

EXISTENCE OF KAM TORI FOR PRESYMPLECTIC VECTOR FIELDS

SEAN BAUER, NIKOLA P. PETROV

ABSTRACT. We prove the existence of a torus that is invariant with respect to the flow of a vector field that preserves the presymplectic form in an exact presymplectic manifold. The flow on this invariant torus is conjugate to a linear flow on a torus with a Diophantine velocity vector. The proof has an “a posteriori” format, the invariant torus is constructed by using a Newton method in a space of functions, starting from a torus that is approximately invariant. The geometry of the problem plays a major role in the construction by allowing us to construct a special adapted basis in which the equations that need to be solved in each step of the iteration have a simple structure. In contrast to the classical methods of proof, this method does not assume that the system is close to integrable, and does not rely on using action-angle variables.

1. INTRODUCTION

The goal of this article is to give a proof of the existence of a torus that is invariant with respect to the flow of a presymplectic vector field V in an exact presymplectic manifold (\mathcal{P}, Ω) , i.e., in a manifold \mathcal{P} endowed with an exact constant-rank 2-form Ω that is preserved under the flow of V .

Perhaps the most prominent occurrence of presymplectic manifolds in physics is in the geometric theory of dynamical systems with constraints. These are systems for which the transition from Lagrangian to Hamiltonian description is non-trivial because some of the relations $p_A := \frac{\partial L}{\partial \dot{q}^A}(q, \dot{q})$ expressing the generalized momenta p_A in terms of generalized velocities \dot{q}^A cannot be solved for \dot{q}^A since the matrix $(\frac{\partial^2 L}{\partial \dot{q}^A \partial \dot{q}^B})$ is degenerate; the relations that cannot be solved play the role of constraints. The modern theory of constrained systems was initiated in the early 1950s by Dirac [23, 24, 25] and developed by Bergmann and his collaborators for purposes of quantization of field theories [4, 10, 44] (the book [52] offers an in-depth exposition). Such situations occur in also classical electromagnetic theory [53, Ch. V], in the description of relativistic particles [36], [53, Ch. VII], gauge fields [38, 47, 53].

A geometric theory of constrained systems was proposed in the late 1970s by Gotay and collaborators [33, 34, 35]. In their approach, the system is transformed in stages, and the process ends up with a manifold that is typically presymplectic.

Presymplectic geometry is also related to equivalence between Lagrangian and Hamiltonian formalisms for constrained systems [7, 8, 15], geometric approach to

2010 *Mathematics Subject Classification*. 34D35, 37J40, 70K43, 70H08.

Key words and phrases. KAM theory; invariant torus; presymplectic manifold; stability.

©2020 Texas State University.

Submitted April 5, 2019. Published December 22, 2020.

maximum principles [6], geometric optics [17, 18, 27], etc. Other topics of interest are canonical transformations in presymplectic systems [14, 16], reduction of presymplectic manifolds [2, 22, 28, 29, 45], etc.

A major achievement in the theory of Hamiltonian systems in the second half of the XX century was the celebrated Kolmogorov-Arnold-Moser (KAM) Theorem. This theorem is the subject of multiple reviews and pedagogical expositions (see, e.g., [5, 11, 19, 20, 41, 46, 50, 51]), and its history is beautifully described in the recent book by Dumas [26].

The standard proofs of the KAM-type theorems perform an infinite sequence of canonical transformations to convert a slightly perturbed integrable system on a symplectic manifold into action-angle variables. González, Jorba, de la Llave and Villanueva developed a new method of proof in their seminal 2005 paper [21]. This method relies heavily on the geometry of the system. One important ingredient in their proof is the so-called *automatic reducibility*: if \mathcal{K} is a torus in the symplectic manifold \mathcal{P} that is invariant with respect to a map $f : \mathcal{P} \rightarrow \mathcal{P}$, then the tangent bundle to \mathcal{K} is preserved under the derivative Tf . This simplifies the structure of the coefficient matrices in certain difference equations which, in turn, dramatically simplifies the solution of the problem.

Methods similar to the ones developed in [21] have since been used in [32] to study the existence of non-twist tori in degenerate Hamiltonian systems, and in [30, 42] to prove the existence of lower dimensional invariant tori that are partially hyperbolic or elliptic. Since these methods are suitable for efficient numerical implementation, they have been used for this purpose in [12, 13, 31, 39]. Many aspects of these methods are considered in the recent book by Haro *et al* [37] (KAM theory is the subject of Chapter 4).

Alishah and de la Llave [3] used the ideas of [21] to prove a KAM theorem for presymplectic systems, for which the degeneracy of the presymplectic form complicates the matters. They considered a family $\{f_\lambda\}$ of maps that preserve the presymplectic form, and found a value $\bar{\lambda}$ of the parameter λ and an embedding K from a torus to the presymplectic manifold such that $f_{\bar{\lambda}} \circ K = K \circ T_\omega$ where $T_\omega : \theta \mapsto \theta + \omega$ is translation on the torus by a Diophantine vector ω .

The main goal of this paper is to prove a KAM theorem for a family $\{V_\lambda\}$ of presymplectic vector fields on an exact presymplectic manifold (\mathcal{P}, Ω) , with $\dim \mathcal{P} = d + 2n$, $\dim \ker \Omega = d$, $\Omega = d\tau$ for some $\tau \in \Omega^1(\mathcal{P})$. For most of the paper we consider $\mathcal{P} \cong \mathbb{T}^d \times T^*\mathbb{T}^n \cong \mathbb{T}^{d+n} \times \mathbb{R}^n$, where $\ker \Omega$ coincides with the first d dimensions. Our goal is to find a value $\bar{\lambda}$ of the parameter $\lambda \in \mathbb{R}^{d+2n}$ and an embedding $K : \mathbb{T}^{d+n} \rightarrow \mathcal{P}$ such that the submanifold $\mathcal{K} := K(\mathbb{T}^{d+n})$ is invariant with respect to the flow $\Phi_{\bar{\lambda}, t}$ of the vector field $V_{\bar{\lambda}}$, and K conjugates $\Phi_{\bar{\lambda}, t}$ to the linear flow $\phi_t : \mathbb{T}^{d+n} \rightarrow \mathbb{T}^{d+n} : \theta \mapsto \theta + t\omega$, where $\omega \in \mathbb{R}^d$ is a constant Diophantine vector:

$$\Phi_{\bar{\lambda}, t} \circ K = K \circ \phi_t, \quad t \geq 0. \quad (1.1)$$

The infinitesimal form of (1.1) is $V_{\bar{\lambda}, K(\theta)} = K_{*\theta} \omega_\theta$, where $K_{*\theta} : T_\theta \mathbb{T}^{d+n} \rightarrow T_{K(\theta)} \mathcal{P}$ is the derivative of K at $\theta \in \mathbb{T}^{d+n}$ and we consider $\omega \in \mathbb{R}^{d+n}$ as $\omega_\theta \in T_\theta \mathbb{T}^{d+n} = \mathbb{R}^{d+n}$.

Our proof of the theorem has an *a posteriori* format (as in [21]). In more detail, we assume the existence of λ_0 and $K_0 : \mathbb{T}^{d+n} \rightarrow \mathcal{P}$ that satisfy (1.1) only approximately, i.e., $\Phi_{\lambda_0, t} \circ K_0 \approx K_0 \circ \phi_t$. Then we start a version of the Newton

method to construct iteratively a sequence of better and better approximations

$$(\lambda_0, K_0) \mapsto (\lambda_1, K_1) \mapsto (\lambda_2, K_2) \mapsto (\lambda_3, K_3) \mapsto \cdots \quad (1.2)$$

whose limit $(\lambda_\infty, K_\infty)$ is the desired solution $(\bar{\lambda}, K)$ satisfying (1.1). The loss of domain accompanying each step of the iteration is compensated by the fast convergence of the Newton method. This method is convenient to implement in numerical computations. Moreover, *a posteriori* theorems are suitable for validation of numerical results, i.e., they can be used to produce computer assisted proofs of existence of invariant manifolds.

In this article we extensively use the geometry of the system to our advantage, inspired by the ideas of [21, 3]. If $\mathcal{K} = K(\mathbb{T}^{d+n})$ is the invariant torus, then at each point $k \in \mathcal{K}$ the kernel $\ker \Omega_k$ of the presymplectic form is a subspace of $T_k\mathcal{K}$, so that we have $\ker \Omega_p \subseteq T_k\mathcal{K} \subseteq T_k\mathcal{P}$. In fact, we have much more – a filtration of subbundles $\ker \Omega \subseteq T\mathcal{K} \subseteq T\mathcal{P}|_{\mathcal{K}}$ which is invariant with respect to the flow $\Phi_{\bar{\lambda}, t}$ of the vector field $V_{\bar{\lambda}}$. This and the invariance (1.1) allow us to construct a special basis adapted to the filtration, in which the matrices of the operators have zero blocks. Even if the torus is only approximately invariant (as in the case of (λ_j, K_j)), these blocks, albeit non-zero, are small and we have good bounds on their norms.

An important role is also played by the fact that \mathcal{K} is isotropic (i.e., that the pull-back of Ω to \mathcal{K} vanishes identically), and the approximately invariant tori are approximately isotropic. We found the following interesting quotations related to this fact. On page 45 of his classic 1973 monograph [43], Moser writes

Actually, more than asserted in Theorem 2.7 can be proven. It turns out that the differential form $\sum_{k=1}^n dy_k \wedge dx_k$ vanishes identically on the tori (3.11), and one calls manifolds with this property and of maximal dimension Lagrange manifolds.

In this quotation, Theorem 2.7 is (as Moser calls it) the Kolmogorov-Arnold Theorem, and the tori (3.11) are the invariant tori whose existence is proved in the KAM theorem. On page 584 of their monumental book [1], Abraham and Marsden write

Moser [1973a] states that the invariant tori are Lagrangian submanifolds [...]. This fact can probably be exploited, although to our knowledge it has not been.

The fact that \mathcal{K} is isotropic and of a maximum dimension (i.e., Lagrangian in the symplectic case) is crucial for the proofs in [21, 3] and in this paper.

The paper is organized as follows. In Sections 2.1–2.5 we introduce some basic definitions and notations, discuss the integrability of the distribution $\ker \Omega$ and construct the symplectic manifold $\mathcal{P}/\ker \Omega$. The main theorem is stated in Section 2.6.

In Section 3 we study the geometric structures occurring when we know the true solution $(\bar{\lambda}, K)$ of the problem – we prove that \mathcal{K} is isotropic (Section 3.1), give a detailed construction of the basis adapted to the filtration $\ker \Omega \subseteq T\mathcal{K} \subseteq T\mathcal{P}|_{\mathcal{K}}$ (Section 3.2), and study the properties of the matrix of transition from a general basis of $T\mathcal{P}|_{\mathcal{K}}$ to the special adapted basis (Section 3.3).

Section 4 is devoted to the properties of approximate solutions. Approximately isotropic tori are studied in Section 4.1. In Section 4.2 we derive an equation for the corrections ε_j and Δ_j needed to obtain a better approximation $\lambda_{j+1} = \lambda_j + \varepsilon_j$, $K_{j+1} = K_j + \Delta_j$. We solve this equation in Section 4.3, relying heavily

on the machinery developed in Section 3. Finally, in Section 5, we collect several lemmata that justify the applicability of the Newton method for performing the iteration (1.2).

2. PRELIMINARIES AND GENERAL SETUP

In this section we set up the problem and state our main result.

2.1. Presymplectic manifolds and vector fields.

Definition 2.1. A *presymplectic manifold* is a pair (\mathcal{P}, Ω) , where \mathcal{P} is a manifold of any (finite) dimension and $\Omega \in \Omega^2(\mathcal{P})$ is a closed 2-form with constant rank. If Ω is exact, i.e., if $\Omega = d\tau$ for some $\tau \in \Omega^1(\mathcal{P})$, then we say that (\mathcal{P}, Ω) is an *exact presymplectic manifold*.

Throughout this paper, we will always assume that

$$\dim \mathcal{P} = d + 2n, \quad \text{rank } \Omega = 2n. \quad (2.1)$$

Most of the time we will consider the specific exact presymplectic manifold

$$\mathcal{P} := \mathbb{T}^d \times T^*\mathbb{T}^n \cong \mathbb{T}^d \times \mathbb{T}^n \times \mathbb{R}^n \quad (2.2)$$

with an exact presymplectic form Ω with $\ker \Omega = \mathbb{T}^d$. We assume that $\mathbb{T}^d \times T^*\mathbb{T}^n$ is endowed with an Euclidean structure, so that we can identify two-forms with linear operators and abstract tangent vectors with column vectors.

In the definition below, $\mathfrak{X}(\mathcal{P})$ stands for the vector fields on \mathcal{P} , \mathcal{L} is the Lie derivative, and ι is the interior product, i.e., the contraction with a vector field.

Definition 2.2. Let $V \in \mathfrak{X}(\mathcal{P})$ and $\Phi_t: \mathcal{P} \rightarrow \mathcal{P}$ be the time- t flow of V . The vector field V is said to be *presymplectic* if $\Phi_t^* \Omega = \Omega$ for all $t \in \mathbb{R}$.

Lemma 2.3. Let (\mathcal{P}, Ω) be a presymplectic manifold, and $V \in \mathfrak{X}(\mathcal{P})$. Then the following conditions are equivalent:

- (a) V is a presymplectic vector field;
- (b) the Lie derivative of the presymplectic form along V vanishes: $\mathcal{L}_V \Omega = 0$;
- (c) the 1-form $\iota_V \Omega$ is closed.

Proof. The equivalence of (a) and (b) comes directly from the definition of a Lie derivative, and the equivalence of (b) and (c) follows from Cartan's magic formula and the closedness of Ω : $0 = \mathcal{L}_V \Omega = \iota_V d\Omega + d(\iota_V \Omega) = d(\iota_V \Omega)$. \square

2.2. Foliation induced by $\ker \Omega$. In this section we discuss some results about a general presymplectic manifold (\mathcal{P}, Ω) (not necessarily (2.2)). For any $p \in \mathcal{P}$, define

$$\begin{aligned} \ker \Omega_p &:= \{W_p \in T_p \mathcal{P} : \iota_{W_p} \Omega_p = 0\} \\ &= \{W_p \in T_p \mathcal{P} : \Omega_p(W_p, U_p) = 0 \forall U_p \in T_p \mathcal{P}\} \subseteq T_p \mathcal{P}. \end{aligned}$$

The subspaces $\ker \Omega_p$ form a differentiable distribution, $\ker \Omega$, of rank d . Define

$$\mathfrak{X}^{\ker \Omega}(\mathcal{P}) := \{W \in \mathfrak{X}(\mathcal{P}) : \iota_W \Omega = 0\} = \{W \in \mathfrak{X}(\mathcal{P}) : W_p \in \ker \Omega_p \forall p \in \mathcal{P}\}.$$

Using the classical Frobenius Theorem (see, e.g., [48, Section 3.5], or, especially, [40, Appendix 3]) and the fact that Ω is closed, one can easily obtain

Lemma 2.4. *If Ω is presymplectic, the distribution $\ker \Omega$ is integrable.*

Lemma 2.4 implies that \mathcal{P} has a foliation with d -dimensional leaves such that the tangent space to the leaf through $p \in \mathcal{P}$ at p is $\ker \Omega_p$. We make the additional assumption that the collection of leaves of the foliation forms a smooth manifold \mathcal{Q} (for a discussion see, e.g., [40, Sec. 4.3.3 of Appendix 3]). Let $\pi^{\mathcal{Q}} : \mathcal{P} \rightarrow \mathcal{Q}$ be the canonical projection taking each point $p \in \mathcal{P}$ to the leaf $(\pi^{\mathcal{Q}})^{-1}(\pi^{\mathcal{Q}}(p))$ through it. If $\pi_*^{\mathcal{Q}} : T\mathcal{P} \rightarrow T\mathcal{Q}$ is the derivative of $\pi^{\mathcal{Q}}$, then $\ker \pi_*^{\mathcal{Q}} = \ker \Omega$. The lemma below (see [40, Section III.7]) states that \mathcal{Q} carries a natural symplectic structure.

Lemma 2.5. *Let (\mathcal{P}, Ω) be a presymplectic manifold such that $\ker \Omega$ determines an integrable foliation of \mathcal{P} whose leaves form a smooth manifold \mathcal{Q} , and let $\pi^{\mathcal{Q}} : \mathcal{P} \rightarrow \mathcal{Q}$ be the canonical projection. Then there exists a unique symplectic form $\tilde{\Omega} \in \Omega^2(\mathcal{Q})$ on the manifold \mathcal{Q} such that $(\pi^{\mathcal{Q}})^*\tilde{\Omega} = \Omega$.*

For the case (2.2) considered in this paper, $\mathcal{Q} \cong T^*\mathbb{R}^n$. It would be interesting to investigate the case when \mathcal{P} has a more complicated structure than (2.2).

2.3. Matrix representation of Ω and $\tilde{\Omega}$. We consider the case when \mathcal{P} has product structure (2.2), so $\ker \Omega_p \cong \mathbb{T}^d$ for every $p \in \mathcal{P}$ and the collection of leaves,

$$\mathcal{Q} = \mathcal{P} / \ker \Omega = T^*\mathbb{T}^n, \tag{2.3}$$

is a symplectic manifold with symplectic form $\tilde{\Omega}$. Because of the assumed Euclidean structure on \mathcal{P} , we can identify a 2-form on \mathcal{P} with a linear operator. For any $p \in \mathcal{P}$, let $J_p : T_p\mathcal{P} \rightarrow T_p\mathcal{P}$ be the linear operator corresponding to Ω_p , defined by

$$\langle U_p, J_p W_p \rangle_{\mathbb{R}^{d+2n}} := \Omega_p(U_p, W_p), \quad U_p, W_p \in T_p\mathcal{P} \cong \mathbb{R}^{d+2n}, \tag{2.4}$$

where $\langle \cdot, \cdot \rangle_{\mathbb{R}^{d+2n}}$ is the Euclidean inner product on \mathbb{R}^{d+2n} . Similarly, for any $q \in \mathcal{Q}$, let $\tilde{J}_q : T_q\mathcal{Q} \rightarrow T_q\mathcal{Q}$ be the linear operator corresponding to the symplectic form $\tilde{\Omega}_q$ on $T_q\mathcal{Q}$ at $q \in \mathcal{Q}$: if $\langle \cdot, \cdot \rangle_{\mathbb{R}^{2n}}$ is the Euclidean inner product on \mathbb{R}^{2n} , then

$$\langle \tilde{U}_q, \tilde{J}_q \tilde{W}_q \rangle_{\mathbb{R}^{2n}} := \tilde{\Omega}_q(\tilde{U}_q, \tilde{W}_q), \quad \tilde{U}_q, \tilde{W}_q \in T_q\mathcal{Q} \cong \mathbb{R}^{2n}. \tag{2.5}$$

If we choose a basis for $T_p\mathcal{P} \cong \mathbb{R}^d \times \mathbb{R}^{2n}$ such that the first d vectors form a basis of $\mathbb{R}^d = T_p \ker \Omega$, and the other $2n$ vectors form a basis of $\mathbb{R}^{2n} \cong T_p T^*\mathbb{T}^n$, then we can write J_p in a matrix form as

$$J_p = \begin{bmatrix} 0 & 0 \\ 0 & \tilde{J}_{\pi^{\mathcal{Q}}(p)} \end{bmatrix}, \quad J_p^\top = -J_p, \quad \tilde{J}_q^\top = -\tilde{J}_q. \tag{2.6}$$

Although J_p is not invertible, we define J_p^{-1} as the Moore-Penrose pseudoinverse:

$$J_p^{-1} := \begin{bmatrix} 0 & 0 \\ 0 & \tilde{J}_{\pi^{\mathcal{Q}}(p)}^{-1} \end{bmatrix}, \quad \text{so that} \quad J_p J_p^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{I}_{2n} \end{bmatrix}. \tag{2.7}$$

2.4. Miscellaneous definitions and results.

Definition 2.6. For $\gamma > 0$ and $\sigma \geq d + n - 1$, the set of all $\omega \in \mathbb{R}^{d+n}$ satisfying

$$|\omega \cdot k| \geq \frac{\gamma}{|k|^\sigma} \quad \forall k \in \mathbb{Z}^{d+n} \setminus \{0\} \tag{2.8}$$

is called the set of *Diophantine vectors* and is denoted by $\mathcal{D}(\gamma, \sigma)$.

For any $\rho > 0$, we define the torus “thickened” into the complex direction,

$$\mathbb{T}_\rho^{d+n} := \{\theta \in \mathbb{C}^{d+n}/\mathbb{Z}^{d+n} : |\operatorname{Im} \theta^\alpha| \leq \rho, \alpha = 1, 2, \dots, d+n\}. \quad (2.9)$$

Let $|\cdot|$ stand for the supremum norm on \mathbb{R}^m or \mathbb{C}^m (for any m). Given $\rho > 0$, we define the set of functions \mathcal{W}_ρ as follows:

$$\begin{aligned} \mathcal{W}_\rho := \{K : \mathbb{T}_\rho^{d+n} \rightarrow \mathcal{P} : & (a) K \text{ is real analytic on the interior of } \mathbb{T}_\rho^{d+n}, \\ & (b) K \text{ is continuous on the boundary of } \mathbb{T}_\rho^{d+n}, \\ & (c) K \text{ is periodic of period 1 in all arguments} \}. \end{aligned} \quad (2.10)$$

Define $\|K\|_\rho = \sup_{\theta \in \mathbb{T}_\rho^{d+n}} |K(\theta)|$, so that $(\mathcal{W}_\rho, \|\cdot\|_\rho)$ is a Banach space. For analytic functions $g : \mathcal{B} \rightarrow \mathbb{C}$ (where $\mathcal{B} \subseteq \mathbb{C}$), and any $\ell \in \mathbb{B}$, define the norms

$$|g|_{C^\ell, \mathcal{B}} := \sup_{0 \leq |k| \leq \ell} \sup_{z \in \mathcal{B}} |D^k g(z)|. \quad (2.11)$$

Lemma 2.7 (Cauchy bound). *For $K \in \mathcal{W}_\rho$ and $0 < \delta < \rho$,*

$$\|DK\|_{\rho-\delta} \leq C\delta^{-1}\|K\|_\rho. \quad (2.12)$$

Lemma 2.8 (Rüssmann [49]). *Let $\omega = [\omega^1 \ \omega^2 \ \dots \ \omega^{d+n}]^\top \in \mathcal{D}(\gamma, \sigma)$ and let the function $h : \mathbb{T}^{d+n} \rightarrow \mathcal{P}$ be analytic on \mathbb{T}_ρ^{d+n} and have zero average. Then for any $0 < \delta < \rho$, the differential equation $\partial_\omega v = h$, where $\partial_\omega := \sum_{\alpha=1}^{d+n} \omega^\alpha \frac{\partial}{\partial \theta^\alpha}$ is the directional derivative in the direction of ω , has a unique average-zero solution $v : \mathbb{T}^{d+n} \rightarrow \mathcal{P}$ which is analytic in $\mathbb{T}_{\rho-\delta}^{d+n}$. The solution v satisfies the estimate*

$$\|v\|_{\rho-\delta} < C\gamma^{-1}\delta^{-\sigma}\|h\|_\rho, \quad (2.13)$$

where C is a constant depending only on d , n , and σ .

2.5. General setup and matrix notation. Let $\{V_\lambda\}$ be a $(d+2n)$ -parameter family of presymplectic vector fields on the exact presymplectic manifold (\mathcal{P}, Ω) . Our goal is to construct a smooth embedding

$$K : \mathbb{T}^{d+n} \rightarrow \mathcal{P} \quad (2.14)$$

such that $\mathcal{K} := K(\mathbb{T}^{d+n})$ be invariant with respect to the flow $\Phi_{\bar{\lambda}, t}$ of $V_{\bar{\lambda}}$ for some value $\bar{\lambda}$ and the flow $\Phi_{\bar{\lambda}, t}$ on \mathcal{K} be conjugate to the linear flow ϕ_t on \mathbb{T}^{d+n} :

$$\Phi_{\bar{\lambda}, t}(K(\theta)) = K(\phi_t(\theta)) \quad \forall t \in \mathbb{R}, \forall \theta \in \mathbb{T}^{d+n}, \quad (2.15)$$

where $\omega \in \mathbb{R}^{d+n}$ is a constant Diophantine vector and $\phi_t : \mathbb{T}^{d+n} \rightarrow \mathbb{T}^{d+n} : \theta \mapsto \theta + t\omega$. (For elements of \mathbb{T}^{d+n} , e.g., $\theta + t\omega$, we assume that we take the fractional part of each component.) Differentiate (2.15) with respect to t and set $t = 0$ to obtain

$$V_{\bar{\lambda}, K(\theta)} = K_{*\theta}\omega_\theta \quad \forall \theta \in \mathbb{T}^{d+n}. \quad (2.16)$$

Here $V_{\bar{\lambda}, K(\theta)} \in T_{K(\theta)}\mathcal{K} \subseteq T_{K(\theta)}\mathcal{P}$ is the value of $V_{\bar{\lambda}}$ at $K(\theta) \in \mathcal{K}$, $\omega_\theta \in T_\theta\mathbb{T}^{d+n}$ is the Diophantine vector ω considered as an element of $T_\theta\mathbb{T}^{d+n} = \mathbb{R}^{d+n}$, and $K_{*\theta} : T_\theta\mathbb{T}^{d+n} \rightarrow T_{K(\theta)}\mathcal{K} \subseteq T_{K(\theta)}\mathcal{P}$ is the derivative of K at θ .

Instead of the differential-geometric notations used in (2.16), we will normally use matrix notations. In these notations (2.16) reads

$$V_{\bar{\lambda}, K(\theta)} = DK_\theta \omega \quad \text{or} \quad V_{\bar{\lambda}, K(\theta)} = \partial_\omega K_\theta, \quad (2.17)$$

where ω is considered as a constant column vector, $\partial_\omega := \omega \cdot \nabla$, and

$$DK_\theta := [(DK_\theta)^\alpha] = \left[\frac{\partial K^A}{\partial \theta^\alpha}(\theta) \right] \in M_{d+2n, d+n}(\mathbb{R}).$$

Throughout the paper we will systematically use the notations collected in Table 1.

TABLE 1. Notation for indices and coordinates

Coordinates	Range of indices	Description
$x = (x^A)$	$A, B = 1, 2, \dots, d + 2n$	Coordinates in $\mathcal{P} \cong \mathbb{T}^d \times T^*\mathbb{T}^n$
$\underline{x} = (\underline{x}^\mu)$	$\mu, \nu = 1, 2, \dots, d$	Coordinates in \mathbb{T}^d (the first d in \mathcal{P})
$\tilde{x} = (\tilde{x}^i)$	$i, j = 1, 2, \dots, 2n$	Coordinates in $T^*\mathbb{T}^n$ (the last $2n$ in \mathcal{P})
$\theta = (\theta^\alpha)$	$\alpha, \beta = 1, 2, \dots, d + n$	Coordinates in \mathbb{T}^{d+n}

2.6. Statement of the main theorem.

Theorem 2.9. *Assume that:*

- (1) $\omega \in \mathcal{D}(\gamma, \sigma)$ is a Diophantine vector;
- (2) $\mathcal{P} = \mathbb{T}^d \times T^*\mathbb{T}^n$;
- (3) Ω is an exact presymplectic form on \mathcal{P} of rank $2n$ such that the kernel of Ω coincides with the first d directions;
- (4) $\{V_\lambda\}$ is a $(d + 2n)$ -parameter family of analytic presymplectic vector fields on \mathcal{P} ;
- (5) $K_0 : \mathbb{T}^{d+n} \rightarrow \mathcal{P}$ is an embedding belonging to the class \mathcal{W}_{ρ_0} (2.10);
- (6) the value λ_0 of the parameter λ is such that the pair (λ_0, K_0) is non-degenerate in the sense of Definition 4.7;
- (7) each vector field from the family $\{V_\lambda\}$ can be holomorphically extended to some complex neighborhood \mathcal{B}_r of $K_0(\mathbb{T}_\rho^{d+n})$, where

$$\mathcal{B}_r := \{z \in \mathbb{C}^{d+2n} \mid \exists \theta \in \mathbb{T}_\rho^{d+n} \text{ such that } |z - K_0(\theta)| < r\}, \tag{2.18}$$

for some $r > 0$ and such that $|V_\lambda|_{C^2, \mathcal{B}_r}$ is finite.

We define the error function $e_{0,\theta} := V_{\lambda, K_0(\theta)} - \partial_\omega K_{0,\theta}$. Then there exists a constant $c > 0$ which depends on $d, n, \sigma, \rho_0, \|DK_0\|_{\rho_0}, r, |V_\lambda|_{C^2, \mathcal{B}_r}, \left\| \frac{\partial V_\lambda}{\partial \lambda} \Big|_{\lambda=\lambda_0} \circ K_0 \right\|_{\rho_0}$, and $|\{\text{avg}(\Lambda_0)\}^{-1}|$ (see (4.32)), such that if $0 < \delta_0 < \max\{1, \frac{\rho_0}{12}\}$ and the error e_0 satisfies

$$\|e_0\|_{\rho_0} \leq \min\{\gamma^4 \delta_0^{4\sigma}, c r \gamma^2 \delta_0^{2\sigma} \|e_0\|_{\rho_0}\},$$

then there exists a mapping $K \in \mathcal{W}_{\rho_0 - 6\delta_0}$ and a vector $\bar{\lambda} \in \mathbb{R}^{d+2n}$ such that (2.16) (or, equivalently, (2.17)) is satisfied. Moreover, the following inequalities hold:

$$\|K - K_0\|_{\rho_0 - 6\delta_0} < \frac{1}{c} \gamma^2 \delta_0^{-2\sigma} \|e_0\|_{\rho_0}, \quad |\bar{\lambda} - \lambda_0| < \frac{1}{c} \gamma^2 \delta_0^{-2\sigma} \|e_0\|_{\rho_0}.$$

3. EXACT SOLUTIONS

In this section we will assume that we know the exact solution of the problem and will use the geometry of the problem to construct bases of TK with special properties that will be utilized in Section 4.

3.1. Invariant tori are isotropic.

Definition 3.1. Let $\{V_\lambda\}$ be a $(d + 2n)$ -parameter family of presymplectic vector fields on the exact presymplectic manifold (\mathcal{P}, Ω) , and $\omega \in \mathbb{R}^{d+n}$ be a Diophantine vector. If for some value $\bar{\lambda}$ of λ there exists an embedding $K : \mathbb{T}^{d+n} \rightarrow \mathcal{P}$ such that (2.17) holds, we call K and \mathcal{K} an *invariant torus (KAM torus)* or a *true solution*.

This terminology is somewhat of a misnomer. We require more than \mathcal{K} being merely invariant under the flow of $V_{\bar{\lambda}}$ – we want that the motion on \mathcal{K} be quasi-periodic, (i.e., we require that the dynamics on \mathcal{K} be conjugate to a linear flow on \mathbb{T}^{d+n} with frequency ω independent over the rationals).

Notational convention. In Section 3 we assume that $\lambda = \bar{\lambda}$, and set $V := V_{\bar{\lambda}}$.

Definition 3.2. An invariant (in the sense of Definition 3.1) torus, $\mathcal{K} = K(\mathbb{T}^{d+n})$, in the presymplectic manifold (\mathcal{P}, Ω) is said to be *isotropic* if the pull-back, $K^*\Omega \in \Omega^2(\mathbb{T}^{d+n})$, of $\Omega \in \Omega^2(\mathcal{P})$ to the torus \mathbb{T}^{d+n} vanishes identically.

In Lemma 3.3 below we will prove that an invariant torus is isotropic. Similar results for maps are well-known for the case of submanifolds invariant with respect to symplectic or presymplectic maps (see, e.g., [21, Section 4, Lemma 1] or [3, Lemma 2.5]). This fact is crucial in the proof of Lemma 4.2 which, in turn, is essential for the bounds needed to solve the linearized equation in Section 4.3.

We introduce the linear operator $L_\theta : T_\theta\mathbb{T}^{d+n} \rightarrow T_\theta\mathbb{T}^{d+n}$ as the matrix representation of the pull-back $(K^*\Omega)_\theta$: for $U_\theta, W_\theta \in T_\theta\mathbb{T}^{d+n}$,

$$\langle U_\theta, L_\theta W_\theta \rangle_{\mathbb{R}^{d+n}} := (K^*\Omega)_\theta(U_\theta, W_\theta) = \Omega_{K(\theta)}(K_{*\theta} U_\theta, K_{*\theta} W_\theta). \tag{3.1}$$

The explicit expression for the matrix elements of L_θ is

$$L_\theta = DK_\theta^\top J_{K(\theta)} DK_\theta \in M_{d+n, d+n}(\mathbb{R}). \tag{3.2}$$

Lemma 3.3. *Let (\mathcal{P}, Ω) be an exact presymplectic manifold, $V \in \mathfrak{X}(\mathcal{P})$ be presymplectic, and $K : \mathbb{T}^{d+n} \rightarrow \mathcal{P}$ be a true solution. Then the invariant torus $\mathcal{K} = K(\mathbb{T}^{d+n})$ is isotropic (i.e., $K^*\Omega$ and, hence, L_θ , vanish identically).*

Proof. We prove the lemma in two steps: first we use the exactness of Ω to show that the average (over \mathbb{T}^{d+n}) of each matrix element of L is zero, and then we use the ergodicity of the flow $\theta \mapsto \theta + t\omega$ on \mathbb{T}^{d+n} to demonstrate that $K^*\Omega$ and, therefore, L , are constant on \mathbb{T}^{d+n} .

Since the presymplectic form Ω is exact, there exists a 1-form $\tau \in \Omega^1(\mathcal{P})$ such that $\Omega = d\tau$, hence $K^*\Omega = K^*(d\tau) = d(K^*\tau)$. If

$$\tau_{K(\theta)} = \sum_{A=1}^{d+2n} \tau_A(K(\theta)) dx^A,$$

then the pull-back $K^*\tau \in \Omega^1(\mathbb{T}^{d+n})$ is given by

$$(K^*\tau)_\theta = \sum_{\alpha=1}^{d+n} C_\alpha(\theta) d\theta^\alpha, \quad C_\alpha(\theta) := \sum_{A=1}^{d+2n} \tau_A(K(\theta)) \frac{\partial K^A}{\partial \theta^\alpha}(\theta),$$

and the matrix representation of $(K^*\Omega)_\theta$ is

$$(L_\theta)^\alpha_\beta = \frac{\partial C_\alpha}{\partial \theta^\beta}(\theta) - \frac{\partial C_\beta}{\partial \theta^\alpha}(\theta).$$

Because of the periodicity of the functions $C_\alpha : \mathbb{T}^{d+n} \rightarrow \mathbb{R}$,

$$\begin{aligned} \text{avg} \left(\frac{\partial C_\alpha}{\partial \theta^\beta} \right) &:= \int_{\mathbb{T}^{d+n}} \frac{\partial C_\alpha}{\partial \theta^\beta}(\theta) \, d\theta^1 \, d\theta^2 \cdots d\theta^{d+n} \\ &= \int_{\mathbb{T}^{d+n-1}} \left(\int_{\mathbb{T}^1} \frac{\partial C_\alpha}{\partial \theta^\beta}(\theta) \, d\theta^\beta \right) d\theta^1 \cdots \widehat{d\theta^\beta} \cdots d\theta^{d+n} = 0 \end{aligned} \tag{3.3}$$

(the term $\widehat{d\theta^\beta}$ is missing), so $\text{avg}(L) = 0$ and, hence, $\text{avg}(K^*\Omega) = 0$.

Now we prove that L and $K^*\Omega$ are constant on \mathbb{T}^{d+n} . Restrict the target space of the map K (2.14) from \mathcal{P} to the image \mathcal{K} of K , to obtain the diffeomorphism $K_r := K|_{\mathbb{T}^{d+n} \rightarrow \mathcal{K}} : \mathbb{T}^{d+n} \rightarrow \mathcal{K}$. Since \mathcal{K} is invariant with respect to the flow of $V \in \mathfrak{X}(\mathcal{P})$, at each $k \in \mathcal{K}$, $V_k \in T_k\mathcal{K}$. Therefore, the restriction $V|_{\mathcal{K}}$ of V to \mathcal{K} can be considered as a vector field on \mathcal{K} : $V|_{\mathcal{K}} \in \mathfrak{X}(\mathcal{K})$. For the same reason, the Lie derivative with respect to V has a natural restriction to sections of any tensor power of the tangent and cotangent bundles of \mathcal{K} . The pull-back of $V|_{\mathcal{K}}$ by the diffeomorphism K_r is $K_r^*V := (K_r^{-1})_* (V|_{\mathcal{K}}) \in \mathfrak{X}(\mathbb{T}^{d+n})$. If we consider the constant $\omega \in \mathbb{R}^{d+n}$ as a tangent vector $\omega_\theta \in T_\theta(\mathbb{T}^{d+n})$, then the pull-back of $V_{K(\theta)} = K_{*\theta} \omega_\theta \in T_{K(\theta)}\mathcal{K}$ is

$$(K_r^*V)_\theta = (K_r^{-1})_{*K(\theta)} K_{*\theta} \omega_\theta = (K_r^{-1} \circ K)_{*\theta} \omega_\theta = \omega_\theta.$$

By a well-known property of the Lie derivative, $K^*\mathcal{L}_V\Omega = \mathcal{L}_{K_r^*V}K^*\Omega = \mathcal{L}_\omega K^*\Omega$ (where all objects and operations are restricted to \mathcal{K}). Since V is presymplectic, $\mathcal{L}_V\Omega = 0$, which implies that its pull-back $K^*\Omega$ to \mathbb{T}^{d+n} is constant on the orbits of the flow $\theta \mapsto \theta + t\omega$, $t \in \mathbb{R}$. But ω is Diophantine, hence this flow is ergodic on \mathbb{T}^{d+n} , therefore $K^*\Omega = \text{const}$ and $L = \text{const}$. This together with the fact that $\text{avg}(L) = 0$ and $\text{avg}(K^*\Omega) = 0$ implies the desired result. \square

3.2. Construction and properties of an adapted basis of $T_{K(\theta)}\mathcal{P}$. In this section we construct a basis of $(T\mathcal{P})|_{\mathcal{K}}$ that is *adapted* to the invariant (with respect to the flow of the presymplectic vector field V) filtration $(\ker \Omega)|_{\mathcal{K}} \subseteq T\mathcal{K} \subseteq (T\mathcal{P})|_{\mathcal{K}}$ of vector bundles over \mathcal{K} .

3.2.1. Adapted coordinates in \mathbb{T}^{d+n} and a basis of $T_{K(\theta)}\mathcal{K}$. We first construct a basis of $T_{K(\theta)}\mathcal{K}$ (at an arbitrary point $K(\theta) \in \mathcal{K}$) as a push-forward $\{K_{*\theta} \left(\frac{\partial}{\partial \theta^\alpha}\right)_\theta\}_{\alpha=1}^{d+n}$ of the basis $\left\{\left(\frac{\partial}{\partial \theta^\alpha}\right)_\theta\right\}_{\alpha=1}^{d+n}$ of $T_\theta\mathbb{T}^{d+n}$.

Every vector in $T_{K(\theta)}\mathcal{K}$ has the form $K_{*\theta}U_\theta$ for some $U_\theta \in T_\theta\mathbb{T}^{d+n}$. The components $(K_{*\theta}U_\theta)^A$ of $K_{*\theta}U_\theta$ in the basis $\left\{\left(\frac{\partial}{\partial x^A}\right)_{K(\theta)}\right\}_{A=1}^{d+2n}$ are related to the components U_θ^α of U_θ in the basis $\left\{\left(\frac{\partial}{\partial \theta^\alpha}\right)_\theta\right\}_{\alpha=1}^{d+n}$ by

$$(K_{*\theta}U_\theta)^A = \sum_{\alpha=1}^{d+n} \frac{\partial K^A}{\partial \theta^\alpha}(\theta) U_\theta^\alpha. \tag{3.4}$$

If we think of $(K_{*\theta}U_\theta)^A$ as a column vector $K_{*\theta}U_\theta = [(K_{*\theta}U_\theta)^1 \cdots (K_{*\theta}U_\theta)^{d+2n}]^\top$ and of the $(d+n)$ columns of the matrix

$$DK_\theta = \left[\frac{\partial K^A}{\partial \theta^\alpha}(\theta)\right] = \left[\frac{\partial K}{\partial \theta^1}(\theta) \cdots \frac{\partial K}{\partial \theta^{d+n}}(\theta)\right] \in M_{d+2n, d+n}(\mathbb{R}), \tag{3.5}$$

as $(d + 2n)$ -dimensional vectors, then we can interpret (3.4) as expressing the arbitrary vector $K_{*\theta} \mathbf{U}_\theta \in T_{K(\theta)}\mathcal{K}$ as a linear combination of $\frac{\partial K}{\partial \theta^1}(\theta), \dots, \frac{\partial K}{\partial \theta^{d+n}}(\theta)$. Therefore, the column vectors $\left\{ \frac{\partial K}{\partial \theta^\alpha}(\theta) \right\}_{\alpha=1}^{d+n}$ form a basis of $T_{K(\theta)}\mathcal{K}$.

To adapt the basis $\left\{ \frac{\partial K}{\partial \theta^\alpha}(\theta) \right\}_{\alpha=1}^{d+n}$ to the special subspace $\ker \Omega_{K(\theta)} \subseteq T_{K(\theta)}\mathcal{K}$, recall that the first d coordinates in $\mathcal{P} = \mathbb{T}^d \times T^*\mathbb{T}^n$ (2.2) correspond to $\ker \Omega$. Because of this, we reorder the coordinates $\theta^\alpha, \alpha = 1, \dots, d + n$, of \mathbb{T}^{d+n} in such a way that the $d \times d$ block in the upper left corner of DK_θ (3.5) has full rank (then the $2n \times n$ block in the lower right corner of DK_θ will also be of full rank). Then $\left\{ \frac{\partial K}{\partial \theta^\mu}(\theta) \right\}_{\mu=1}^d$ is a basis of $\ker \Omega_{K(\theta)}$, while the vectors $\frac{\partial K}{\partial \theta^{d+1}}(\theta), \dots, \frac{\partial K}{\partial \theta^{d+n}}(\theta)$ span an n -dimensional subspace that is transversal to $\ker \Omega_{K(\theta)}$ in $T_{K(\theta)}\mathcal{K}$.

Define the matrices $Z_\theta \in M_{d+2n,d}(\mathbb{R})$ and $X_\theta \in M_{d+2n,n}(\mathbb{R})$ as the first d , respectively the last n , columns of DK_θ (3.5), so that $DK_\theta = \begin{bmatrix} Z_\theta & X_\theta \end{bmatrix}$. Recall the notation from Table 1 for the coordinates $x = (x^A) = (\underline{x}, \tilde{x})$, in \mathcal{P} (2.2):

$$\underline{x} = (\underline{x}^\mu) = (x^1, \dots, x^d), \quad \tilde{x} = (\tilde{x}^i) = (\tilde{x}^1, \dots, \tilde{x}^{2n}) = (x^{d+1}, \dots, x^{d+2n}).$$

We will use these notation in matrices with $d + 2n$ rows – the underscore for the first d rows, and the tilde for the remaining $2n$ rows:

$$Z_\theta = \begin{bmatrix} \underline{Z}_\theta \\ \tilde{Z}_\theta \end{bmatrix}, \quad X_\theta = \begin{bmatrix} \underline{X}_\theta \\ \tilde{X}_\theta \end{bmatrix}, \quad DK_\theta = \begin{bmatrix} Z_\theta & X_\theta \end{bmatrix} = \begin{bmatrix} \underline{Z}_\theta & \underline{X}_\theta \\ \tilde{Z}_\theta & \tilde{X}_\theta \end{bmatrix}. \quad (3.6)$$

3.2.2. *Adapted basis of $T_{K(\theta)}\mathcal{P}$.* Having constructed a basis of $T_{K(\theta)}\mathcal{K}$, we need n more vectors that span the complement of $T_{K(\theta)}\mathcal{K}$ in $T_{K(\theta)}\mathcal{P}$. We will construct them in such a way that, together with the columns $\frac{\partial K}{\partial \theta^{d+1}}(\theta), \dots, \frac{\partial K}{\partial \theta^{d+n}}(\theta)$ of X_θ , they form a symplectic basis of $T_{K(\theta)}\mathcal{Q}$ (2.3). To this end we will use the matrix representation $\tilde{J}_{K(\theta)}$ (2.5) of the symplectic form $\tilde{\Omega}$ on \mathcal{Q} , as well as the Gramian matrix of the vectors $\frac{\partial K}{\partial \theta^{d+1}}(\theta), \dots, \frac{\partial K}{\partial \theta^{d+n}}(\theta)$, i.e., the matrix $X_\theta^\top X_\theta \in M_{n,n}(\mathbb{R})$ of their inner products with respect to the Euclidean inner product on $T^*\mathbb{T}^n$. Define

$$R_\theta := (\tilde{X}_\theta^\top \tilde{X}_\theta)^{-1} \in M_{n,n}(\mathbb{R}), \quad (3.7)$$

$$\tilde{Y}_\theta := \tilde{J}_{K(\theta)}^{-1} \tilde{X}_\theta R_\theta \in M_{2n,n}(\mathbb{R}), \quad (3.8)$$

$$Y_\theta := \begin{bmatrix} 0 \\ \tilde{Y}_\theta \end{bmatrix} = \begin{bmatrix} 0 \\ \tilde{J}_{K(\theta)}^{-1} \tilde{X}_\theta R_\theta \end{bmatrix} \in M_{d+2n,n}(\mathbb{R}). \quad (3.9)$$

Since $\tilde{J}_{K(\theta)}^{-1}$, \tilde{X}_θ , and R_θ are of maximal rank, \tilde{Y}_θ and Y_θ are of maximal rank: $\text{rank } \tilde{Y}_\theta = \text{rank } Y_\theta = n$. We think of the n columns of Y_θ as of as vectors in $T_{K(\theta)}\mathcal{P}$.

Let $Z_{\theta,\mu}, X_{\theta,a}, Y_{\theta,a}$, with $\mu = 1, \dots, d, a = 1, \dots, n$, stand for the columns of Z_θ, X_θ , and Y_θ . These $d + 2n$ vectors are a basis of $T_{K(\theta)}\mathcal{P}$ with the properties

$$\begin{aligned} \text{span}\{Z_{\theta,\mu}\}_{\mu=1}^d &= \ker \Omega_{K(\theta)}, \\ \text{span}\left\{ \{Z_{\theta,\mu}\}_{\mu=1}^d, \{X_{\theta,a}\}_{a=1}^n \right\} &= T_{K(\theta)}\mathcal{K}. \end{aligned}$$

The first property implies that $\Omega_{K(\theta)}(Z_{\theta,\mu}, \cdot) = 0$. The construction of $Y_{\theta,a}$ yields (using (2.4), (2.5), (2.6), (3.7), (3.8), and (3.9))

$$\Omega_{K(\theta)}(X_{\theta,a}, Y_{\theta,b}) = \langle X_{\theta,a}, J_{K(\theta)} Y_{\theta,b} \rangle_{\mathbb{R}^{d+2n}} = X_{\theta,a}^\top J_{K(\theta)} Y_{\theta,b} = (X_\theta^\top J_{K(\theta)} Y_\theta)_{ab}$$

$$= \left(\left[\begin{array}{c} \underline{X}_\theta^\top \tilde{X}_\theta^\top \\ \tilde{X}_\theta R_\theta \end{array} \right] \begin{bmatrix} 0 \\ \tilde{X}_\theta R_\theta \end{bmatrix} \right)_{ab} = (\tilde{X}_\theta^\top \tilde{X}_\theta R_\theta)_{ab} = (\mathbb{I}_n)_{ab} = \delta_{ab}$$

($\mathbb{I}_n = (\delta_{ab})$ is the unit $n \times n$ matrix), hence the vectors $X_{\theta,a}$ and $Y_{\theta,b}$ form a symplectic basis of $T_{K(\theta)}\mathcal{Q} \cong (T_{K(\theta)}\mathcal{P})/\ker \Omega_{K(\theta)}$. We write this symbolically as

$$\Omega_{K(\theta)}(X_\theta, Y_\theta) = X_\theta^\top J_{K(\theta)} Y_\theta = \tilde{X}_\theta^\top \tilde{J}_{K(\theta)} \tilde{Y}_\theta = \mathbb{I}_n. \tag{3.10}$$

Below we summarize the properties of the basis of $T_{K(\theta)}\mathcal{P}$ constructed above, using matrix notations as in (3.10):

$$\begin{aligned} Z_\theta^\top J_{K(\theta)} Z_\theta &= \tilde{Z}_\theta^\top \tilde{J}_{K(\theta)} \tilde{Z}_\theta = 0, \\ Z_\theta^\top J_{K(\theta)} X_\theta &= \tilde{Z}_\theta^\top \tilde{J}_{K(\theta)} \tilde{X}_\theta = 0, \\ Z_\theta^\top J_{K(\theta)} Y_\theta &= \tilde{Z}_\theta^\top \tilde{J}_{K(\theta)} \tilde{Y}_\theta = \tilde{Z}_\theta^\top \tilde{X}_\theta R_\theta, \\ X_\theta^\top J_{K(\theta)} X_\theta &= \tilde{X}_\theta^\top \tilde{J}_{K(\theta)} \tilde{X}_\theta = 0, \\ X_\theta^\top J_{K(\theta)} Y_\theta &= \tilde{X}_\theta^\top \tilde{J}_{K(\theta)} \tilde{Y}_\theta = \mathbb{I}_n, \\ Y_\theta^\top J_{K(\theta)} Y_\theta &= \tilde{Y}_\theta^\top \tilde{J}_{K(\theta)} \tilde{Y}_\theta = -R_\theta \tilde{X}_\theta^\top \tilde{J}_{K(\theta)}^{-1} \tilde{X}_\theta R_\theta. \end{aligned} \tag{3.11}$$

3.2.3. *Presymplecticity of V at \mathcal{K} in adapted coordinates.* If $U, V, W \in \mathfrak{X}(\mathcal{P})$ with V presymplectic, i.e., $\mathcal{L}_V \Omega = 0$ (recall Lemma 2.3), then

$$\begin{aligned} 0 &= (\mathcal{L}_V \Omega)(U, W) = \mathcal{L}_V(\Omega(U, W)) - \Omega(\mathcal{L}_V U, W) - \Omega(U, \mathcal{L}_V W) \\ &= \sum_{A,B,C=1}^{d+2n} U^A \left(\frac{\partial \Omega_{AB}}{\partial x^C} V^C + \frac{\partial V^C}{\partial x^A} \Omega_{CB} + \Omega_{AC} \frac{\partial V^C}{\partial x^B} \right) W^B \\ &= \sum_{A,B=1}^{d+2n} U_A \left((DJ)V + (DV)^\top J + JDV \right)^A{}_B W^B, \end{aligned}$$

where we used the operator $J = (J^A{}_B)$ (2.4), lowering an index of a vector signifies transposition (i.e., contraction with the Euclidean metric tensor), and

$$(DV)^C{}_B = \frac{\partial V^C}{\partial x^B}, \quad ((DJ)V)^A{}_B := \sum_{C=1}^{d+2n} \frac{\partial J^A{}_B}{\partial x^C} V^C.$$

Therefore in matrix notation the presymplecticity condition reads

$$(DJ)V + (DV)^\top J + JDV = 0. \tag{3.12}$$

Writing the derivative of V at $K(\theta) \in \mathcal{K}$ as

$$DV_{K(\theta)} =: \begin{bmatrix} \frac{\partial V}{\partial \underline{x}} & \frac{\partial V}{\partial \tilde{x}} \\ \frac{\partial \tilde{V}}{\partial \underline{x}} & \frac{\partial \tilde{V}}{\partial \tilde{x}} \end{bmatrix}_{K(\theta)}, \tag{3.13}$$

we can easily show that (3.12) is equivalent to the conditions

$$\left[\frac{\partial \tilde{V}}{\partial \underline{x}} \right]_{K(\theta)} = 0, \quad (D\tilde{J})_{K(\theta)} V_{K(\theta)} + \left[\frac{\partial \tilde{V}}{\partial \tilde{x}} \right]_{K(\theta)}^\top \tilde{J}_{K(\theta)} + \tilde{J}_{K(\theta)} \left[\frac{\partial \tilde{V}}{\partial \tilde{x}} \right]_{K(\theta)} = 0. \tag{3.14}$$

3.3. Change of basis matrix M_θ .

3.3.1. *Definition of M_θ .* The adapted basis $\{Z_{\theta,\mu}\}_{\mu=1}^d, \{X_{\theta,a}\}_{a=1}^n, \{Y_{\theta,a}\}_{a=1}^n$ of $T_{K(\theta)}\mathcal{P}$ constructed in Section 3.2 has properties that are very useful for our analysis. Given an arbitrary column vector U_θ , considered as an element of $T_{K(\theta)}\mathcal{P}$, we can find its components in the adapted basis as follows. Define the *change of basis matrix* M_θ of all vectors from the adapted basis, written as column vectors:

$$M_\theta := [DK_\theta \ Y_\theta] = [Z_\theta \ X_\theta \ Y_\theta] = \begin{bmatrix} Z_\theta & X_\theta & 0 \\ \tilde{Z}_\theta & \tilde{X}_\theta & \tilde{Y}_\theta \end{bmatrix} \in M_{d+2n,d+2n}(\mathbb{R}). \tag{3.15}$$

Then the vector U_θ can be written as a superposition of the vectors from the adapted basis as follows:

$$U_\theta = M_\theta \xi_\theta = \sum_{\mu=1}^d Z_{\theta,\mu} \xi_\theta^\mu + \sum_{a=1}^n X_{\theta,a} \xi_\theta^{d+a} + \sum_{a=1}^n Y_{\theta,a} \xi_\theta^{d+n+a}. \tag{3.16}$$

In the adapted basis, if we write the $(d + 2n)$ components of the vector ξ_θ as three blocks of length d , n , and n , as in (3.16), then the vectors from $T_{K(\theta)}\mathcal{K}$ have the form $\xi_\theta = [* * 0]^\top$, where the stars represent numbers that are generally non-zero.

3.3.2. *Computing $(DV_{K(\theta)} - \partial_\omega)M_\theta$.* In this section we will perform some computations related to the change of basis matrix M_θ (3.15), which will be needed in Section 4. Differentiating the invariance condition (2.17), we obtain

$$DV_{K(\theta)}DK_\theta = \partial_\omega DK_\theta. \tag{3.17}$$

This, together with the definition (3.15) of M_θ , gives us

$$(DV_{K(\theta)} - \partial_\omega)M_\theta = (DV_{K(\theta)} - \partial_\omega)[DK_\theta \ Y_\theta] = \begin{bmatrix} 0 & 0 & (DV_{K(\theta)} - \partial_\omega)Y_\theta \end{bmatrix}.$$

Our first goal is to find an explicit expression for $(DV_{K(\theta)} - \partial_\omega)Y_\theta$. To this end we have to compute

$$(DV_{K(\theta)} - \partial_\omega)Y_\theta = \begin{bmatrix} [\frac{\partial V}{\partial \tilde{x}}]_{K(\theta)} \tilde{Y}_\theta \\ [\frac{\partial \tilde{V}}{\partial \tilde{x}}]_{K(\theta)} \tilde{Y}_\theta - \partial_\omega \tilde{Y}_\theta \end{bmatrix}.$$

By the Leibniz rule,

$$\partial_\omega \tilde{Y}_\theta = \partial_\omega (\tilde{J}_{K(\theta)}^{-1} \tilde{X}_\theta R_\theta) = \partial_\omega (\tilde{J}_{K(\theta)}^{-1}) \tilde{X}_\theta R_\theta + \tilde{J}_{K(\theta)}^{-1} \partial_\omega (\tilde{X}_\theta) R_\theta + \tilde{J}_{K(\theta)}^{-1} \tilde{X}_\theta \partial_\omega R_\theta.$$

The elementary identity $0 = \partial_\omega (\mathbb{I}_{2n}) = \partial_\omega (\tilde{J}_{K(\theta)}^{-1} \tilde{J}_{K(\theta)})$, the invariance (2.17), and the presymplecticity condition (3.14) yield

$$\begin{aligned} \partial_\omega (\tilde{J}_{K(\theta)}^{-1}) &= -\tilde{J}_{K(\theta)}^{-1} \partial_\omega (\tilde{J}_{K(\theta)}) \tilde{J}_{K(\theta)}^{-1} = -\tilde{J}_{K(\theta)}^{-1} (D\tilde{J})_{K(\theta)} (\partial_\omega K_\theta) \tilde{J}_{K(\theta)}^{-1} \\ &= -(\tilde{J}^{-1} (D\tilde{J}) V \tilde{J}^{-1})_{K(\theta)} = -\left(\tilde{J} \left[\frac{\partial \tilde{V}}{\partial \tilde{x}} \right] + \left[\frac{\partial \tilde{V}}{\partial \tilde{x}} \right]^\top \tilde{J} \right)_{K(\theta)}. \end{aligned} \tag{3.18}$$

The invariance (3.17) and the expressions (3.13) and (3.14) give us

$$\partial_\omega \tilde{X}_\theta = \left[\frac{\partial \tilde{V}}{\partial \tilde{x}} \right]_{K(\theta)} \tilde{X}_\theta. \tag{3.19}$$

From the definition (3.7) of R_θ , the expression (3.19) for $\partial_\omega \tilde{X}_\theta$, we easily obtain

$$\partial_\omega R_\theta = -2R_\theta \tilde{X}_\theta^\top \left[\frac{\partial \tilde{V}}{\partial \tilde{x}} \right]_{K(\theta)}^{\text{sym}} \tilde{X}_\theta R_\theta, \tag{3.20}$$

where

$$\left[\frac{\partial \tilde{V}}{\partial \tilde{x}}\right]_{K(\theta)}^{\text{sym}} := \frac{1}{2} \left(\left[\frac{\partial \tilde{V}}{\partial \tilde{x}}\right]_{K(\theta)} + \left[\frac{\partial \tilde{V}}{\partial \tilde{x}}\right]_{K(\theta)}^\top \right).$$

It will be convenient to introduce the operator

$$\tilde{\Pi}_\theta := \mathbb{I}_{2n} - \tilde{X}_\theta R_\theta \tilde{X}_\theta^\top : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}. \tag{3.21}$$

We collect some properties of $\tilde{\Pi}_\theta$ and $\tilde{J}_{K(\theta)}^{-1} \tilde{\Pi}_\theta \tilde{J}_{K(\theta)} = \mathbb{I}_{2n} - \tilde{Y}_\theta \tilde{X}_\theta^\top \tilde{J}_{K(\theta)}$ which follow easily from (3.7), (3.8), and (3.11):

- $\tilde{\Pi}_\theta$ is symmetric;
- both $\tilde{\Pi}_\theta$ and $\tilde{J}_{K(\theta)}^{-1} \tilde{\Pi}_\theta \tilde{J}_{K(\theta)}$ are idempotent;
- the n columns of the matrix \tilde{X}_θ are in the kernel of $\tilde{\Pi}_\theta$;
- the n columns of \tilde{X}_θ are eigenvectors of $\tilde{J}_{K(\theta)}^{-1} \tilde{\Pi}_\theta \tilde{J}_{K(\theta)}$ with eigenvalue 1, while the n columns of \tilde{Y}_θ are in the kernel of $\tilde{J}_{K(\theta)}^{-1} \tilde{\Pi}_\theta \tilde{J}_{K(\theta)}$:

$$\tilde{J}_{K(\theta)}^{-1} \tilde{\Pi}_\theta \tilde{J}_{K(\theta)} \tilde{X}_\theta = \tilde{X}_\theta, \quad \tilde{J}_{K(\theta)}^{-1} \tilde{\Pi}_\theta \tilde{J}_{K(\theta)} \tilde{Y}_\theta = 0, \tag{3.22}$$

hence $\tilde{J}_{K(\theta)}^{-1} \tilde{\Pi}_\theta \tilde{J}_{K(\theta)}$ and $(\mathbb{I}_{2n} - \tilde{J}_{K(\theta)}^{-1} \tilde{\Pi}_\theta \tilde{J}_{K(\theta)})$ are projection operators corresponding to the splitting

$$T_{K(\theta)} \mathcal{Q} = \text{span}\{\tilde{X}_{\theta,a}\}_{a=1}^n \oplus \text{span}\{\tilde{Y}_{\theta,a}\}_{a=1}^n.$$

Putting together (3.18), (3.19), and (3.20), and using the definition (3.21), we obtain

$$\partial_\omega \tilde{Y}_\theta = \left[\frac{\partial \tilde{V}}{\partial \tilde{x}}\right]_{K(\theta)} \tilde{Y}_\theta + 2 \tilde{J}_{K(\theta)}^{-1} \tilde{\Pi}_\theta \left[\frac{\partial \tilde{V}}{\partial \tilde{x}}\right]_{K(\theta)}^{\text{sym}} \tilde{X}_\theta R_\theta \tag{3.23}$$

so, finally,

$$(DV_{K(\theta)} - \partial_\omega) M_\theta = \begin{bmatrix} 0 & 0 & \left[\frac{\partial \tilde{V}}{\partial \tilde{x}}\right]_{K(\theta)} \tilde{Y}_\theta \\ 0 & 0 & -2 \tilde{J}_{K(\theta)}^{-1} \tilde{\Pi}_\theta \left[\frac{\partial \tilde{V}}{\partial \tilde{x}}\right]_{K(\theta)}^{\text{sym}} \tilde{X}_\theta R_\theta \end{bmatrix}. \tag{3.24}$$

3.3.3. *Writing $(DV_{K(\theta)} - \partial_\omega)M_\theta$ as $M_\theta C_\theta$.* Having computed $(DV_{K(\theta)} - \partial_\omega)M_\theta$, we will rewrite it in a form that plays a crucial role in Section 4:

$$(DV_{K(\theta)} - \partial_\omega)M_\theta = M_\theta C_\theta := M_\theta \begin{bmatrix} 0 & 0 & T_\theta \\ 0 & 0 & S_\theta \\ 0 & 0 & U_\theta \end{bmatrix}. \tag{3.25}$$

Substitute (3.24) and (3.15) in (3.25) to obtain that T_θ , S_θ , and U_θ should satisfy

$$\underline{Z}_\theta T_\theta + \underline{X}_\theta S_\theta = \left[\frac{\partial \tilde{V}}{\partial \tilde{x}}\right]_{K(\theta)} \tilde{J}_{K(\theta)}^{-1} \tilde{X}_\theta R_\theta, \tag{3.26}$$

$$\tilde{Z}_\theta T_\theta + \tilde{X}_\theta S_\theta + \tilde{Y}_\theta U_\theta = -2 \tilde{J}_{K(\theta)}^{-1} \tilde{\Pi}_\theta \left[\frac{\partial \tilde{V}}{\partial \tilde{x}}\right]_{K(\theta)}^{\text{sym}} \tilde{X}_\theta R_\theta. \tag{3.27}$$

Multiplying (3.27) separately by $\tilde{X}_\theta^\top \tilde{J}_{K(\theta)}$ and $\tilde{Y}_\theta^\top \tilde{J}_{K(\theta)}$ on the left and using (3.11) and the definition of R_θ (3.7), we obtain

$$U_\theta = -2 \tilde{X}_\theta^\top \tilde{\Pi}_\theta \left[\frac{\partial \tilde{V}}{\partial \tilde{x}}\right]_{K(\theta)}^{\text{sym}} \tilde{X}_\theta R_\theta, \tag{3.28}$$

$$-R_\theta \tilde{X}_\theta^\top \tilde{Z}_\theta T_\theta - S_\theta - R_\theta \tilde{X}_\theta^\top \tilde{J}_{K(\theta)}^{-1} \tilde{X}_\theta R_\theta U_\theta = -2 \tilde{Y}_\theta^\top \tilde{\Pi}_\theta \left[\frac{\partial \tilde{V}}{\partial \tilde{x}}\right]_{K(\theta)}^{\text{sym}} \tilde{X}_\theta R_\theta. \tag{3.29}$$

Since $\tilde{X}_\theta^\top \tilde{\Pi}_\theta = (\tilde{\Pi}_\theta \tilde{X}_\theta)^\top = 0$, (3.28) yields $U_\theta = 0$. From (3.26) and (3.29) we obtain

$$\begin{aligned} T_\theta &= \mathcal{Z}_\theta^{-1} \left(\left[\frac{\partial V}{\partial \tilde{x}} \right]_{K(\theta)} \tilde{Y}_\theta - 2 \underline{X}_\theta \tilde{Y}_\theta^\top \tilde{\Pi}_\theta \left[\frac{\partial \tilde{V}}{\partial \tilde{x}} \right]_{K(\theta)}^{\text{sym}} \tilde{X}_\theta R_\theta \right), \\ S_\theta &= -\tilde{Y}_\theta^\top \tilde{J}_{K(\theta)} \tilde{Z}_\theta \mathcal{Z}_\theta^{-1} \left[\frac{\partial V}{\partial \tilde{x}} \right]_{K(\theta)} \tilde{Y}_\theta \\ &\quad + 2 \left(\mathbb{I}_n - \tilde{Y}_\theta^\top \tilde{J}_{K(\theta)} \tilde{Z}_\theta \mathcal{Z}_\theta^{-1} \underline{X}_\theta \right) \tilde{Y}_\theta^\top \tilde{\Pi}_\theta \left[\frac{\partial \tilde{V}}{\partial \tilde{x}} \right]_{K(\theta)}^{\text{sym}} \tilde{X}_\theta R_\theta, \end{aligned} \tag{3.30}$$

where we have set

$$\mathcal{Z}_\theta := \underline{Z}_\theta - \underline{X}_\theta R_\theta \tilde{X}_\theta^\top \tilde{Z}_\theta = \underline{Z}_\theta + \underline{X}_\theta \tilde{Y}_\theta^\top \tilde{J}_{K(\theta)} \tilde{Z}_\theta \in M_{d,d}(\mathbb{R}). \tag{3.31}$$

The geometric meaning of \mathcal{Z}_θ is discussed in detail in [9, Section 3.5.4]. We summarize our findings in the following lemma.

Lemma 3.4. *Let (\mathcal{P}, Ω) be an exact presymplectic manifold, $V \in \mathfrak{X}(\mathcal{P})$ be a presymplectic vector field, $K : \mathbb{T}^{d+n} \rightarrow \mathcal{P}$ be an invariant torus in the sense of Definition 3.1, and the matrix M_θ be defined by (3.15). Then the equality (3.25) holds, with $U_\theta = 0$ and T_θ and S_θ given by (3.30).*

3.3.4. *Factorization of M_θ .* Later we will need the representation of M_θ^{-1} that follows from the lemma below.

Lemma 3.5. *If the $(d + 2n) \times (d + 2n)$ matrices Q_θ and W_θ are defined by*

$$Q_\theta := \begin{pmatrix} \mathbb{I}_d & 0 \\ 0 & \tilde{X}_\theta^\top \\ 0 & \tilde{Y}_\theta^\top \end{pmatrix} \begin{pmatrix} \mathbb{I}_d & 0 \\ 0 & \tilde{J}_{K(\theta)} \end{pmatrix} = \begin{pmatrix} \mathbb{I}_d & 0 \\ 0 & \tilde{X}_\theta^\top \tilde{J}_{K(\theta)} \\ 0 & \tilde{Y}_\theta^\top \tilde{J}_{K(\theta)} \end{pmatrix}, \tag{3.32}$$

$$W_\theta = \begin{pmatrix} \underline{Z}_\theta & \underline{X}_\theta & 0 \\ 0 & 0 & \mathbb{I}_n \\ \tilde{Y}_\theta^\top \tilde{J}_{K(\theta)} \tilde{Z}_\theta & -\mathbb{I}_n & \tilde{Y}_\theta^\top \tilde{J}_{K(\theta)} \tilde{Y}_\theta \end{pmatrix}, \tag{3.33}$$

then the following identity holds

$$Q_\theta M_\theta = W_\theta. \tag{3.34}$$

This implies that M_θ (3.15) is invertible if and only if W_θ is invertible.

Proof. The columns of \tilde{X}_θ and \tilde{Y}_θ form a (symplectic) basis of \mathbb{R}^{2n} , which implies that the rows of \tilde{X}_θ^\top and \tilde{Y}_θ^\top form a basis of \mathbb{R}^{2n} . Since $\tilde{J}_{K(\theta)}$ is an invertible matrix (it corresponds to the symplectic form $\tilde{\Omega}$ on \mathcal{Q} , recall (2.3) and (2.5)), the rows of $\tilde{X}_\theta^\top \tilde{J}_{K(\theta)}$ and $\tilde{Y}_\theta^\top \tilde{J}_{K(\theta)}$ form a basis of \mathbb{R}^{2n} , so that the matrix Q_θ given by (3.32) is invertible. The identity (3.34) follows directly from (3.11). \square

4. APPROXIMATE SOLUTIONS

In this section we examine the case when K is merely an approximate solution as defined below. We will build off of the results in Section 3 for true solutions to derive similar results for approximate solutions. We start with the definition for approximate solution.

Definition 4.1. Let (\mathcal{P}, Ω) be an exact presymplectic manifold, $\{V_\lambda\}$ be a $(d+2n)$ -parameter family of presymplectic vector fields, $\omega \in \mathcal{D}(\gamma, \sigma)$, and $K_0: \mathbb{T}^{d+n} \rightarrow \mathcal{P}$ be an embedding. For a value λ_0 of the parameter λ , define the error,

$$e_{0,\theta} := V_{\lambda_0, K_0(\theta)} - \partial_\omega K_{0,\theta} \in T_{K_0(\theta)}\mathcal{P} \cong \mathbb{R}^{d+2n}. \tag{4.1}$$

If some appropriately defined norm of e_0 is sufficiently small, then we say that K_0 is an *approximate solution*.

4.1. Approximately isotropic tori. In Lemma 3.3 we showed that if K_θ is a true solution (i.e., if (2.17) is satisfied), then the invariant manifold $\mathcal{K} = K(\mathbb{T}^{d+n})$ is isotropic, i.e., $K^*\Omega = 0$. The analogous result for this section will be that if $K_{0,\theta}$ is an approximate solution, then K_0 is *approximately isotropic*, i.e., $K_0^*\Omega$ is small.

Lemma 4.2. *Let (\mathcal{P}, Ω) be an exact presymplectic manifold, $\{V_\lambda\}$ be a $(d+2n)$ -parameter family of analytic presymplectic vector fields, and $K_0 \in \mathcal{W}_\rho$ (2.10) be an approximate solution with Diophantine frequency $\omega \in \mathcal{D}(\gamma, \sigma)$. Assume that V_λ extends holomorphically to some complex neighborhood \mathcal{B}_r (2.18) of the image of \mathbb{T}_ρ^{d+n} under K_0 , for some $r > 0$. Let $L_{0,\theta} : \mathbb{T}_\theta \mathbb{T}^{d+n} \rightarrow \mathbb{T}_\theta \mathbb{T}^{d+n}$ be the matrix representation of the pull-back $(K_0^*\Omega)_\theta$ as in (3.1) and (3.2):*

$$L_{0,\theta} = DK_{0,\theta}^\top J_{K_0(\theta)} DK_{0,\theta}. \tag{4.2}$$

Then there exists a constant $C > 0$ depending on $d, n, \sigma, \rho, \|DK_0\|_\rho, |V_{\lambda_0}|_{C^1, \mathcal{B}_r}$, and $|J|_{C^1, \mathcal{B}_r}$, such that for every δ satisfying $0 < \delta < \frac{\rho}{2}$, the following bound holds:

$$\|L_0\|_{\rho-2\delta} < C\gamma^{-1}\delta^{-(\sigma+1)}\|e_0\|_\rho. \tag{4.3}$$

Proof. For the directional derivative of $L_{0,\theta}$ (4.2), using (4.6), (4.1), and (3.12), we have

$$\begin{aligned} & \partial_\omega L_{0,\theta} \\ &= \partial_\omega (DK_{0,\theta}^\top J_{K_0(\theta)} DK_{0,\theta}) \\ &= \partial_\omega (DK_{0,\theta}^\top) J_{K_0(\theta)} DK_{0,\theta} + DK_{0,\theta}^\top \partial_\omega (J_{K_0(\theta)}) DK_{0,\theta} + DK_{0,\theta}^\top J_{K_0(\theta)} \partial_\omega (DK_{0,\theta}) \\ &= (DV_{\lambda_0, K_0(\theta)} DK_{0,\theta} - De_{0,\theta})^\top J_{K_0(\theta)} DK_{0,\theta} \\ &\quad + DK_{0,\theta}^\top DJ_{K_0(\theta)} (V_{\lambda_0, K_0(\theta)} - e_{0,\theta}) DK_{0,\theta} \\ &\quad + DK_{0,\theta}^\top J_{K_0(\theta)} (DV_{\lambda_0, K_0(\theta)} DK_{0,\theta} - De_{0,\theta}) \\ &= DK_{0,\theta}^\top \left(DV_{\lambda_0, K_0(\theta)}^\top J_{K_0(\theta)} + DJ_{K_0(\theta)} V_{\lambda_0, K_0(\theta)} + J_{K_0(\theta)} DV_{\lambda_0, K_0(\theta)} \right) DK_{0,\theta} \\ &\quad - (De_{0,\theta}^\top J_{K_0(\theta)} DK_{0,\theta} + DK_{0,\theta}^\top DJ_{K_0(\theta)} e_{0,\theta} DK_{0,\theta} + DK_{0,\theta}^\top J_{K_0(\theta)} De_{0,\theta}) \\ &= - (De_{0,\theta}^\top J_{K_0(\theta)} DK_{0,\theta} + DK_{0,\theta}^\top DJ_{K_0(\theta)} e_{0,\theta} DK_{0,\theta} + DK_{0,\theta}^\top J_{K_0(\theta)} De_{0,\theta}). \end{aligned}$$

From this and the Cauchy bound (2.12) we obtain

$$\|\partial_\omega L_0\|_{\rho-\delta} \leq C_1\|e_0\|_{\rho-\delta} + C_2\|De_0\|_{\rho-\delta} \leq C\delta^{-1}\|e_0\|_\rho. \tag{4.4}$$

Although K_0 is only an approximate solution, the exactness of Ω implies that $\text{avg}(L_0) = 0$ (the proof of this repeats part of the proof of Lemma 3.3, with K replaced by K_0). We apply (4.4) and the Rüssmann estimate (2.13) to obtain

$$\|L_0\|_{\rho-2\delta} \leq C\gamma^{-1}\delta^{-\sigma}\|\partial_\omega L_0\|_{\rho-\delta} \leq C\gamma^{-1}\delta^{-(\sigma+1)}\|e_0\|_\rho.$$

□

4.2. Linearized equation for the corrections. Given a family of presymplectic vector fields $\{V_\lambda\}$, the equation (2.17) can be difficult to solve for a value $\bar{\lambda}$ of the parameter and an embedding $K : \mathbb{T}^{d+n} \rightarrow \mathcal{P}$. So instead of solving it directly for λ and K , we start with an approximate solution (λ_0, K_0) and construct an iterative process that produces better approximate solutions. As a result, we obtain a sequence (λ_j, K_j) (1.2) that converges to $(\lambda_\infty, K_\infty) = (\bar{\lambda}, K)$.

Let (λ_j, K_j) be an approximate pair. Define the error (cf. (4.1))

$$e_{j,\theta} := V_{\lambda_j, K_j(\theta)} - \partial_\omega K_{j,\theta} \in T_{K_j(\theta)}\mathcal{P} \cong \mathbb{R}^{d+2n}. \tag{4.5}$$

We will usually consider e_j as a mapping $e_j : \mathbb{T}^{d+n} \rightarrow \mathbb{R}^{d+2n}$, whose derivative, $De_j : \mathbb{T}^{d+n} \rightarrow M_{d+2n, d+n}(\mathbb{R})$, is given by

$$De_{j,\theta} = (DV_{\lambda_j, K_j(\theta)} - \partial_\omega) DK_{j,\theta} \in M_{d+2n, d+n}(\mathbb{R}). \tag{4.6}$$

Note that in (4.6), $DV_{\lambda_j, K_j(\theta)}$ stands for the derivative of the vector field V_{λ_j} with respect to the spatial variables $x \in \mathbb{R}^{d+2n}$:

$$DV_{\lambda_j, K_j(\theta)} = [(DV_{\lambda_j, K_j(\theta)})^A{}_B] = \left[\frac{\partial V_{\lambda_j, x}^A}{\partial x^B} \Big|_{x=K_j(\theta)} \right] \in M_{d+2n, d+2n}(\mathbb{R}).$$

Recall that the presymplecticity of V_{λ_j} imply that $DV_{\lambda_j, K_j(\theta)}$ (3.13) satisfies (3.14).

Let ε_j and Δ_j be $(j + 1)$ st correction terms, i.e.,

$$\lambda_{j+1} := \lambda_j + \varepsilon_j, \quad K_{j+1,\theta} := K_{j,\theta} + \Delta_{j,\theta}. \tag{4.7}$$

To derive an equation for the corrections ε_j and $\Delta_{j,\theta}$, we set $\mathcal{F}[\lambda, K] := V_\lambda \circ K - \partial_\omega K$, so that $e_{j,\theta} = \mathcal{F}[\lambda_j, K_j](\theta)$. Expanding $\mathcal{F}[\lambda_{j+1}, K_{j+1}]$ about (λ_j, K_j) , we obtain

$$e_{j+1,\theta} = e_{j,\theta} + DV_{\lambda_j, K_j(\theta)} \Delta_{j,\theta} - \partial_\omega \Delta_{j,\theta} + \left[\frac{\partial V_\lambda}{\partial \lambda} \right]_{\lambda_j, K_j(\theta)} \varepsilon_j + \mathcal{O}(|\Delta_j, \varepsilon_j|^2). \tag{4.8}$$

Therefore, if the corrections ε_j and Δ_j satisfy the linear equation

$$(DV_{\lambda_j, K_j(\theta)} - \partial_\omega) \Delta_{j,\theta} = -e_{j,\theta} - \left[\frac{\partial V_\lambda}{\partial \lambda} \right]_{\lambda_j, K_j(\theta)} \varepsilon_j, \tag{4.9}$$

the cancelations in the right-hand side of (4.8) guarantee quadratic convergence.

The system (4.9) of $(d + 2n)$ equations for the corrections ε_j and Δ_j is a linear algebraic equation with respect to ε_j , and a linear first-order partial differential equation with respect to Δ_j . Since $DV_{\lambda_j, K_j(\theta)} \in M_{d+2n, d+2n}(\mathbb{R})$ is of a general form, is not easy to solve (4.9) and to obtain estimates on its solution. In Section 4.3 we will use the matrix M_θ (3.15) of change of basis from a general one to the adapted basis constructed in Section 3.2; the calculations from Section 3.3 will be very useful.

4.3. Solving the linearized equation.

4.3.1. Change of basis. We use an adapted basis in \mathbb{R}^{d+2n} , so that instead of the unknown function Δ_j we introduce the unknown function $\xi_j : \mathbb{T}^{d+n} \rightarrow \mathbb{R}^{d+2n}$ through the linear change of basis

$$\Delta_{j,\theta} =: M_{j,\theta} \xi_{j,\theta}. \tag{4.10}$$

The change of basis matrix $M_{j,\theta} \in M_{d+2n, d+2n}(\mathbb{R})$ is constructed similarly to the matrix M_θ in (3.15), but by using the approximate value λ_j and the approximate

embedding K_j . Namely, given an approximate invariant torus K_j (which we treat as a map $K_j: \mathbb{T}^{d+n} \rightarrow \mathbb{R}^{d+2n}$), we define

$$\begin{bmatrix} \widetilde{Z}_{j,\theta} & \widetilde{X}_{j,\theta} \\ \widetilde{Z}_{j,\theta} & \widetilde{X}_{j,\theta} \end{bmatrix} := [Z_{j,\theta} \ X_{j,\theta}] := DK_{j,\theta} \in M_{d+2n,d+n}(\mathbb{R})$$

as in (3.6). If the matrix $\widetilde{X}_{j,\theta}^\top \widetilde{X}_{j,\theta}$ is invertible (cf. Definition 4.3 below), define $R_{j,\theta} := (\widetilde{X}_{j,\theta}^\top \widetilde{X}_{j,\theta})^{-1} \in M_{n,n}(\mathbb{R})$ as in (3.7),

$$\widetilde{Y}_{j,\theta} := \widetilde{J}_{K_j(\theta)}^{-1} \widetilde{X}_{j,\theta} R_{j,\theta} \in M_{2n,n}(\mathbb{R}), \quad Y_{j,\theta} := \begin{bmatrix} 0 \\ \widetilde{Y}_{j,\theta} \end{bmatrix} \in M_{d+2n,n}(\mathbb{R})$$

as in (3.8) and (3.9), and the approximate change of basis matrix $M_{j,\theta}$ as in (3.15):

$$M_{j,\theta} := [DK_{j,\theta} \ Y_{j,\theta}] = [Z_{j,\theta} \ X_{j,\theta} \ Y_{j,\theta}] = \begin{pmatrix} \widetilde{Z}_{j,\theta} & \widetilde{X}_{j,\theta} & 0 \\ \widetilde{Z}_{j,\theta} & \widetilde{X}_{j,\theta} & \widetilde{Y}_{j,\theta} \end{pmatrix} \in M_{d+2n,d+2n}(\mathbb{R}).$$

Below we will need to invert the matrices $M_{j,\theta}$ for solving the linearized equation, which motivates the following definition.

Definition 4.3. The map $K \in \mathcal{W}_\rho$ is said to be a non-degenerate torus if the matrices $\widetilde{X}_\theta^\top \widetilde{X}_\theta \in M_{n,n}(\mathbb{R})$ and $M_\theta \in M_{d+2n,d+2n}(\mathbb{R})$ defined above (and therefore, the matrix W_θ as in (3.33)) are invertible.

We will always assume that K_j is a non-degenerate torus (this is a part of Definition 4.7 below which is one of the assumptions in the Main Theorem).

As before, we think of the columns $Z_{j,\theta,\mu}$, $X_{j,\theta,a}$, and $Y_{j,\theta,a}$ ($\mu = 1, \dots, d$, $a = 1, \dots, n$) of the matrices $Z_{j,\theta}$, $X_{j,\theta}$, and $Y_{j,\theta}$ as vectors in $T_{K_j(\theta)}\mathcal{P}$. If K_j is close to the true solution K , then these vectors still form a basis of $T_{K_j(\theta)}\mathcal{P}$ as in the true case. By construction, the columns of $Z_{j,\theta}$ and $X_{j,\theta}$ span the tangent space to the approximately invariant torus $\mathcal{K}_j := K_j(\mathbb{T}^{d+n})$:

$$\text{span} \{ \{Z_{j,\theta,\mu}\}_{\mu=1}^d, \{X_{j,\theta,a}\}_{a=1}^n \} = T_{K_j(\theta)}\mathcal{K}_j. \tag{4.11}$$

However, unlike the case of a true solution, \mathcal{K}_j is *not* invariant with respect to the flow of V_{λ_j} , and $\ker \Omega_{K_j(\theta)}$ is generally not a subspace of $T_{K_j(\theta)}\mathcal{K}_j$.

Making the substitution (4.10) in (4.9) and assuming that $M_{j,\theta}$ is invertible, we obtain the following equation for the new unknown function $\xi_{j,\theta}$:

$$\begin{aligned} & M_{j,\theta}^{-1} (DV_{\lambda_j, K_j(\theta)} M_{j,\theta} - \partial_\omega M_{j,\theta}) \xi_{j,\theta} - \partial_\omega \xi_{j,\theta} \\ &= -M_{j,\theta}^{-1} \left(e_{j,\theta} + \left[\frac{\partial V_\lambda}{\partial \lambda} \right]_{\lambda_j, K_j(\theta)} \varepsilon_j \right). \end{aligned} \tag{4.12}$$

To rewrite the coefficient of $\xi_{j,\theta}$ in (4.12) in a simple form, we want that

$$\begin{aligned} & (DV_{\lambda_j, K_j(\theta)} - \partial_\omega) M_{j,\theta} \\ &= M_{j,\theta} (C_{j,\theta} + B_{j,\theta}) := M_{j,\theta} \left(\begin{bmatrix} 0 & 0 & T_{j,\theta} \\ 0 & 0 & S_{j,\theta} \\ 0 & 0 & 0 \end{bmatrix} + B_{j,\theta} \right) \end{aligned} \tag{4.13}$$

where $B_{j,\theta}$ is “small,” i.e., vanishing if e_j becomes identically zero (cf. (3.25)). For the left-hand side, long and unenlightening calculations (cf. Section 3.3.2) yield

$$\begin{aligned} & (DV_{\lambda_j, K_j(\theta)} - \partial_\omega) M_{j,\theta} \\ &= \left[D e_{j,\theta} \ (DV_{\lambda_j, K_j(\theta)} - \partial_\omega) Y_{j,\theta} \right] \end{aligned}$$

$$\begin{aligned}
 &= \left[(DV_{\lambda_j, K_j(\theta)} - \partial_\omega) DK_{j,\theta} \quad (DV_{\lambda_j, K_j(\theta)} - \partial_\omega) \begin{bmatrix} 0 \\ \tilde{Y}_{j,\theta} \end{bmatrix} \right] \\
 &= \begin{bmatrix} 0 & 0 & \left[\frac{\partial V_{\lambda_j}}{\partial \tilde{x}} \right]_{K_j(\theta)} \left[\frac{\partial V_{\lambda_j}}{\partial \tilde{x}} \right]_{K_j(\theta)} \tilde{Y}_{j,\theta} \\ 0 & 0 & -2\tilde{J}_{K_j(\theta)}^{-1} \tilde{\Pi}_{j,\theta} \left[\frac{\partial \tilde{V}_{\lambda_j}}{\partial \tilde{x}} \right]_{K_j(\theta)}^{\text{sym}} \tilde{X}_{j,\theta} R_{j,\theta} \end{bmatrix} + \left[De_{j,\theta} \begin{bmatrix} 0 \\ \mathcal{E}[e_j](\theta) \end{bmatrix} \right].
 \end{aligned}$$

Here we have set $\tilde{\Pi}_{j,\theta} := \mathbb{I}_{2n} - \tilde{X}_{j,\theta} R_{j,\theta} \tilde{X}_{j,\theta}^\top : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ (cf. (3.21)),

$$\begin{aligned}
 \mathcal{E}[e_j](\theta) &:= \tilde{J}_{K_j(\theta)}^{-1} D\tilde{J}_{K_j(\theta)} e_{j,\theta} \tilde{Y}_{j,\theta} - \tilde{J}_{K_j(\theta)}^{-1} \left[\frac{\partial \tilde{e}_{j,\theta}}{\partial \theta} \right] R_{j,\theta} \\
 &\quad + \tilde{Y}_{j,\theta} \left(\left[\frac{\partial \tilde{e}_{j,\theta}}{\partial \theta} \right]^\top \tilde{X}_{j,\theta} + \tilde{X}_{j,\theta}^\top \left[\frac{\partial \tilde{e}_{j,\theta}}{\partial \theta} \right] \right) R_{j,\theta} \in M_{2n,n}(\mathbb{R}),
 \end{aligned} \tag{4.14}$$

where $\left[\frac{\partial \tilde{e}_{j,\theta}}{\partial \theta} \right]$ is the $2n \times n$ matrix with entries

$$\left[\frac{\partial \tilde{e}_{j,\theta}}{\partial \theta} \right]_a^i := \frac{\partial e_{j,\theta}^{d+i}}{\partial \theta^{d+a}}, \quad i = 1, \dots, 2n; \quad a = 1, \dots, n;$$

and $D\tilde{J}_{K_j(\theta)} e_{j,\theta}$ stands for the $2n \times 2n$ matrix with entries

$$(D\tilde{J}_{K_j(\theta)} e_{j,\theta})_k^i := \sum_{A=1}^{d+2n} \frac{\partial \tilde{J}_k^i}{\partial x^A} \Big|_{K_j(\theta)} e_{j,\theta}^A, \quad i, k = 1, \dots, 2n,$$

For the entries of the matrix $C_{j,\theta}$ (4.13), we obtain (similarly to Section 3.3.3)

$$\begin{aligned}
 T_{j,\theta} &= \mathcal{Z}_{j,\theta}^{-1} \left(\left[\frac{\partial V_{\lambda_j}}{\partial \tilde{x}} \right]_{K_j(\theta)} \tilde{Y}_{j,\theta} - 2\tilde{X}_{j,\theta} \tilde{Y}_{j,\theta}^\top \tilde{\Pi}_{j,\theta} \left[\frac{\partial \tilde{V}_{\lambda_j}}{\partial \tilde{x}} \right]_{K_j(\theta)}^{\text{sym}} \tilde{X}_{j,\theta} R_{j,\theta} \right), \\
 S_{j,\theta} &= -\tilde{Y}_{j,\theta}^\top \tilde{J}_{K_j(\theta)} \tilde{Z}_{j,\theta} \mathcal{Z}_{j,\theta}^{-1} \left[\frac{\partial V_{\lambda_j}}{\partial \tilde{x}} \right]_{K_j(\theta)} \tilde{Y}_{j,\theta} \\
 &\quad + 2 \left(\mathbb{I}_n - \tilde{Y}_{j,\theta}^\top \tilde{J}_{K_j(\theta)} \tilde{Z}_{j,\theta} \mathcal{Z}_{j,\theta}^{-1} \tilde{X}_{j,\theta} \right) \tilde{Y}_{j,\theta}^\top \tilde{\Pi}_{j,\theta} \left[\frac{\partial \tilde{V}_{\lambda_j}}{\partial \tilde{x}} \right]_{K_j(\theta)}^{\text{sym}} \tilde{X}_{j,\theta} R_{j,\theta},
 \end{aligned} \tag{4.15}$$

where we have set $\mathcal{Z}_{j,\theta} := \tilde{Z}_{j,\theta} + \tilde{X}_{j,\theta} \tilde{Y}_{j,\theta}^\top \tilde{J}_{K_j(\theta)} \tilde{Z}_{j,\theta} \in M_{d,d}(\mathbb{R})$ (cf. (3.30), (3.31)).

Summarizing, with the help of (4.13), (4.14), and (4.15), we rewrote (4.12) as

$$(C_{j,\theta} + B_{j,\theta}) \xi_{j,\theta} - \partial_\omega \xi_{j,\theta} = -M_{j,\theta}^{-1} \left(e_{j,\theta} + \left[\frac{\partial V_{\lambda_j}}{\partial \lambda} \right]_{\lambda_j, K_j(\theta)} \varepsilon_j \right), \tag{4.16}$$

where $C_{j,\theta}$ is defined in (4.13) and (4.15), and $B_{j,\theta}$ is a “small” matrix, given by

$$B_{j,\theta} = M_{j,\theta}^{-1} \left[De_{j,\theta} \begin{bmatrix} 0 \\ \mathcal{E}[e_j](\theta) \end{bmatrix} \right]. \tag{4.17}$$

4.3.2. Invertibility issues. We define the $(d+2n) \times (d+2n)$ matrices $Q_{j,\theta}$ and $W_{j,\theta}$ in exactly the same way as Q_θ (3.32) and W_θ (3.33) but with $X_\theta, Y_\theta, Z_\theta,$ and $J_{K(\theta)}$ replaced by $X_{j,\theta}, Y_{j,\theta}, Z_{j,\theta},$ and $J_{K_j(\theta)}$, respectively. Since the rank of the $(2n \times 2n)$ matrix $[\tilde{X}_{j,\theta} \tilde{Y}_{j,\theta}]$ is maximal and $\tilde{J}_{K_j(\theta)}$ is non-degenerate, $Q_{j,\theta}$ is non-degenerate. By a direct calculation we obtain (cf. (3.34))

$$Q_{j,\theta} M_{j,\theta} = W_{j,\theta} + P_{j,\theta} := W_{j,\theta} + \begin{bmatrix} 0 & 0 & 0 \\ \tilde{X}_{j,\theta}^\top \tilde{J}_{K_j(\theta)} \tilde{Z}_{j,\theta} & \tilde{X}_{j,\theta}^\top \tilde{J}_{K_j(\theta)} \tilde{X}_{j,\theta} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{4.18}$$

where the matrix $P_{j,\theta}$ is small (if K_j were a true solution, $P_{j,\theta}$ would be zero).

Lemma 4.4. *Assume that the hypotheses of Lemma 4.2 hold. Then there exists a constant C depending on $d, n, \sigma, \rho, \|DK_j\|_\rho, |V_{\lambda_j}|_{C^1, \mathcal{B}_r},$ and $|J|_{C^1, \mathcal{B}_r},$ such that for every δ satisfying $0 < \delta < \rho/2,$ we have the bound*

$$\|W_j^{-1}P_j\|_{\rho-2\delta} \leq C\gamma^{-1}\delta^{-(\sigma+1)}\|e_j\|_\rho. \tag{4.19}$$

Proof. Recalling the bound (4.3) on the norm of the pull-back L_j (4.2) of the presymplectic form Ω to the torus $\mathcal{K}_j = K_j(\mathbb{T}^{d+n}),$ we obtain

$$\begin{aligned} \|W_j^{-1}P_j\|_{\rho-2\delta} &\leq C\|P_j\|_{\rho-2\delta} \\ &\leq C_1\|\tilde{X}_j^\top(\tilde{J} \circ K_j)\tilde{Z}_j\|_{\rho-2\delta} + C_2\|\tilde{X}_j^\top(\tilde{J} \circ K_j)\tilde{X}_j\|_{\rho-2\delta} \\ &\leq C\|DK_j^\top(\tilde{J} \circ K_j)DK_j\|_{\rho-2\delta} = C\|L_j\|_{\rho-2\delta} \\ &\leq C\gamma^{-1}\delta^{-(\sigma+1)}\|e_j\|_\rho. \end{aligned}$$

□

The approximate factorization (4.18) can be used to write the inverse matrix $M_{j,\theta}^{-1}$ in a convenient form, and Lemma 4.4 yields some useful bounds.

Lemma 4.5. *Assume that the hypotheses of Lemma 4.2 hold. Let $0 < \delta < \frac{\rho}{2}$ and the error e_j satisfy the bound*

$$C\gamma^{-1}\delta^{-(\sigma+1)}\|e_j\|_\rho \leq \frac{1}{2}, \tag{4.20}$$

where C is the same constant as in (4.19). Then the matrix $M_{j,\theta}$ is invertible, and

$$M_{j,\theta}^{-1} = W_{j,\theta}^{-1}Q_{j,\theta} + M_{j,\theta}^E, \tag{4.21}$$

where the error term is

$$M_{j,\theta}^E = -(\mathbb{I}_{d+2n} + W_{j,\theta}^{-1}P_{j,\theta})^{-1}W_{j,\theta}^{-1}P_{j,\theta}W_{j,\theta}^{-1}Q_{j,\theta}, \tag{4.22}$$

and satisfies the bound

$$\|M_{j,\theta}^E\|_{\rho-2\delta} \leq C'\gamma^{-1}\delta^{-(\sigma+1)}\|e_j\|_\rho; \tag{4.23}$$

here C' is a constant that depends on the same parameters as the constant C in (4.19).

Proof. The expression (4.22) comes from (4.18), and (4.23) follows from (4.19). □

4.3.3. Bounds on the “small” parts. Recall that, to find an approximate solution of the linearized equation (4.9), we changed the variable $\Delta_{j,\theta}$ to $\xi_{j,\theta}$ by (4.10) to transform it to the form (4.12). Then we rewrote the coefficient of $\xi_{j,\theta}$ in (4.12) as a sum of a “big” part, $C_{j,\theta}$ (given by (4.13) and (4.15)), and a “small” part, $B_{j,\theta}$ (4.17). In the Lemma below we give bounds on the “small” terms in (4.12).

Lemma 4.6. *Let $K_j \in \mathcal{W}_\rho$ and the error e_j be defined by (4.5). Let the pair (λ_j, K_j) be non-degenerate (in the sense of Definition 4.7 below) for the family $\{V_\lambda\}$ of presymplectic analytic vector fields. If the error e_j satisfies (4.20), then the change of variables (4.10) transforms the linearized equation (4.9) to*

$$\begin{aligned} (C_{j,\theta} + B_{j,\theta})\xi_{j,\theta} - \partial_\omega\xi_{j,\theta} &= -M_{j,\theta}^{-1}(e_{j,\theta} + \partial_\lambda V_{\lambda_j, K_j(\theta)}\varepsilon_j) \\ &= -W_{j,\theta}^{-1}Q_{j,\theta}e_{j,\theta} - W_{j,\theta}^{-1}Q_{j,\theta}\left[\frac{\partial V_\lambda}{\partial \lambda}\right]_{\lambda_j, K_j(\theta)}\varepsilon_j \\ &\quad - M_{j,\theta}^E e_{j,\theta} - M_{j,\theta}^E\left[\frac{\partial V_\lambda}{\partial \lambda}\right]_{\lambda_j, K_j(\theta)}\varepsilon_j, \end{aligned} \tag{4.24}$$

where $C_{j,\theta}$ is defined by (4.13) and (4.15), $B_{j,\theta}$ by (4.17), M_j^E by (4.22). Furthermore,

$$\begin{aligned} \|B_j\|_{\rho-2\delta} &\leq C\delta^{-1}\|e_j\|_\rho, \\ \|M_j^E e_j\|_{\rho-2\delta} &\leq C\gamma^{-1}\delta^{-(\sigma+1)}\|e_j\|_\rho^2, \\ \|M_j^E \left[\frac{\partial V_\lambda}{\partial \lambda}\Big|_{\lambda_j} \circ K_j\right] \varepsilon_j\|_{\rho-2\delta} &\leq C\gamma^{-1}\delta^{-(\sigma+1)}\left\|\frac{\partial V_\lambda}{\partial \lambda}\Big|_{\lambda_j} \circ K_j\right\| \|e_j\|_\rho |\varepsilon_j|. \end{aligned} \tag{4.25}$$

Proof. Equation (4.24) follows directly from (4.16) and (4.21), so we only need to derive the bounds (4.25). Combining (4.17) and (4.21), we obtain

$$B_{j,\theta} = \left(W_{j,\theta}^{-1}Q_{j,\theta} + M_{j,\theta}^E\right) \left[De_{j,\theta} \begin{pmatrix} 0 \\ \mathcal{E}[e_j](\theta) \end{pmatrix}\right].$$

From the definition (4.14) of $\mathcal{E}[e_j](\theta)$ and the Cauchy bound (2.12),

$$\|\mathcal{E}[e_j]\|_{\rho-2\delta} \leq C_1\|e_j\|_{\rho-2\delta} + C_2\delta^{-1}\|e_j\|_{\rho-\delta} \leq C\delta^{-1}\|e_j\|_{\rho-\delta},$$

which, together with the bound (4.23) on M_j^E , yields the first bound in (4.25):

$$\begin{aligned} \|B_j\|_{\rho-2\delta} &\leq (\|W_j^{-1}Q_j\|_{\rho-2\delta} + \|M_j^E\|_{\rho-2\delta}) (\|De_j\|_{\rho-2\delta} + \|\mathcal{E}[e_j]\|_{\rho-2\delta}) \\ &\leq \left(C_1 + C_2\gamma^{-1}\delta^{-(\sigma+1)}\|e_j\|_\rho\right) \gamma^{-1}\|e_j\|_{\rho-\delta} \leq C\gamma^{-1}\|e_j\|_\rho. \end{aligned}$$

The remaining two bounds in (4.25) are direct consequences of (4.23). □

4.3.4. *Solving the simplified equation.* To use the Newton method for finding $\xi_{j,\theta}$, it is enough to solve (4.24) retaining only the ‘‘big’’ terms, i.e., ignoring all terms of higher order with respect to $\|e_j\|$. As we will show below (see (4.37)), the term ε_j is of order of $\|e_j\|$. Lemma 4.6 allows us to keep only the leading terms in (4.24):

$$C_{j,\theta}\xi_{j,\theta} - \partial_\omega \xi_{j,\theta} = -W_{j,\theta}^{-1}Q_{j,\theta}e_{j,\theta} - \Lambda_{j,\theta}\varepsilon_j, \tag{4.26}$$

where we have set

$$\Lambda_{j,\theta} := W_{j,\theta}^{-1}Q_{j,\theta} \left[\frac{\partial V_\lambda}{\partial \lambda}\Big|_{\lambda_j, K_j(\theta)}\right]. \tag{4.27}$$

Let us denote the first d components of $\xi_{j,\theta}$ by $\xi_{j,\theta}^z$, the next n components by $\xi_{j,\theta}^x$, and the last n components by $\xi_{j,\theta}^y$. Using the specific form of $C_{j,\theta}$ (4.13), we rewrite (4.26) in the form

$$\partial_\omega \begin{bmatrix} \xi_{j,\theta}^z \\ \xi_{j,\theta}^x \\ \xi_{j,\theta}^y \end{bmatrix} = W_{j,\theta}^{-1}Q_{j,\theta}e_{j,\theta} + \Lambda_{j,\theta}\varepsilon_j + \begin{bmatrix} T_{j,\theta} \xi_{j,\theta}^y \\ S_{j,\theta} \xi_{j,\theta}^y \\ 0 \end{bmatrix}. \tag{4.28}$$

Integrating both sides of (4.28) over \mathbb{T}^{d+n} , we obtain that (4.28) has a solution if and only if the average over \mathbb{T}^{d+n} of its right-hand side is 0:

$$\text{avg} \left(W_j^{-1}Q_j e_j\right) + \text{avg} \left(\Lambda_j\right) \varepsilon_j + \begin{bmatrix} \text{avg}(T_j \xi_j^y) \\ \text{avg}(S_j \xi_j^y) \\ 0 \end{bmatrix} = 0. \tag{4.29}$$

Observe that the right-hand side of the last n equations of the system (4.28) does not involve ξ_j^y , so that the last n equations of (4.28) have the form

$$\partial_\omega \xi_{j,\theta}^y = \left(W_{j,\theta}^{-1}Q_{j,\theta}e_{j,\theta} + \Lambda_{j,\theta}\varepsilon_j\right)^y, \tag{4.30}$$

where the superscript y stands for the last n components. The system (4.30) for $\xi_{j,\theta}^y$ has a solution if and only if the last-hand side of (4.30) has zero average. To satisfy this condition, we will require initially that ε_j satisfy the linear system

$$\text{avg} (W_j^{-1}Q_j e_j) + \text{avg} (\Lambda_j) \varepsilon_j = 0, \tag{4.31}$$

which in turn has a solution if and only if the matrix $\text{avg} (\Lambda_j)$ is invertible. In order to guarantee that we can solve (4.31) for ε_j , we introduce the following

Definition 4.7. The pair (λ, K) is said to be *non-degenerate* for a $(d + 2n)$ -parameter family of vector fields V_λ if $K \in \mathcal{W}_\rho$ is a non-degenerate torus (recall Definition 4.3) and the matrix Λ (as in (4.27)) has a non-singular average:

$$\text{rank avg} (\Lambda) = d + 2n, \quad \text{avg} (\Lambda) := \int_{\mathbb{T}^{d+2n}} \Lambda_\theta \, d\theta^1 \cdots d\theta^{d+2n}. \tag{4.32}$$

We assume that the pair (λ_j, K_j) is non-degenerate, and set ε_j to be equal to the preliminary value

$$\varepsilon_j^{\text{prelim}} := -\{\text{avg} (\Lambda_j)\}^{-1} \text{avg} (W_j^{-1}Q_j e_j), \tag{4.33}$$

so that (4.31) is satisfied; $\varepsilon_j^{\text{prelim}}$ satisfies the bound

$$|\varepsilon_j^{\text{prelim}}| \leq C \text{avg} (e_j) \leq C \|e_j\|_\rho. \tag{4.34}$$

This choice of ε_j guarantees the existence of a solution ξ_j^y of (4.30) which (thanks to (4.34) and the Rüssmann’s inequality (2.13)) satisfies the bound

$$\|\xi_j^y\|_{\rho-\delta} \leq C \gamma^{-1} \delta^{-\sigma} \|W_j^{-1}Q_j e_j + \Lambda_j \varepsilon_j^{\text{prelim}}\|_\rho \leq C \gamma^{-1} \delta^{-\sigma} \|e_j\|_\rho. \tag{4.35}$$

Having found ξ_j^y from solving (4.30), we redefine ε_j as

$$\varepsilon_j := -\{\text{avg} (\Lambda_j)\}^{-1} \left(\text{avg} (W_j^{-1}Q_j e_j) + \begin{bmatrix} \text{avg}(T_j \xi_j^y) \\ \text{avg}(S_j \xi_j^y) \\ 0 \end{bmatrix} \right) \tag{4.36}$$

to satisfy the solvability condition (4.29) for (4.28). Note that the change from $\varepsilon_j^{\text{prelim}}$ (4.33) to ε_j (4.36) does not affect the component ξ_j^y . Thanks to (4.35), ε_j satisfies

$$|\varepsilon_j| \leq C (\|e_j\|_{\rho-\delta} + \|\xi_j^y\|_{\rho-\delta}) \leq C \gamma^{-1} \delta^{-\sigma} \|e_j\|_\rho. \tag{4.37}$$

With the new value of ε_j from (4.36), we solve (4.28) to find ξ_j which, according to the Rüssmann’s inequality (2.13) and the bound (4.37), satisfies

$$\begin{aligned} \|\xi_j\|_{\rho-2\delta} &\leq C \gamma^{-1} \delta^{-\sigma} (\|W_j^{-1}Q_j e_j + \Lambda_j \varepsilon_j\|_{\rho-\delta} + C \|\xi_j^y\|_{\rho-\delta}) \\ &\leq C \gamma^{-1} \delta^{-\sigma} (\|e_j\|_\rho + |\varepsilon_j| + \gamma^{-1} \delta^{-\sigma} \|e_j\|_\rho) \leq C \gamma^{-2} \delta^{-2\sigma} \|e_j\|_\rho. \end{aligned} \tag{4.38}$$

We summarize our findings in the following lemma.

Lemma 4.8. *Assume the hypotheses of Lemma 4.6. Then there exist a parameter ε_j and a function ξ_j that solve the reduced linear equation (4.26) and satisfy the bounds (4.37) and (4.38).*

5. NEWTON’S METHOD

In this section we will collect the estimates for the j th step of the iterative scheme and show that the Newton method generates a Cauchy sequence of approximate solutions in a Banach space which converges to a true solution. We only give brief sketches, referring the reader to [21, 3] for details.

Lemma 5.1. *If the assumptions of Lemma 4.8 are satisfied and $r_j := \|K_j - K_0\|_{\rho_j} < r$, then there exist a function Δ_j and a parameter $\varepsilon_j \in \mathbb{R}^{d+2n}$ such that*

$$\begin{aligned} \|\Delta_j\|_{\rho_j-2\delta_j} &\leq c_j \gamma^{-2} \delta_j^{-2\sigma} \|e_j\|_{\rho_j} \\ \|D\Delta_j\|_{\rho_j-3\delta_j} &\leq c_j \gamma^{-2} \delta_j^{-(2\sigma+1)} \|e_j\|_{\rho_j} \\ |\varepsilon_j| &\leq c_j |\text{avg}(\Lambda_j)^{-1}| \|e_j\|_{\rho_j}, \end{aligned} \tag{5.1}$$

where c_j is a constant that depends on $n, d, r, \rho, |V_{\lambda_j}|_{C^2, \mathcal{B}_r}, \|DK_j\|_{\rho_j}, \|R_j\|_{\rho_j}$, and $\|\frac{\partial V_\lambda}{\partial \lambda}\|_{\rho_j}$. Additionally, if

$$r_j + c_j \gamma^{-2} \delta_j^{-(2\sigma-1)} \|e_j\|_{\rho_j} < r, \tag{5.2}$$

then

$$\|e_{j+1}\|_{\rho_{j+1}} \leq c_j \gamma^{-4} \delta_j^{-4\sigma} \|e_j\|_{\rho_j}^2. \tag{5.3}$$

Proof. The inequalities in (5.1) follow from Lemmata 2.7 and 4.6, and (4.10).

To see that $K_{j+1} \in \mathcal{B}_r$, that is, K_{j+1} stays within the neighborhood where V is holomorphically extended, we use (5.1) and (5.2):

$$\begin{aligned} \|K_{j+1} - K_0\|_{\rho_{j+1}-2\delta_{j+1}} &= \|K_j + \Delta_j - K_0\|_{\rho_{j+1}-2\delta_{j+1}} \\ &\leq \|K_j - K_0\|_{\rho_j} + \|\Delta_j\|_{\rho_j-2\delta_j} \\ &\leq r_j + c_j \gamma^{-2} \delta_j^{-(2\sigma+1)} \|e_j\|_{\rho_j} < r. \end{aligned}$$

To prove (5.3), recall from (4.8) that $\xi_j = M_j^{-1} \Delta_j$ was found by solving (4.26), so

$$\begin{aligned} DV_{\lambda_j, K_j(\theta)} \Delta_{j,\theta} - \partial_\omega \Delta_{j,\theta} + \left[\frac{\partial V_\lambda}{\partial \lambda}\right]_{\lambda_j, K_j(\theta)} \varepsilon_j + e_j \\ = M_{j,\theta} \left(B_{j,\theta} \xi_{j,\theta} + M_{j,\theta}^E e_{j,\theta} + M_{j,\theta}^E \left[\frac{\partial V_\lambda}{\partial \lambda}\right]_{\lambda_j} \circ K_j \right), \end{aligned}$$

and each term on the right hand side is quadratically small from Lemma 4.6, hence

$$\|DV_{\lambda_j, K_j} \Delta_j - \partial_\omega \Delta_j + \left[\frac{\partial V_\lambda}{\partial \lambda}\right]_{\lambda_j, K_j} \varepsilon_j + e_j\|_{\rho_j-2\delta_j} \leq C \gamma^{-3} \delta^{-(3\sigma+1)} \|e_j\|_{\rho_j}^2.$$

Finally, recalling the Taylor expansion (4.8) of $e_{j+1,\theta}$, we see that the remainder term is on the order of $\|\Delta_j\|_{\rho_j-2\delta_j}^2$. Thus we get the estimate (5.3). \square

The lemma below guarantees that, if the error is small and some invertibility conditions are met at the j th step, then the invertibility holds at the $(j+1)$ st step.

Lemma 5.2. *Assume the setup of Lemma 5.1. If $c_j \gamma^{-2} \delta_j^{-(\sigma+1)} \|e_j\|_{\rho_j} \leq 1/2$, then:*

- (i) if $\tilde{X}_j^\top \tilde{X}_j$ is invertible, then $\tilde{X}_{j+1}^\top \tilde{X}_{j+1}$ is invertible;
- (ii) if W_j is invertible, then W_{j+1} is invertible;
- (iii) if $\text{avg}(\Lambda_j)$ is invertible, then $\text{avg}(\Lambda_{j+1})$ is invertible.

Proof. Recalling that \tilde{X}_j is a part of the matrix DK_j , we obtain $\tilde{X}_{j+1}^\top \tilde{X}_{j+1} = \tilde{X}_j^\top \tilde{X}_j + P_j$, with $P_j := \tilde{X}_j^\top \tilde{\Delta}_{j,\tilde{x}} + \tilde{\Delta}_{j,\tilde{x}}^\top \tilde{X}_j + \tilde{\Delta}_{j,\tilde{x}}^\top \tilde{\Delta}_{j,\tilde{x}}$, where $\tilde{\Delta}_{j,\tilde{x}} \in M_{2n,2n}(\mathbb{R})$ has entries $(\tilde{\Delta}_{j,\tilde{x}})^i_k = \partial(\tilde{\Delta}_j)^i / \partial \tilde{x}^k$. The bounds (5.1) give an upper bound on the size of P_j . The matrix $\tilde{X}_j^\top \tilde{X}_j$ is invertible by assumption, $\mathbb{I}_n + (\tilde{X}_j^\top \tilde{X}_j)^{-1} P_j$ is invertible by the Neumann series, so $\tilde{X}_{j+1}^\top \tilde{X}_{j+1} = \tilde{X}_j^\top \tilde{X}_j (\mathbb{I}_n + (\tilde{X}_j^\top \tilde{X}_j)^{-1} P_j)$ is invertible, which completes the proof of (i). The proofs of (ii) and (iii) are similar. \square

The lemma below shows how close the initial approximation has to be for the Newton method to be iterated indefinitely and to converge to a true solution K_∞ , and gives a bound on the difference between K_∞ and the initial approximation K_0 . The proof can be found in [21, Lemma 13].

Lemma 5.3. *Let $\{c_j\}_{j \geq 0}$ be the sequence of constants from Lemmata 5.1 and 5.2. For $0 < \delta_0 < \min(\rho_0/12, 1)$ we define*

$$\delta_j := \delta_0 2^{-j}, \quad \rho_j := \rho_{j-1} - 6\delta_{j-1}, \quad r_j := \|K_j - K_0\|_{\rho_j},$$

$\rho_\infty := \lim_{j \rightarrow \infty} \rho_j$, $K_\infty := \lim_{j \rightarrow \infty} K_j$. Then there exists a constant $C > 0$ depending on d , n , $|V_\lambda|_{C^2, \mathcal{B}_r}$, $|J_0|_{C^1, \mathcal{B}_r}$, $\|DK_0\|_{\rho_0}$, and $|\{\text{avg}(\Lambda_0)\}^{-1}|$ such that if $\|e_0\|_{\rho_0}$ satisfies the conditions

$$C 2^{4\sigma} \gamma^{-4} \delta_0^{-4\sigma} \|e_0\|_{\rho_0} \leq \frac{1}{2},$$

$$C \left(1 + \frac{2^{4\sigma}}{2^{2\sigma} - 1}\right) \gamma^{-2} \delta_0^{-2\sigma} \|e_0\|_{\rho_0} < r,$$

then the Newton method converges to a true solution $(\lambda_\infty, K_\infty)$. Furthermore,

$$\|K_\infty - K_0\|_{\rho_0 - 6\delta_0} \leq \frac{2^{2\sigma}}{2^{2\sigma} - 1} c \gamma^{-2} \delta_0^{-2\sigma} \|e_0\|_{\rho_0}.$$

Acknowledgments. The research of the authors was partially supported by NSF grant DMS-0807658. NPP also acknowledges the generous support by the Nancy Scofield Hester Presidential Professorship. We would like to thank Rafael de la Llave, Garrett Alston, and Mahesh Sunkula for encouragement and enlightening discussions.

REFERENCES

- [1] R. Abrahamm J. E. Marsden; *Foundations of Mechanics*, second ed., Benjamin/Cummings, Reading, Mass., 1978. MR 515141 (81e:58025)
- [2] L. Abrunheiro, M. Camarinha, J. Clemente-Gallardo; *Cubic polynomials on Lie groups: reduction of the Hamiltonian system*, J. Phys. A **44** (2011), no. 35, 355203, 16, Corrigendum: 46(18):189501, 2, 2013. MR 2823461 (2012k:37116)
- [3] H. N. Alishah, R. de la Llave; *Tracing KAM tori in presymplectic dynamical systems*, J. Dynam. Differential Equations **24** (2012), no. 4, 685–711. MR 3000600
- [4] J. L. Anderson, P. G. Bergmann; *Constraints in covariant field theories*, Physical Rev. (2) **83** (1951), 1018–1025. MR 0044382
- [5] V. I. Arnold, V. V. Kozlov, A. I. Neishtadt; *Mathematical Aspects of Classical and Celestial Mechanics*, third ed., Encyclopaedia of Mathematical Sciences, vol. 3, Springer-Verlag, Berlin, 2006, [Dynamical systems. III]. MR 2269239
- [6] M. Barbero-Liñán, M. C. Muñoz-Lecanda; *Geometric approach to Pontryagin's maximum principle*, Acta Appl. Math. **108** (2009), no. 2, 429–485. MR 2551483
- [7] C. Batlle, J. Gomis, J. M. Pons, N. Roman; *Lagrangian and Hamiltonian constraints*, Lett. Math. Phys. **13** (1987), no. 1, 17–23. MR 878657 (88h:58041)

- [8] C. Batlle, J. Gomis, J. M. Pons, N. Román-Roy; *Equivalence between the Lagrangian and Hamiltonian formalism for constrained systems*, J. Math. Phys. **27** (1986), no. 12, 2953–2962. MR 866595 (88a:58060)
- [9] S. Bauer; *On the existence of KAM Tori for Presymplectic Vector Fields*, ProQuest LLC, Ann Arbor, MI, 2016, Thesis (Ph.D.)—The University of Oklahoma.
- [10] P. G. Bergmann, I. Goldberg; *Dirac bracket transformations in phase space*, Phys. Rev. (2) **98** (1955), 531–538. MR 0074314
- [11] H. Broer, M. Sevryuk; *KAM theory: quasi-periodicity in dynamical systems*, Handbook of Dynamical Systems. Vol. 3, Elsevier/North-Holland, Amsterdam, 2010, pp. 249–344. MR 3292649
- [12] R. Calleja, R. de la Llave; *A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification*, Nonlinearity **23** (2010), no. 9, 2029–2058. MR 2672635
- [13] R. C. Calleja, A. Celletti, R. de la Llave; *A KAM theory for conformally symplectic systems: efficient algorithms and their validation*, J. Differential Equations **255** (2013), no. 5, 978–1049. MR 3062760
- [14] J. F. Cariñena, J. Gomis, L. A. Ibort, N. Román; *Canonical transformations theory for presymplectic systems*, J. Math. Phys. **26** (1985), no. 8, 1961–1969. MR 796226 (86j:58041)
- [15] J. F. Cariñena, L. A. Ibort; *Geometric theory of the equivalence of Lagrangians for constrained systems*, J. Phys. A **18** (1985), no. 17, 3335–3341. MR 817862
- [16] J. F. Cariñena, C. Lopez, N. Román-Roy; *Geometric study of the connection between the Lagrangian and Hamiltonian constraints*, J. Geom. Phys. **4** (1987), no. 3, 315–334. MR 957017 (89i:58031)
- [17] J. F. Cariñena, J. Nasarre; *On the symplectic structures arising in geometric optics*, Fortschr. Phys. **44** (1996), no. 3, 181–198. MR 1400305
- [18] J. F. Cariñena, J. Nasarre; *Presymplectic geometry and Fermat’s principle for anisotropic media*, J. Phys. A **29** (1996), no. 8, 1695–1702. MR 1395798
- [19] L. Chierchia; *KAM lectures*, Dynamical Systems. Part I, Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa, 2003, Hamiltonian systems and celestial mechanics, Selected papers from the Research Trimester held in Pisa, February 4–April 26, 2002, pp. 1–55. MR 2071231
- [20] L. Chierchia; *Kolmogorov-Arnold-Moser (KAM) theory*, Encyclopedia of Complexity and Systems Science (R. A. Meyers, ed.), Springer, 2009, pp. 5064–5091.
- [21] R. de la Llave, A. González, À. Jorba, J. Villanueva; *KAM theory without action-angle variables*, Nonlinearity **18** (2005), no. 2, 855–895. MR 2122688 (2005k:37131)
- [22] M. de León, D. Martín de Diego; *Almost product structures and Poisson reduction of presymplectic systems*, Extracta Math. **10** (1995), no. 1, 37–45. MR 1359590
- [23] P. A. M. Dirac; *Generalized Hamiltonian dynamics*, Canadian J. Math. **2** (1950), 129–148. MR 0043724 (13,306b)
- [24] P. A. M. Dirac; *The Hamiltonian form of field dynamics*, Canadian J. Math. **3** (1951), 1–23. MR 0043725 (13,306c)
- [25] P. A. M. Dirac; *Generalized Hamiltonian dynamics*, Proc. Roy. Soc. London. Ser. A **246** (1958), 326–332. MR 0094205 (20 #724)
- [26] H. S. Dumas; *The KAM story: A Friendly Introduction to the Content, History, and Significance of Classical Kolmogorov-Arnold-Moser Theory*, World Scientific, New Jersey, 2014.
- [27] C. Duval; *Polarized spinoptics and symplectic physics*, Preprint (2013), arXiv:1312.4486 [math-ph].
- [28] A. Echeverría-Enríquez, J. Marín-Solano, M. C. Muñoz-Lecanda, N. Román-Roy; *Geometric reduction in optimal control theory with symmetries*, Rep. Math. Phys. **52** (2003), no. 1, 89–113. MR 2006728 (2004h:49037)
- [29] A. Echeverría-Enríquez, M. C. Muñoz-Lecanda, N. Román-Roy; *Reduction of presymplectic manifolds with symmetry*, Rev. Math. Phys. **11** (1999), no. 10, 1209–1247. MR 1734712 (2001b:53106)
- [30] E. Fontich, R. de la Llave, Y. Sire; *Construction of invariant whiskered tori by a parameterization method. I. Maps and flows in finite dimensions*, J. Differential Equations **246** (2009), no. 8, 3136–3213. MR 2507954 (2010c:37139)
- [31] A. M. Fox, J. D. Meiss; *Critical invariant circles in asymmetric and multiharmonic generalized standard maps*, Commun. Nonlinear Sci. Numer. Simul. **19** (2014), no. 4, 1004–1026. MR 3119277

- [32] A. González-Enríquez, A. Haro, R. de la Llave; *Singularity theory for non-twist KAM tori*, Mem. Amer. Math. Soc. **227** (2014), no. 1067, vi+115. MR 3154587
- [33] M. Gotay; *Presymplectic Manifolds, Geometric Constraint Theory and Dirac-Bergmann Theory of Constraints*, Ph.D. thesis, University of Maryland, 1979.
- [34] M. J. Gotay, J. M. Nester; *Presymplectic Lagrangian systems. I. The constraint algorithm and the equivalence theorem*, Ann. Inst. H. Poincaré Sect. A (N.S.) **30** (1979), no. 2, 129–142. MR 535369 (80j:58035)
- [35] M. J. Gotay, J. M. Nester, G. Hinds; *Presymplectic manifolds and the Dirac-Bergmann theory of constraints*, J. Math. Phys. **19** (1978), no. 11, 2388–2399. MR 506712 (80e:58025)
- [36] A. Hanson, T. Regge, C. Teitelboim; *Constrained Hamiltonian Systems*, Accademia Nazionale dei Lincei, Roma, 1976.
- [37] À. Haro, M. Canadell, A. Luque, J. M. Mondelo, J.-L. Figueras; *The Parameterization Method for Invariant Manifolds: From Rigorous Results to Effective Computations*, Applied Mathematical Sciences, vol. 195, Springer-Verlag, 2016.
- [38] M. Henneaux, C. Teitelboim; *Quantization of Gauge Systems*, Princeton University Press, Princeton, NJ, 1992. MR 1191617 (94h:81003)
- [39] G. Huguet, R. de la Llave, Y. Sire; *Computation of whiskered invariant tori and their associated manifolds: new fast algorithms*, Discrete Contin. Dyn. Syst. **32** (2012), no. 4, 1309–1353. MR 2851901 (2012k:37126)
- [40] P. Libermann, C.-M. Marle; *Symplectic Geometry and Analytical Mechanics*, D. Reidel, Dordrecht, 1987. MR 882548 (88c:58016)
- [41] R. de la Llave; *A tutorial on KAM theory*, Smooth Ergodic Theory and Its Applications (Seattle, WA, 1999), Proc. Sympos. Pure Math., vol. 69, Amer. Math. Soc., Providence, RI, 2001, pp. 175–292. MR 2002h:37123
- [42] A. Luque, J. Villanueva; *A KAM theorem without action-angle variables for elliptic lower dimensional tori*, Nonlinearity **24** (2011), no. 4, 1033–1080. MR 2773779
- [43] J. Moser; *Stable and Random Motions in Dynamical Systems*, Annals of Mathematics Studies, No. 77, Princeton University Press, Princeton, N. J.; University of Tokyo Press, Tokyo, 1973. MR 0442980
- [44] E. Newman, P. G. Bergmann; *Lagrangians linear in the “velocities”*, Phys. Rev. (2) **99** (1955), 587–592. MR 0072039
- [45] T. Noda, M. Oda; *Geometry of moment maps and reductions for presymplectic manifolds*, Georgian Math. J. **16** (2009), no. 2, 343–352. MR 2562366
- [46] J. Pöschel; *A lecture on the classical KAM theorem*, Smooth Ergodic Theory and Its Applications (Seattle, WA, 1999), Proc. Sympos. Pure Math., vol. 69, Amer. Math. Soc., Providence, RI, 2001, pp. 707–732. MR 1858551
- [47] H. J. Rothe, K. D. Rothe; *Classical and Quantum Dynamics of Constrained Hamiltonian Systems*, World Scientific Lecture Notes in Physics, vol. 81, World Scientific, Hackensack, NJ, 2010. MR 2723931
- [48] G. Rudolph, M. Schmidt; *Differential Geometry and Mathematical Physics, I: Manifolds, Lie Groups and Hamiltonian Systems*, Springer, 2013.
- [49] H. Rüssmann; *On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus*, Dynamical Systems, Theory and Applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974), Lecture Notes in Phys., vol. 38, Springer, Berlin, 1975, pp. 598–624. MR 0467824 (57 #7675)
- [50] D. A. Salamon; *The Kolmogorov-Arnold-Moser theorem*, Math. Phys. Electron. J. **10** (2004), Paper 3, 37 pp. (electronic). MR 2111297 (2005h:37135)
- [51] M. B. Sevryuk; *The classical KAM theory at the dawn of the twenty-first century*, Mosc. Math. J. **3** (2003), no. 3, 1113–1144, 1201–1202. MR 2078576
- [52] E. C. G. Sudarshan, N. Mukunda; *Classical Dynamics: A Modern Perspective*, Wiley-Interscience, New York-London-Sydney, 1974. MR 0434047 (55 #7016)
- [53] K. Sundermeyer; *Constrained Dynamics*, Lecture Notes in Physics, vol. 169, Springer-Verlag, Berlin-New York, 1982. MR 678773 (84f:58051)

SEAN BAUER

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OKLAHOMA, NORMAN, OK 73019, USA.

AMERICAN FIDELITY, 9000 CAMERON PARKWAY, OKLAHOMA CITY, OK 73114, USA

Email address: sean.michael.bauer@gmail.com

NIKOLA P. PETROV
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OKLAHOMA, NORMAN, OK 73019, USA
Email address: npetrov@ou.edu