\tau+1}}]$ and $\beta=\rho^{2\tau+2}$ we continue

$$K^n\,\mathrm{e}^{-K\rho}\,\varepsilon \quad \le. \quad \rho^{-2n}\,\mathrm{e}^{-\frac{1}{2\rho}}\,\varepsilon \quad \le. \quad s\varepsilon$$

where the last inequality is given by Lemma 5.10 of [Braaksma,Broer;87]. The condition of that lemma leads to the inequality (6.18). For the second term we use that we may shrink the domain in the y,p,q and λ direction and get

$$\begin{aligned} |\sum_{\substack{m>2d\\ k\in\mathbb{Z}^n}} P_{km}\mathrm{e}^{\mathrm{i}(k|x)}| \quad &\le \quad \left| \int \frac{\partial^{|l|+i+j+|h|}}{\partial y^l\partial p^i\partial q^j\partial\lambda^h} \sum_{\substack{m>2d\\ k\in\mathbb{Z}^n}} P_{km}\mathrm{e}^{\mathrm{i}(k|x)}\,\mathrm{d}y^l\mathrm{d}p^i\mathrm{d}q^j\mathrm{d}\lambda^h \right| \\ &\le. \quad \frac{(\alpha s)^{2d+1}}{s^{2d+1}} |\sum_{m=0}^{\infty}\sum_{|k|\in\mathbb{Z}^n} P_{km}\mathrm{e}^{\mathrm{i}(k|x)}| \quad \le. \quad \alpha^{2d+1}\varepsilon \end{aligned}$$

in $\mathcal{D}(r-\rho,\alpha s,\beta)$, where $\int$ stands for

$$\underbrace{\int_0^{y_1}\cdots\int_0^{y_1}}_{l_1}\cdots\underbrace{\int_0^{y_n}\cdots\int_0^{y_n}}_{l_n}\underbrace{\int_0^{p}\cdots\int_0^{p}}_{i}\underbrace{\int_0^{q}\cdots\int_0^{q}}_{j}\underbrace{\int_0^{\lambda_1}\cdots\int_0^{\lambda_1}}_{h_1}\cdots\underbrace{\int_0^{\lambda_{d-2}}\cdots\int_0^{\lambda_{d-2}}}_{h_{d-2}}$$

with $\|(l,i,j,h)\|=2d+1$; **q.e.d.**

Our next goal is to solve the "intermediate" homological equation (6.16). To this end we add the average of (6.22) to N , *i.e.* we let

$$\bar{N} \;=\; N \;+\; P_0(y,p,q,\lambda) \quad , \tag{6.26}$$

where

$$P_0(y,p,q,\lambda) \;=\; \sum_{j=1}^{d} P_j(\lambda)q^j \;+\; \sum_{j=0}^{[\frac{d}{2}]} Q_j(\lambda)q^j p \;+\; P_{00200}p^2 \;+\; (P_{01000}\mid y)$$

is given by (6.23) with

$$\begin{aligned} P_j(\lambda) \;&=\; \sum_{2(d-1)h_1+\ldots+4h_{d-2}\le 2d-2j} P_{000jh}\lambda^h \\ Q_j(\lambda) \;&=\; \sum_{2(d-1)h_1+\ldots+4h_{d-2}\le d-2j} P_{001jh}\lambda^h \quad . \end{aligned}$$

Here and below we completely suppress the ω–dependence, in particular the coefficients $P_{0lijh}=P_{0lijh}(\omega)$ are functions on $U=U_\beta(\mathcal{O}_1')$. Recall that we ignore the constant terms $P_{0000h}\lambda^h$. Since we cannot solve (6.16) completely, we let

$$\begin{aligned} F \;&=\; \sum_{0<|k|\le K} F_k\mathrm{e}^{\mathrm{i}(k|x)} \\ F_k \;&=\; \sum_{m\le 2d} F_{km} \;=\; \sum_{m\le 2d}\Big(\sum_{\|(l,i,j,h)\|=m} F_{klijh}\,y^lp^iq^j\lambda^h\Big) \end{aligned}$$

be the solution of

$$N \;+\; R \;+\; \{N,F\} \;=\; \bar{N} \quad (\mathrm{mod}\,\mathcal{F}_{2d}) \quad ,$$

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

$$\begin{aligned}
\mathrm{i}(k \mid \omega) F_{km} \;&=\; P_{km} \;+\; \left(\tfrac{b}{(d-1)!} q^{d-1} + \sum_{j=1}^{d-2} \tfrac{1}{(j-1)!} \lambda_j q^{j-1} \right) \frac{\partial F_{k,m+2-d}}{\partial p} \;-\; ap \frac{\partial F_{k,m+2-d}}{\partial q} \\
&=\; P_{km} \;+\; \{N_0, F_{k,m+2-d}\} \quad ,
\end{aligned}$$

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

$$F_{km} \;=\; \Delta P_{km} \;+\; \sum_{i=1}^{6} \Delta^{i+1} \underbrace{\{N_0, \ldots, \{N_0}_{i}, P_{k,m-i(d-2)} \underbrace{\} \cdots \}}_{i} \quad . \tag{6.27}$$

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

$$m - 6(d-2) \;\leq\; 2d - 6(d-2) \;=\; 12 - 4d \;\leq\; 0$$

the right hand side of (6.27) contains at most 7 terms.

Remark : Since we only solved the equation (6.16) up to the weighted order $2d$, the higher order terms $\{N_0, \sum_{\substack{0<|k|\leq K \\ d+2<m\leq 2d}} F_{km} \mathrm{e}^{\mathrm{i}(k|x)}\}$ have to be included in the new perturbation.

To estimate the nested Poisson brackets in (6.27) we work on the nested domains

$$\mathcal{D}^i \;=\; \mathcal{D}(r - \frac{3+i}{4}\rho, \frac{1}{1+i}s, \frac{\beta}{2}) \;\subset\; \mathcal{D}^1 \;=\; \mathcal{D}(r-\rho, \frac{1}{2}s, \frac{\beta}{2}) \quad , \quad i = 1, \ldots, 8$$

and later use the four domains

$$\mathcal{D}^{i\alpha} \;=\; \mathcal{D}(r - \frac{11+i}{4}\rho, 2^{1-i}\alpha s, \frac{7-i}{12}\beta) \;\subset\; \mathcal{D}^{\alpha} \;=\; \mathcal{D}(r - 3\rho, \alpha s, \frac{\beta}{2}) \quad , \quad i = 1, \ldots, 4$$

to define the normalizing co–ordinate transformation. For Poisson brackets with N_0 we have the inequality

$$|\{N_0, G\}|_{\mathcal{D}^i} \;=\; |\frac{\partial N_0}{\partial q}\frac{\partial G}{\partial p} \;-\; \frac{\partial N_0}{\partial p}\frac{\partial G}{\partial q}|_{\mathcal{D}^i} \;\leq\!\cdot\; s^{d-2}\, |G|_{\mathcal{D}^{i-1}} \quad .$$

Lemma 6.4 *Under the conditions of Lemma 6.1 we have*

$$|F|_{\mathcal{D}^7} \;\leq\!\cdot\; \varepsilon \quad .$$

Proof. Expanding

$$F = \sum_{0<|k|\le K} F_k \mathrm{e}^{\mathrm{i}(k|x)} = \sum_{\substack{0<|k|\le K\\ m\le 2d}} F_{km} \mathrm{e}^{\mathrm{i}(k|x)}$$

$$= \sum_{i=0}^{6} \Delta^{i+1} \underbrace{\{N_0, \ldots, \{N_0,}_{i} \sum_{\substack{0<|k|\le K\\ m\le 2d}} P_{k,m-i(d-2)}\, \mathrm{e}^{\mathrm{i}(k|x)} \underbrace{\}\cdots\}}_{i}$$

we obtain

$$|F|_{\mathcal{D}^7} \le \left| \sum_{\substack{0<|k|\le K\\ m\le 2d}} \sum_{i=0}^{6} \Delta^{i+1} \underbrace{\{N_0, \ldots, \{N_0,}_{i} P_{k,m-i(d-2)} \underbrace{\}\cdots\}}_{i} \mathrm{e}^{\mathrm{i}(k|x)} \right|_{\mathcal{D}^7}$$

$$\le\cdot \sum_{\substack{0<|k|\le K\\ m\le 2d}} \sum_{i=0}^{6} |\Delta|^{i+1} s^{(d-2)i} \left| P_{k,m-i(d-2)} \right|_{\mathcal{D}^{7-i}} \mathrm{e}^{|k|(r-\rho)}$$

$$\le\cdot \sum_{|k|\le K} \sum_{i=0}^{6} |(k\mid\omega)|^{-i-1} s^{(d-2)i} \left| P \right|_{\mathcal{D}} \mathrm{e}^{-|k|\rho}$$

where we again used the Paley–Wiener estimate. Given ω there exists $w \in \mathcal{O}'_1$ with $|\omega - w| \le \frac{1}{2}\beta$ and from $K = [\beta^{\frac{-1}{\tau+1}}]$ we obtain

$$|(k\mid\omega)| \ge |(k\mid w)| - |\omega - w| \cdot |k| \ge |k|^{-\tau} - \frac{1}{2}\beta\, K \ge |k|^{-\tau} - \frac{1}{2}K^{-\tau} \ge \frac{1}{2}|k|^{-\tau}$$

whence the inequality

$$\sum_{k\in\mathbf{Z}} |k|^{7\tau} \mathrm{e}^{-|k|\rho} \le\cdot \rho^{-8\tau}$$

allows to conclude the proof ; **q.e.d.**

Remark : Since $(\rho_\nu)_\nu$ decreases geometrically and $(s_\nu)_\nu$ decreases exponentially fast we also include factors $\frac{1}{\rho_\nu}$ in the "generic constant c", taking care that the total number of such factors remains finite (and independent of ν).

By the Cauchy estimate we have

$$\left|\frac{\partial^{|l|+i+j+|h|}F}{\partial y^l \partial p^i \partial q^j \partial \lambda^h}\right|_{\mathcal{D}^8} \le\cdot s^{-m}\varepsilon \tag{6.28}$$

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

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

$$\Uparrow D_\mu X_F \Uparrow_{\mathcal{D}} := \max_{|l|+i+j\le\mu} \left\{ \left|\frac{\partial^{|l|+i+j}G}{\partial y^l \partial p^i \partial q^j}\right|_{\mathcal{D}} \right\} \quad \text{for} \quad \mu \ge 1$$

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 cs^σ 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 $\frac{\partial^{|l|+i+j}}{\partial y^l \partial p^i \partial q^j} X_F$.

Lemma 6.5 *Under the conditions of Lemma 6.1 we have*

$$\|X_F\|_{\mathcal{D}^8} \leq. s^\sigma \quad , \qquad \Uparrow D_\mu X_F \Uparrow_{\mathcal{D}^8} \leq. s^\sigma \quad \forall_{\mu \geq 1} \quad .$$

Proof. This is an immediate consequence of (6.28) and $\varepsilon = s^{2d+\sigma}$. Indeed, the expressions $\frac{\partial^{|l|+i+j}}{\partial y^l \partial p^i \partial q^j} G$ are either partial derivatives of F of weighted order $\leq 2d$, or vanish identically ; **q.e.d.**

Hence, the flow φ_t^F of X_F satisfies $\|\varphi_t^F - \mathrm{id}\|_{\mathcal{D}^8} \leq c\,|t|\,s^\sigma$ as well, *i.e.* the first, second, third and fourth component of $\varphi_t^F - \mathrm{id}$ are bounded by $c|t|s^\sigma$, $c|t|\varepsilon$, $c|t|s^{-2}\varepsilon$ and $c|t|s^{d+\sigma}$, respectively. Therefore, the inequality $s^{\frac{2d(\sigma-\kappa)}{2d+\sigma}} \leq \frac{1}{2c}$ implies that, for $-1 \leq t \leq 1$, the flow φ_t^F not only maps $\mathcal{D}^8$ into $\mathcal{D}^7$, but also maps $\mathcal{D}^{2\alpha}$ into $\mathcal{D}^\alpha$. Here we slightly abuse notation in that the same symbol φ_t^F is used for the mapping acting as the identity in the fifth and sixth component. Furthermore, we have the following estimate for $\varphi_F = \varphi_{t=1}^F$, the time one map of the Hamiltonian flow φ_t^F . The norm for φ_F is defined by

$$\|\varphi_F\|_{C^{lij}(\mathcal{D})} \;=\; \max_{0 \leq t \leq 1} \left\| \frac{\partial^{|l|+i+j} \varphi_t^F}{\partial y^l \partial p^i \partial q^j} \right\|_{\mathcal{D}} \quad .$$

We also define

$$(A \bullet B)(t) \;:=\; A(t) \circ B(t)$$

for two mapping–valued mappings A and B .

Lemma 6.6 *For any given l, i, j there is a constant s_0 , depending only on n, τ and $|l| + i + j$, such that if $s \leq s_0$*

$$\|\varphi_F - \mathrm{id}\|_{C^{lij}(\mathcal{D}^{2\alpha})} \quad \leq. \quad s^\sigma \quad .$$

Proof. Note that φ_t^F satisfies the integral equation

$$\varphi_t^F \;=\; \mathrm{id} + \int_0^t X_F \circ \varphi_{\tilde{t}}^F \, \mathrm{d}\tilde{t}$$

from which we derive for the (total) derivatives of order $\mu \geq 2$

$$D^\mu \varphi_F \;=\; 0 \;+\; \int_0^1 \sum c_{\ell\iota}^\mu ((D^\ell X_F) \circ \varphi_t^F) \bullet (D^{\iota_1} \varphi_t^F, \ldots, D^{\iota_\ell} \varphi_t^F) \, \mathrm{d}t \tag{6.29}$$

where the sum is taken over $\ell = 1, \ldots, \mu$ and $\iota_1 \geq \ldots \geq \iota_\ell \geq 1$ with $\iota_1 + \ldots + \iota_\ell = \mu$. Partial derivatives on the left hand side that do not differentiate with respect to x involve on the right hand side only partial derivatives with at most one differentiation with respect to x . For the first derivative we have

$$\Uparrow D\varphi_F \Uparrow_{\mathcal{D}^{2\alpha}} \;\leq\; 1 + \Uparrow DX_F \Uparrow_{\mathcal{D}^{\alpha}} \cdot \Uparrow D\varphi_F \Uparrow_{\mathcal{D}^{2\alpha}} \quad .$$

When $|l|+i+j=1$ Lemma 6.5 yields

$$\left\| \frac{\partial^{|l|+i+j}\varphi_F}{\partial y^l \partial p^i \partial q^j} \right\|_{\mathcal{D}^{2\alpha}} \;\leq\; \Uparrow D\varphi_F \Uparrow_{\mathcal{D}^{2\alpha}} \;\leq\; \frac{1}{1-cs^\sigma} \;\leq\; 2 \quad ,$$

if s_0 is sufficiently small. This immediately implies the desired result for $|l|+i+j=1$. Inductively, assuming the proper bound on $\frac{\partial^{|l|+i+j}}{\partial y^l \partial p^i \partial q^j}\varphi_F$ holds for $|l|+i+j \leq \mu-1$, from (6.29) we have

$$\begin{aligned} \left\| \frac{\partial^\mu(\varphi_F - \mathrm{id})}{\partial y^l \partial p^i \partial q^j} \right\|_{\mathcal{D}^{2\alpha}} \;\leq\!\cdot\;\; & \Uparrow D_\mu X_F \Uparrow_{\mathcal{D}^{\alpha}} \Uparrow D\varphi_F \Uparrow^{\mu}_{\mathcal{D}^{2\alpha}} \\ & + \Uparrow D_{\mu-1}X_F \Uparrow_{\mathcal{D}^{\alpha}} \; s^\sigma + \Uparrow DX_F \Uparrow_{\mathcal{D}^{\alpha}} \left\|\partial^\mu \varphi_F\right\|_{\mathcal{D}^{2\alpha}} \quad . \end{aligned}$$

Here $\partial^\mu\varphi_F$ denote partial derivatives of order μ that do not differentiate with respect to x ; those that do have been included in the second term. It follows that

$$\left\| \frac{\partial^\mu(\varphi_F - \mathrm{id})}{\partial y^l \partial p^i \partial q^j} \right\|_{\mathcal{D}^{2\alpha}} \;\leq\!\cdot\;\; \frac{2^\mu + \tilde{c}}{1-cs^\sigma}\, s^\sigma$$

for sufficiently small s_0 ; **q.e.d.**

6.1.3 Transformation of $\bar{N}$ into normal form

So far we have solved the small divisor problem (6.16) to construct a symplectic change of variables φ_F that transforms away the x–dependence of the lower order terms entering in $\bar{N}$. The second part of the iteration step consists in finding a symplectic change of variables $\phi_1 \circ \phi_2$ which transforms $\bar{N}$ of (6.26) into the normal form (1.1) up to some small terms, *i.e.* $\bar{N} \circ \phi_1 \circ \phi_2 = N_+ + O(\varepsilon_+)$.

Since $\bar{N}$ 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 [Broer,Chow,Kim,Vegter;93,95] one can apply the machinery of (planar) singularity theory to solve normalization problems (like the passage from $\bar{N}$ 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 ϕ_1 and ϕ_2 in an explicit way.

First we use the shear transformation

$$\phi_1 \ : \ \begin{cases} q_1 \ = \ q \\ p_1 \ = \ p \ + \ (a+2P_{00200})^{-1} \sum_{j=0}^{[\frac{d}{2}]} Q_j(\lambda) q^j \end{cases} \tag{6.30}$$

to kill the crossing terms $\sum_{j=0}^{[\frac{d}{2}]} Q_j(\lambda) q^j p$ in $\bar{N}$ (see (6.26)). We obtain

$$\begin{aligned} \bar{N} \circ \phi_1 \ &= \ (\omega_+ \mid y) \ + \ \frac{a_+}{2} p_1^2 \ + \ \frac{b_+}{d!} q_1^d \ + \ P_{d-1}(\lambda) q_1^{d-1} \\ &\quad + \ \sum_{j=1}^{d-2} (\frac{\lambda_j}{j!} + P_j(\lambda)) q_1^j \ - \ \tfrac{1}{2a_+} (\sum_j Q_j(\lambda) q_1^j)^2 \\ &=: \ \tilde{N} \ - \ \tfrac{1}{2a_+} (\sum_j Q_j(\lambda) q_1^j)^2 \end{aligned} \tag{6.31}$$

with $\omega_+ = \omega + P_{01000}$, $a_+ = a + 2P_{00200}$ and $b_+ = b + d!\, P_{000d0}$.

Remark : The squared sum in (6.31) can be estimated by $(\frac{\varepsilon}{s^d})^2 = s^{2d+2\sigma}$, and can therefore be included in the new perturbation.

The remaining term $P_{d-1}(\lambda) q_1^{d-1}$ of order $d-1$ in the part $\tilde{N}$ of (6.31) can be dealt with by the standard translation

$$\phi_2 \ : \ \begin{cases} q_2 \ = \ q_1 \ + \ \frac{(d-1)!}{b_+} P_{d-1}(\lambda) \\ p_2 \ = \ p_1 \end{cases} \tag{6.32}$$

which is well–known from singularity theory (or algebra). We arrive at

$$\begin{aligned} \bar{N} \circ \phi_1 \circ \phi_2 \ &= \ (\omega_+ \mid y) \ + \ \frac{a_+}{2} p_2^2 \ + \ \frac{b_+}{d!} q_2^d \ + \ \sum_{j=1}^{d-2} \frac{\lambda_j^+}{j!} q_2^j \\ &\quad - \ \tfrac{1}{2a_+} (\sum_j Q_j(\lambda) q_1^j)^2 \circ \phi_2 \\ &= \ N_+ \ - \ \tfrac{1}{2a_+} (\sum_j Q_j(\lambda) q_1^j)^2 \circ \phi_2 \end{aligned}$$

with

$$\begin{aligned} \lambda_i^+ \ &= \ \lambda_i \ + \ i! P_i(\lambda) \ + \ \frac{(-1)^{d-i+1}(d-i-1)}{(d-i)!\, b_+^{d-i-1}} ((d-1)!\, P_{d-1}(\lambda))^{d-i} \\ &\quad + \ \sum_{j=i+1}^{d-2} \frac{(-1)^{j-i}}{(j-i)!\, b_+^{j-i}} ((d-1)!\, P_{d-1}(\lambda))^{j-i} (\lambda_j \ + \ j! P_j(\lambda)) \quad . \end{aligned}$$

Recall that we always ignore constant terms. Since the Jacobian of $\lambda \mapsto \lambda^+$ is nonsingular, we can replace λ by λ^+ in the next KAM–step. Thus we get the desired new normal form.

6.1.4 Estimates of the iteration step

We now compose our map $\Psi : \mathcal{D}_+ \longrightarrow \mathcal{D}$ using $\mathcal{D}_+ \subseteq \mathcal{D}^{4\alpha}$ and $\mathcal{D}^{\alpha} \subseteq \mathcal{D}$. We have already remarked that $\varphi_F : \mathcal{D}^{2\alpha} \longrightarrow \mathcal{D}^{\alpha}$. The inequalities

$$\begin{aligned} |P_{01000}|_U &\leq\cdot \quad \frac{\varepsilon}{s^{2d}} \quad < \quad \frac{\beta}{12} \\ |a_+^{-1} \sum_j Q_j(\lambda) q^j| &\leq\cdot \quad \frac{\varepsilon}{s^d} \quad < \quad \frac{(\alpha s)^d}{4} \end{aligned}$$

imply $\phi_1 : \mathcal{D}^{3\alpha} \longrightarrow \mathcal{D}^{2\alpha}$ where we have subsumed $\omega \mapsto \omega_+$ into this mapping. Similarly we subsume $\lambda \mapsto \lambda^+$ into ϕ_2 and obtain $\phi_2 : \mathcal{D}^{4\alpha} \longrightarrow \mathcal{D}^{3\alpha}$ from

$$\begin{aligned} |(d-1)!\, b_+^{-1} P_{d-1}(\lambda)| &\leq\cdot \quad s^3 \quad < \quad \frac{(\alpha s)^2}{8} \\ |i!\, P_i(\lambda)| &\leq\cdot \quad \frac{\varepsilon}{s^{2i}} \quad < \quad \frac{(\alpha s)^{2d-2i}}{8} \qquad . \end{aligned}$$

Together we have that

$$\Psi \quad = \quad \varphi_F \circ \phi_1 \circ \phi_2 \quad : \quad \mathcal{D}^{4\alpha} \quad \longrightarrow \quad \mathcal{D}^{\alpha} \qquad .$$

This defines the desired co–ordinate transformation for one iteration step. Similar to Lemma 6.6, we have the estimates for Ψ :

Lemma 6.7 *For any given l, i, j there is a constant s_0 , depending only on n, τ and $|l| + i + j$, such that if $s \leq s_0$*

$$\|\Psi - \mathrm{id}\|_{C^{lij}(\mathcal{D}^{4\alpha})} \quad \leq\cdot \quad s^{\sigma} \qquad .$$

Proof. Note that ϕ_1 and ϕ_2 are linear maps, *i.e.*

$$D^{\mu}\Psi \quad = \quad D^{\mu}\varphi_F \circ \phi_1 \circ \phi_2 \bullet \underbrace{(\phi_1 \circ \phi_2, \ldots, \phi_1 \circ \phi_2)}_{\mu \text{ entries}} \qquad .$$

The estimate follows by Lemma 6.6, the chain rule and the inductive principle **;** **q.e.d.**

The new perturbation is

$$\begin{aligned} P_+ \quad = \quad & \bar{P} \circ \phi_1 \circ \phi_2 \; + \; \{ N_0, \sum_{\substack{0<|k|\leq K \\ d+2<m\leq 2d}} F_{km} \mathrm{e}^{\mathrm{i}(k|x)} \} \circ \phi_1 \circ \phi_2 \\ & - \; \tfrac{1}{2a_+} (\textstyle\sum_j Q_j(\lambda) q^j)^2 \circ \phi_2 \quad , \end{aligned} \tag{6.33}$$

where

$$\bar{P} \;=\; \{R,F\} \;+\; \int_0^1 (1-t)\{\{N+R,F\},F\}\circ\varphi_t^{F}\,\mathrm{d}t \;+\; (P-R)\circ\varphi_{F}$$

was defined in (6.14). We have to estimate all terms by $\varepsilon_+ = s^\kappa\varepsilon$. Using the Cauchy inequality we immediately have

$$\begin{aligned} |\{R,F\}\circ\phi_1\circ\phi_2|_{\mathcal{D}^{4\alpha}} &\leq |\{R,F\}|_{\mathcal{D}^{2\alpha}} \\ &\leq\!\cdot\; s^{-2d}\,|R|_{\mathcal{D}^{7}}\,|F|_{\mathcal{D}^{7}} \\ &\leq\!\cdot\; s^{2d+2\sigma} \end{aligned}$$

and

$$\begin{aligned} |\int_0^1 (1-t)\{\{N+R,F\},F\}\circ\varphi_t^{F}\,\mathrm{d}t|_{\mathcal{D}^{2\alpha}} &\leq |\{\{N+R,F\},F\}|_{\mathcal{D}^{\alpha}}\cdot\int_0^1(1-t)\mathrm{d}t \\ &\leq\!\cdot\; s^{-2d}\,|\{N+R,F\}|_{\mathcal{D}^{7}}\,|F|_{\mathcal{D}^{7}} \\ &\leq\!\cdot\; s^{-2d}\,|F|^2_{\mathcal{D}^{6}} \\ &\leq\!\cdot\; s^{2d+2\sigma} \quad . \end{aligned}$$

Furthermore (6.25) yields

$$|(P-R)\circ\varphi_{F}\circ\phi_1\circ\phi_2|_{\mathcal{D}^{4\alpha}} \;\leq\!\cdot\; \alpha^{1-\sigma}s^\kappa\varepsilon \quad .$$

We have already remarked that

$$|\tfrac{1}{2a_+}(\sum_{j=0}^{[\frac{d}{2}]} Q_j(\lambda)q^j)^2|_{\mathcal{D}^{3\alpha}} \;\leq\!\cdot\; s^{\sigma-\kappa}s^\kappa\varepsilon$$

and finally we obtain

$$\begin{aligned} |\{\, N_0, \sum_{\substack{0<|k|\leq K\\ d+2<m\leq 2d}} F_{km}\mathrm{e}^{\mathrm{i}(k|x)}\,\}|_{\mathcal{D}^{2\alpha}} &\leq\!\cdot\; s^{d-2}\,|\sum_{\substack{0<|k|\leq K\\ d+2<m\leq 2d}} F_{km}\mathrm{e}^{\mathrm{i}(k|x)}|_{\mathcal{D}^{7}} \\ &\leq\!\cdot\; s^{d-2}\varepsilon \quad . \end{aligned}$$

In total, we get

$$|P_+|_{\mathcal{D}_+} \;\leq\; |P_+|_{\mathcal{D}^{4\alpha}} \;\leq\; c\,\alpha^{(1-\sigma)(\sigma-\kappa)}\,s^\kappa\,\varepsilon \;<\; \varepsilon_+$$

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

$$|\frac{\partial^{|l|+i+j+|h|}P_+}{\partial y^l\partial p^i\partial q^j\partial\lambda^h}|_{\mathcal{D}_+} \;\leq\; s_+^{-m}\varepsilon_+$$

by Cauchy's inequality. This concludes the proof of Lemma 6.1.

Remark : The latter inequality will be used to prove that all the coefficients of P_∞ with weighted order less than $2d$ vanish.

6.2 Iteration and Convergence

In the previous subsection we were concerned with one step of the iteration process. Thus, given a small perturbation $H_\nu = N_\nu + P_\nu$ of our normal form N_ν , we now know how to construct a co–ordinate change Ψ_ν such that $H_{\nu+1} := H_\nu \circ \Psi_\nu$ is an even smaller perturbation of the adapted normal form $N_{\nu+1}$. 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_∞ no longer forms an obstruction for the desired conclusions.

By composition $\Phi_{\nu+1} := \Psi_0 \circ \Psi_1 \circ \ldots \circ \Psi_\nu$ we obtain a co–ordinate transformation that turns the given $H_0 = N_0 + P_0$ into $N_{\nu+1} + P_{\nu+1}$. Our aim is to find a "limit" Φ_∞ with

$$H_0 \circ \Phi_\infty \quad = \quad N_\infty \; + \; P_\infty \quad .$$

The occurrence of P_∞ reflects that $\lim\limits_{\nu\to\infty} \Phi_\nu$ is only defined on $\bigcap \mathcal{D}_\nu = U_{\frac{r_0}{2}}(\mathsf{T}) \times \{0\} \times \{0\} \times \{0\} \times \mathcal{O}'_1$. To obtain the desired convergence we will need a bound on the C^μ–norm

$$\|\Phi_\nu\|_{C^\mu(\mathcal{D}_\nu)} \quad = \quad \max_{|l|+i+j\le\mu} \left\| \frac{\partial^{|l|+i+j}\Phi_\nu}{\partial y^l \partial p^i \partial q^j} \right\|_{\mathcal{D}_\nu} \quad .$$

Lemma 6.8 *A constant $c > 0$ exists, depending only on n, τ, d and μ , such that* $\|\Phi_\nu\|_{C^\mu(\mathcal{D}_\nu^{\alpha_\nu})} \;\le\; c$ *for every* $\nu \in \mathbb{N}$.

Proof. Firstly, we note that the estimate in this lemma holds for $\Psi_\nu = \varphi_{F_\nu} \circ \phi_1^\nu \circ \phi_2^\nu$ by Lemma 6.7. From $\Phi_{\nu+1} = \Phi_\nu \circ \Psi_\nu$ it follows that $\frac{\partial}{\partial z}\Phi_{\nu+1} = \left(\frac{\partial}{\partial z}\Phi_\nu \circ \Psi_\nu\right) \bullet D\Psi_\nu$ and we have to estimate products of the form

$$\Big(\frac{\partial^{|l|+i+j}\Phi_\nu}{\partial y^l \partial p^i \partial q^j} \circ \Psi_\nu \Big) \bullet (D^{\iota_1}\Psi_\nu, \ldots, D^{\iota_\ell}\Psi_\nu)$$

with $\ell = |l| + i + j$. The estimates for Φ_ν can be proven inductively **;** **q.e.d.**

In the case of *e.g.* normally elliptic tori the transformations one works with form a group, cf. [Pöschel;89]. 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 Ψ_ν 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_ν that generates the first part φ_{F_ν} of the co–ordinate transformation Ψ_ν and by its second part defined explicitly in (6.30) and (6.32). The problem is now that one cannot restrict to the fixed weighted order $2d$ in (y, p, q, λ) imposed by the type of the bifurcation, as the composition already of Ψ_ν and $\Psi_{\nu+1}$ would increase this order to $4d$. Therefore, we have to pass to a polynomial truncation of fixed degree in order to define Φ_∞ 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" Υ_ν have to be symplecto–morphisms as well.
2. The estimates implied by Lemma 6.1 should remain valid after the transformed Hamiltonian functions $H_0 \circ \Phi_\nu$ are replaced by the Hamiltonians $H_0 \circ \Upsilon_\nu$.

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

$$\sum_{i=1}^{n} (y_i - Y_i)\,\mathrm{d}x_i \;+\; (X_i - x_i)\,\mathrm{d}Y_i \;\;+\;\; (Q-q)\,\mathrm{d}P \;+\; (p-P)\,\mathrm{d}q \qquad (6.34)$$

is closed and can therefore be written as $\mathrm{d}S_\nu$. Indeed, being composed from finitely many translations ϕ_2^μ , shear transformations ϕ_1^μ and time one maps φ_F^μ , the transformation Φ_ν is homotopic to the identity. Thus, the closed one–form (6.34) is exact, *i.e.* S_ν is one–valued. The function $S_\nu = S_\nu(x,Y,P,q)$ is a generating function for Φ_ν , cf. [Arnol'd;74]. Note that S_ν itself is only determined up to a constant and that all partial derivatives are 2π–periodic in the toral co–ordinates $x_1,\ldots,x_n$.

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

To be precise, we write

$$\Phi_\nu(x,y,p,q,\lambda,\omega) \;\;=\;\; ((x,y,p,q) + W_\nu(x,y,p,q,\lambda,\omega), \lambda + \widetilde{\Lambda}_\nu(\lambda,\omega), \omega + \widetilde{\Omega}_\nu(\lambda,\omega))$$

and let $\mathcal{F}_\nu : \mathcal{D}_\nu \longrightarrow \mathcal{D}_0$ denote the transformation of (x,y,p,q,λ,ω) into

$$(\,x\,,\,y + W_\nu^2(x,y,p,q,\lambda,\omega)\,,\,p + W_\nu^3(x,y,p,q,\lambda,\omega)\,,q\,,\;\omega\,,\,\lambda\,) \;\stackrel{!}{=}\; (\,x\,,\,Y\,,\,P\,,\,q\,,\,\omega\,,\,\lambda\,)$$

and $\mathcal{G}_\nu := \mathcal{F}_\nu^{-1}$. The truncations $\widetilde{S}_\nu$ are polynomials in Y,P,q and λ , the coefficients of which are holomorphic functions in x and ω . To truncate we write S_ν as a Taylor series at $\mathcal{F}_\nu(x,0,0,0,0,\omega) =: (x,Y_\nu,P_\nu,0,0,\omega)$. Therefore,

$$S_\nu^{lijh}(x,\omega) \;\;=\;\; \frac{\partial^{|l|+i+j+|h|}S_\nu}{\partial Y^l \partial P^i \partial q^j \partial \lambda^h}(x,Y_\nu,P_\nu,0,0,\omega) \quad ,$$

and we define

$$\widetilde{S}_\nu(x,y,p,q,\lambda,\omega) \;\;:=\;\; \sum_{|l|+i+j=0}^{d+1} \;\; \sum_{|h|=0}^{\min(|l|+i+j,1)} S_\nu^{lijh}(x,\omega)\cdot(Y-Y_\nu)^l(P-P_\nu)^i q^j \lambda^h \quad .$$

Lemma 6.9 *Under the conditions of Lemma 6.1 the sequence* $(\widetilde{S}_\nu)_{\nu\in\mathbf{N}}$ *of truncations is uniformly convergent on* $\overline{U_{\frac{r_0}{2}}(\mathsf{T})} \times \mathcal{O}'_\gamma$.

Proof. For $i \geq 1$ we can use $\dfrac{\partial S_\nu}{\partial P} = W_\nu^4 \circ \mathcal{G}_\nu$. We immediately get

$$\begin{aligned} |\, S_{\nu+1}^{0100}(x,\omega) - S_\nu^{0100}(x,\omega)\,| &= |\, W_{\nu+1}^4(x,0,0,0,0,\omega) - W_\nu^4(x,0,0,0,0,\omega)\,| \\ &\leq \|\Phi_\nu\|_{C^2(\mathcal{D}_\nu)} \, |(\Psi_\nu - \mathrm{id})(x,0,0,0,0,\omega)\,| \;\; \leq. \;\; s_\nu^\sigma \end{aligned}$$

and this exponential decay yields a limit $\widetilde{S}_\infty^{0100}$. For higher derivatives we use the chain rule to write $S_\nu^{lijh}(x,\omega)$ as a sum of terms

$$\Big(\frac{\partial^\mu W_\nu^4}{\partial y^{l'} \partial p^{i'} \partial q^{j'} \partial \lambda^{h'}} \circ \mathcal{G}_\nu\Big)(D^{\iota_1}\mathcal{G}_\nu, \ldots, D^{\iota_\mu}\mathcal{G}_\nu)$$

with $(l',i',j',h') \leq (l,i,j,h)$ in all components and $\iota_1 + \ldots + \iota_\mu = |l| + i + j - 1$. Using $D\mathcal{G}_{\nu+1} - D\mathcal{G}_\nu = D\mathcal{G}_\nu(D\mathcal{F}_\nu - D\mathcal{F}_{\nu+1})D\mathcal{G}_{\nu+1}$ we obtain again an exponential decay for the difference $S_{\nu+1}^{lijh}(x,\omega) - S_\nu^{lijh}(x,\omega)$ and thus a limit coefficient function $\widetilde{S}_\infty^{lijh}(x,\omega)$.

By means of $\dfrac{\partial S_\nu}{\partial Y} = W_\nu^1 \circ \mathcal{G}_\nu$ and $\dfrac{\partial S_\nu}{\partial q} = -W_\nu^3 \circ \mathcal{G}_\nu$ we obtain this same result for $|l| \geq 1$ and $j \geq 1$ as well.

We are left with the coefficient functions $S_\nu^{0000}(x,\omega)$. Their averages vanish and they are determined by $\dfrac{\partial S_\nu}{\partial x} = -W_\nu^1 \circ \mathcal{G}_\nu$. We conclude

$$\begin{aligned} |\, S_{\nu+1}^{0000}(x,\omega) - S_\nu^{0000}(x,\omega)\,| &\leq \max_{i=1,\ldots,n} \Big|\int_0^{x_i} W_\nu^{2,i} \; - \; W_\nu^{2,i} \circ \mathcal{G}_\nu \, \mathrm{d}\tilde{x}_i \,\Big| \\ &\leq |W_\nu^2 \; - \; W_\nu^2 \circ \mathcal{G}_\nu| \;\; \leq. \;\; s_\nu^\sigma \quad ; \end{aligned}$$

q.e.d.

Instead of investigating $\Phi_\infty = \lim_{\nu\to\infty} \Upsilon_\nu$ directly, we work with the (truncated) generating functions $\widetilde{S}_\nu$. Using the Inverse Approximation Lemma, cf. [Zehnder;75], we obtain Whitney–C^∞–smooth limit functions $\widetilde{S}_\infty^{lijh}$ on $\overline{U_{\frac{r_0}{2}}(\mathsf{T})} \times \mathcal{O}'_\gamma$. They constitute the coefficients of a generating function

$$\widetilde{S}_\infty \; : \; U_{\frac{r_0}{2}}(\mathsf{T}) \times \mathbb{C}^n \times \mathbb{C}^2 \times \mathbb{C}^{d-2} \times \mathcal{O}'_\gamma \longrightarrow \mathbb{C}$$

which is analytic (since polynomial) in x, Y, P, q and λ . With Whitney's Extension Theorem, cf. [Whitney;34], we get $\widetilde{S}_\infty(x,Y,P,q,\lambda,\omega)$ for all $\omega \in \mathbb{R}^n$. This defines for every (λ,ω) a symplecto–morphism on $\mathsf{T} \times \mathbb{R}^n \times \mathbb{R}^2$. To obtain Φ_∞ we have to complete these symplecto–morphisms by $\mathrm{id} + (\widetilde{\Lambda}_\infty, \widetilde{\Omega}_\infty) = \mathrm{id} + \lim_{\nu\to\infty} (\widetilde{\Lambda}_\nu, \widetilde{\Omega}_\nu)$. This latter convergence to Whitney–C^∞–smooth functions is an immediate consequence of Lemma 6.8 and the Inverse Approximation Lemma.

To conclude the proof of Theorem 2.1 (with $\gamma = 1$) we apply the Inverse Approximation Lemma to the coefficient functions a_ν, b_ν of the normal forms N_ν and obtain a Whitney–C^∞–smooth Hamiltonian function N_∞ which is (again) analytic in y, p, q and λ . Letting $P_\infty := H_0 \circ \Phi_\infty - N_\infty$ we have, according to our choice of the truncations $\widetilde{S}_\nu$ of S_ν at order $d+1$,

$$P_{\infty,\,lijh} = \lim_{\nu\to\infty} P_{\nu,\,lijh}$$

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 2.1 (with $\gamma = 1$).

Remark : While the normal form (1.1) comes from singularity theory, the above proof does not use the preparation theorem (the main tool to deal with universal unfoldings of singularities, cf. [Bröcker,Lander;75]). Indeed, we could use the explicit transformations (6.30) and (6.32) to get rid of some lower order terms, whereas occurring higher order terms were included in the new perturbation (6.33) of the KAM iteration scheme. Note that both the preparation theorem and KAM theory ultimately rely on the implicit mapping theorem.

6.3 Scaling properties

The final step in our proof of Theorem 2.1 consists in re–installing the Diophantine constant γ by means of a suitable scaling. The scaling properties of Theorem 2.1 with respect to γ are also important in applications. While γ will always be chosen as small as the perturbation allows, one often has to choose γ as a function of occurring parameters that for certain values force $\gamma \searrow 0$. Therefore, we address this question in a slightly more general fashion and consider the co–ordinate transformation

$$(x, y, p, q, \lambda, \omega) \;\; \mapsto \;\; (X, Y, P, Q, \Lambda, \Omega)$$

with

$$\begin{aligned} X &= x \\ \Omega &= \gamma^{-1}\omega \end{aligned}$$

and, using $\Gamma = \gamma^{\varsigma}$ with a fixed exponent ς that may be specified later on,

$$\begin{aligned} Y &= \Gamma^{-2d+1} y \\ P &= \Gamma^{-d} p \\ Q &= \Gamma^{-2} q \\ \Lambda_j &= \Gamma^{-2d+2j} \lambda_j \quad . \end{aligned}$$

In the new co–ordinates the normal form (1.1) becomes

$$\gamma^{2d\varsigma + (1-\varsigma)} \hat{N}(X, Y, P, Q; \Lambda, \Omega)$$

where the coefficient functions $\hat{a}$ and $\hat{b}$ of $\hat{N}$ are related to those of N through

$$\begin{aligned} \hat{a}(\Omega) &= \gamma^{\varsigma - 1} \, a(\gamma\Omega) \\ \hat{b}(\Omega) &= \gamma^{\varsigma - 1} \, b(\gamma\Omega) \quad . \end{aligned}$$

For $\varsigma = 1$ this yields the desired re–installation of the Diophantine constant γ .

As already remarked in Section 3.1, perturbations of superintegrable systems lead to dynamics with two time scales : superposed to the fast unperturbed quasi–periodic motion one has a slow movement of *e.g.* the angular momentum in the rigid body example of Section 3.1. The slow dynamics is induced by the perturbation and may involve a bifurcation scenario, with fast dynamics on the invariant tori. The coefficient functions a and b of a normal form (1.1) would then be small as well, *i.e.* of the order δ of the perturbation. However, this can easily be remidied through a scaling with $\varsigma = \frac{1}{2d-1}$, leading to coefficient functions $\hat{a}$ and $\hat{b}$ that do satisfy the necessary estimates from below. In this process we end up with a very small Diophantine constant $\gamma \sim \delta^{1+\frac{1}{2d-2}}$, expressing the well–known fact that much more invariant tori become Diophantine in systems with two time scales.

Acknowledgments. We thank Florian Wagener for valuable discussions during the preparation of this paper. The figures were produced with Mathematica [Wolfram;88], with the exception of *figure 2b* for which we thank Carles Simó. JY was supported by the Alexander von Humboldt Foundation in Germany. HWB and HH acknowledge support from the FOM grant MFGBb1.

Bibliography

V.I. Arnol'd, V.A. Vasil'ev, V.V. Goryunov, O.V. Lyashko [88] Singularity theory I : Singularities, local and global theory ; in *Dynamical Systems VI* (ed. V.I. Arnol'd) Springer (1993)

V.I. Arnol'd [74] *Mathematical Methods of Classical Mechanics* ; Springer (1978)

G. Benettin, F. Fassò [96] Fast rotations of the rigid body : a study by Hamiltonian perturbation theory. Part I ; *Nonlinearity* **9**, p. 137 – 186 (1996)

N. Bourbaki [85] *Commutative Algebra* ; Springer (1989)

B.J.L. Braaksma, H.W. Broer [87] On a quasi–periodic Hopf bifurcation ; *Ann. Inst. H. Poincaré, Analyse non linéaire* **4**(2), p. 115 – 168 (1987)

B.J.L. Braaksma, H.W. Broer, G.B. Huitema [90] Toward a quasi–periodic bifurcation theory ; *Mem. AMS* **83** #421, p. 83 – 167 (1990)

Th. Bröcker, L. Lander [75] *Differentiable Germs and Catastrophes* ; Cambridge Univ. Press (1975)

H.W. Broer, S.–N. Chow, Y. Kim, G. Vegter [93] A normally elliptic Hamiltonian bifurcation ; *Z. angew. Math. Phys.* **44**, p. 389 – 432 (1993)

H.W. Broer, S.-N. Chow, Y. Kim, G. Vegter [95] The Hamiltonian Double–Zero Eigenvalue ; p. 1 – 19 in *Normal Forms and Homoclinic Chaos, Waterloo 1992* (eds. W.F. Langford, W. Nagata) Fields Institute Communications **4** (1995)

H.W. Broer, H. Hanßmann, À. Jorba, J. Villanueva, F.O.O. Wagener [02] Normal–internal resonances in quasi–periodically forced oscillators : a conservative approach ; preprint, Universiteit van Amsterdam (2002)

H.W. Broer, G.B. Huitema [91] A proof of the isoenergetic KAM–theorem from the "ordinary" one ; *J. Diff. Eq.* **90**(1), p. 52 – 60 (1991)

H.W. Broer, G.B. Huitema [95] Unfoldings of quasi–periodic tori in reversible systems ; *J. Dynamics Diff. Eq.* **7**(1), p. 191 – 212 (1995)

H.W. Broer, G.B. Huitema, M.B. Sevryuk [96] *Quasi–Periodic Motions in Families of Dynamical Systems : Order amidst Chaos* ; LNM **1645**, Springer (1996)

H.W. Broer, G.B. Huitema, F. Takens [90] Unfoldings of quasi–periodic tori ; *Mem. AMS* **83** #421, p. 1 – 82 (1990)

H.W. Broer, R. Roussarie [01] Exponential confinement of chaos in the bifurcation set of real analytic diffeomorphisms ; p. 167 – 210 in *Global Analysis of Dynamical Systems, Leiden 2001* (eds. H.W. Broer, B. Krauskopf, G. Vegter) IoP (2001)

H.W. Broer, R. Roussarie, C. Simó [96] Invariant circles in the Bogdanov–Takens bifurcation for diffeomorphisms ; *Erg. Th. Dyn. Syst.* **16**, p. 1147 – 1172 (1996)

H.W. Broer, F. Takens, F.O.O. Wagener [99] Integrable and non–integrable deformations of the skew–Hopf bifurcation ; *Reg. & Chaot. Dyn.* **4**(2), p. 17 – 43 (1999)

H.W. Broer, G. Vegter [92] Bifurcational Aspects of Parametric Resonance ; *Dynamics Rep., new ser.* **1**, p. 1 – 53 (1992)

H.W. Broer, F.O.O. Wagener [00] Quasi–periodic stability of subfamilies of an unfolded skew–Hopf bifurcation ; *Arch. Rat. Mech. An.* **152**, p. 283 – 326 (2000)

S. Ferrer, H. Hanßmann, J. Palacián, P. Yanguas [02] On perturbed oscillators in 1-1-1 resonance : the case of axially symmetric cubic potentials ; *J. Geom. Phys.* **40**, p. 320 – 369 (2002)

C.G. Gibson, K. Wirthmüller, A.A. du Plessis, E.J.N. Looijenga [76] *Topological stability of smooth mappings* ; LNM **552**, Springer (1976)

M. Golubitsky, D. Schaeffer [85] *Singularities and Groups in Bifurcation Theory* ; Springer (1985)

M. Golubitsky, D. Schaeffer, I. Stewart [88] *Singularities and Groups in Bifurcation Theory II* ; Springer (1988)

H. Hanßmann [95] *Quasi–periodic Motions of a Rigid Body — A case study on perturbations of superintegrable systems* ; Ph.D. thesis, Rijksuniversiteit Groningen (1995)

H. Hanßmann [96] Equivariant Perturbations of the Euler Top ; p. 227 – 253 in *Nonlinear Dynamical Systems and Chaos, Groningen 1995* (eds. H.W. Broer et al.) Birkhäuser (1996)

H. Hanßmann [98] The Quasi–Periodic Centre–Saddle Bifurcation ; *J. Diff. Eq.* **142**(2), p. 305 – 370 (1998)

H. de Jong [99] *Quasiperiodic breathers in systems of weakly coupled pendulums* ; Ph.D. thesis, Rijksuniversiteit Groningen (1999)

I. Kan, H. Koçak, J.A. Yorke [92] Antimonotonicity : Concurrent creation and annihilation of periodic orbits ; *Ann. Math.* **136**(2), p. 219 – 252 (1992)

M. Mazzocco [97] KAM theorem for generic analytic perturbations of the Euler system ; *Z. angew. Math. Phys.* **48**, p. 193 – 219 (1997)

K.R. Meyer [70] Generic bifurcation of periodic points ; *Trans. AMS* **149**, p. 95 – 107 (1970)

K.R. Meyer [75] Generic Bifurcation in Hamiltonian Systems ; p. 62 – 70 in *Dynamical Systems — Warwick 1974* (ed. A. Manning) LNM **468**, Springer (1975)

J. Moser [65] Combination tones for Duffings equation ; *Comm. Pure Appl. Math.* **18**, p. 167 – 181 (1965)

J. Moser [67] Convergent Series Expansion for Quasi–Periodic Motions ; *Math. Ann.* **169**(1), p. 136 – 176 (1967)

C. Nelk [98] *Bewegung eines starren Körpers in einem symmetrischen affinen Kraftfeld* ; Diplomarbeit, RWTH Aachen (1998)

J. Palis, F. Takens [93] *Hyperbolicity & sensitive chaotic dynamics at homoclinic bifurcations* ; Cambridge Univ. Press (1993)

V. Poènaru [76] *Singularités C^∞ en Présence de Symétrie* ; LNM **510**, Springer (1976)

J. Pöschel [82] Integrability of Hamiltonian Systems on Cantor Sets ; *Comm. Pure Appl. Math.* **35**, p. 653 – 696 (1982)

J. Pöschel [89] On Elliptic Lower Dimensional Tori in Hamiltonian Systems ; *Math. Z.* **202**, p. 559 – 608 (1989)

M.B. Sevryuk [86] *Reversible Systems* ; LNM **1211**, Springer (1986)

J.J. Stoker [50] *Nonlinear Vibrations* ; Interscience (1950)

F. Takens, F.O.O. Wagener [00] Resonances in skew and reducible quasi–periodic Hopf bifurcations ; *Nonlinearity* **13**, p. 377 – 396 (2000)

F.O.O. Wagener [98] *On the skew Hopf bifurcation* ; Ph.D. thesis, Rijksuniversiteit Groningen (1998)

H. Whitney [34] Analytic extensions of differentiable functions defined in closed sets ; *Trans. AMS* **36**, p. 63 – 89 (1934)

H. Whitney [65] Local properties of analytic varieties ; p. 205 – 244 in *Differential and Combinatorial Topology* (ed. S.S. Cairns) Princeton Univ. Press (1965)

S. Wolfram [88] *Mathematica : A System for Doing Mathematics by Computer* ; Addison-Wesley (1988)

J. Xu, J. You [01] Persistence of lower–dimensional tori under the first Melnikov's non–resonance condition ; *J. Math. Pures Appl.* **80**(10), p. 1045 – 1067 (2001)

J. You [98] A KAM Theorem for Hyperbolic–Type Degenerate Lower Dimensional Tori in Hamiltonian Systems ; *Comm. Math. Phys.* **192**(1), p. 145 – 168 (1998)

J. You [99] Lower dimensional tori of Reversible Hamiltonian Systems in the resonant zone ; p. 301 – 314 in *Dynamical Systems in Honor of Prof. Liao Shantao*, World Scientific (1999)

E. Zehnder [75] Generalized Implicit Function Theorems with Applications to Some Small Divisor Problems I ; *Comm. Pure Appl. Math.* **28**, p. 91 – 140 (1975)