UNIVERSITAT DE BARCELONA

Flow map parameterization methods for invariant tori in Hamiltonian systems

Álvaro Fernández-Mora

Flow map parameterization methods for invariant tori in Hamiltonian systems

by

Álvaro Fernández-Mora

A thesis submitted for the PhD programme
in Mathematics and Computer Science

Advisors:
Àlex Haro
Josep-Maria Mondelo

Departament de Matemàtiques i Informàtica
Universitat de Barcelona
Barcelona, March 19, 2025

Memòria presentada per aspirar al grau de Doctor en Matemàtiques i Informàtica per la Universitat de Barcelona.

Certifico que aquesta memòria ha estat realitzada per en Álvaro Fernández-Mora, sota la meva direcció.

Barcelona, 19 de març de 2025

Àlex Haro Josep-Maria Mondelo

Contents

A mis abuelos María y Aquilino.

Agradecimientos

Me gustaría tomar un momento para agradecer a todos aquellos que me han ayudado en este viaje de una manera u otra.

A Àlex, que durante los últimos años me has ayudado en infinitas cuestiones de temas muy diversos: académicos, burocráticos, docentes, personales, gastronómicos, etc. Te agradezco enormemente la paciencia que has tenido y el entusiasmo que me has transmitido desde aquel día que te mandé el primer mail.

A Josep-María, por todo lo que me has enseñado sobre astrodinámica y computación. Por tener la paciencia de sentarte conmigo a leer el Kernighan and Ritchie con alguien que no sabía C. Si hoy administro el clúster de la facultad esto también es gracias a ti.

Me gustaría agradecer también a Ariadna por dirigirme al grupo de Barcelona cuando nos conocimos en Goddard allá por el 2017. Yo era aún un estudiante de máster y nuestras tertulias matemáticas me motivaron a empezar esta nueva etapa en Barcelona.

A Arturo, porque ir a tomar café nunca fue algo tan prolífero. Por poder compartir frikadas de vi y bash. Gracias por siempre cuidar de los doctorantes del Xalet. A Àngel por siempre resolver cualquier cuestión con la que iba a molestarte y a Joan Carles y Susana porque siempre hicisteis del Xalet un lugar agradable. Agradezco también a Ainoa y Pedro, expertos en graficación, por estar allí y sacarme del despacho a por una cerveza. A Miquel, porque trabajar y organizar cosas contigo siempre fue muy fácil y cómodo. También a los nuevos doctorantes Rafa y Philip aunque haya tenido menos tiempo con vosotros.

Me gustaría agradecer también a Rafael de la Llave por todo lo que me enseñaste esos meses en Atlanta y los consejos sobre la vida académica que me has transmitido. Agradezco también a la gente de Georgia Tech su hospitalidad.

A Jordi-Lluís por esos días que me recibiste en Upsala dónde programabamos escuchando a Seguridad Social. Te agradezco tus innumerables consejos y tu cercanía.

A mis amigos de Sevilla, por siempre estar ahí pese a que lleve ya una década de idas y venidas. Agradezco también a los nuevos amigos que he hecho en Barcelona, en especial a Danai, Enrico, y Celia.

Agradezco con especial cariño a mis padres Rosario y Javier. Por su apoyo incondicional pese a no saber qué hago exactamente. También a mis hermanas Cristina y Natalia a las que siempre mando mis papers solo para que digan que no entienden nada.

Por último, agradezco a todos los que he conocido en la UB y a los organismos que han financiado este proyecto. Esta tesis ha sido financiada por la beca 2021 FI_B 00122 además de por los proyectos MCINN-AEI PID2020-118281GB-C31, PID2021-125535NB-I00, 2021 SGR 00113, y Marie Skłodowska-Curie grant agreement 734557.

Introducció

Des de la revolució newtoniana, molts problemes físics poden ser formulats com a solucions d'equacions diferencials que defineixen sistemes dinàmics *continus*. Anem a introduir un escenari senzill per a aquests problemes. Sigui $\mathcal{M}$ una varietat suau i $X : \mathcal{M} \to T\mathcal{M}$ un camp vectorial analític real. Ens ocuparem de les solucions d'equacions diferencials ordinaries donades per

$$\dot{z} = X(z), \tag{0.1}$$

per a $z \in \mathcal{M}$. El flux de (0.1), $\phi : \mathbb{R} \times \mathcal{M} \to \mathcal{M}$, defineix un grup uniparamètric de difeomorfismes suaus $\{\phi_t\}_{t\in\mathbb{R}}$ que determinen la dinàmica sobre $\mathcal{M}$ sota X. Una altra perspectiva per estudiar l'evolució d'un sistema consisteix a estudiar les òrbites d'un d'una aplicació que actua sobre $\mathcal{M}$ com

$$\begin{aligned} F: \quad \mathcal{M} &\longrightarrow \mathcal{M} \\ z &\longmapsto F(z), \end{aligned}$$

on les òrbites són donades per les iteracions de F. Aquest sistema dinàmic es diu *discret.* És possible construir sistemes dinàmics discrets a partir dels continus de manera senzilla. Per exemple, es podria considerar l'aplicació de temps-1, ϕ_1, o l'aplicació de retorn de primer ordre de Poincaré. La construcció de camps vectorials compatibles a partir d'aplicacions és una qüestió molt més subtil. Subratllem que, tot i que aquesta tesi tracta principalment sistemes Hamiltonians no autònoms, per ara ens concentrarem en els sistemes *autònoms* per claredat. Podríem considerar que això no és una restricció, perquè recordem que qualsevol sistema dependent del temps pot ser autonomitzat estenent l'espai de fase.

Certament podríem demanar estructura addicional sobre $\mathcal{M}, X$, i F. Per exemple, en mecànica Hamiltoniana, $\mathcal{M}$ és una varietat suau equipada amb una 2$-$forma diferencial ω que és simplèctica, és a dir, tancada ($d\omega = 0$) i no degenerada. La condició de no degeneració implica que $\mathcal{M}$ ha de ser de dimensió parella. També necessitem una funció suau $H : \mathcal{M} \to \mathbb{R}$, el Hamiltonià. El camp vectorial $X : \mathcal{M} \to T\mathcal{M}$ associat a H es defineix com

$$\iota_X \boldsymbol{\omega} = -\mathrm{d}H, \tag{0.2}$$

on $\iota_X \boldsymbol{\omega} = \boldsymbol{\omega}(X, \cdot)$ és la contracció de la forma simplèctica amb el camp vectorial X.

A causa d'un teorema de Darboux, sempre existeixen coordenades canòniques (q, p) que transformen localment la forma simplèctica $\boldsymbol{\omega}$ en la forma simplèctica estàndard

$$\boldsymbol{\omega}_o = \mathrm{d}p \wedge \mathrm{d}q = \sum_{i=1}^{d} \mathrm{d}p_i \wedge \mathrm{d}q_i,$$

on $\wedge$ és el producte exterior. En aquestes coordenades, les equacions de moviment donades per (0.2) són

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}.$$

Aquestes són les equacions clàssiques (canòniques) de Hamilton.

A partir de la màgica fórmula de Cartan podem veure immediatament que el flux de X preserva l'estructura simplèctica ja que

$$\mathcal{L}_X \boldsymbol{\omega} = \mathrm{d}\iota_X \boldsymbol{\omega} + \iota_X \mathrm{d}\boldsymbol{\omega} = 0,$$

on $\mathcal{L}_X$ és la derivada de Lie respecte al camp vectorial X i d és la derivada exterior. El primer terme és zero perquè $\iota_X \boldsymbol{\omega}$ és una 1−forma exacta i el segon és zero perquè $\boldsymbol{\omega}$ és tancada per suposició. Llavors, per a qualsevol $t \in \mathbb{R}$, ϕ_t és un simplectomorfisme i $\{\phi_t\}_{t\in\mathbb{R}}$ és un grup uniparamètric de simplectomorfismes. Que ϕ_t preservi la forma simplèctica també es pot escriure com

$$\phi_t^* \boldsymbol{\omega} = \boldsymbol{\omega}, \tag{0.3}$$

on $\phi_t^* \boldsymbol{\omega}$ denota el pullback de la forma simplèctica pel difeomorfisme ϕ_t, que ja sabem que és un simplectomorfisme.

Si suposem que $\boldsymbol{\omega}$ és exacta, és a dir, existeix una 1−forma $\boldsymbol{\alpha}$ tal que $\boldsymbol{\omega} = \mathrm{d}\boldsymbol{\alpha}$, aleshores observem que

$$\begin{aligned} 0 =& \phi_t^* \mathrm{d}\boldsymbol{\alpha} - \mathrm{d}\boldsymbol{\alpha} \\ =& \mathrm{d}(\phi_t^* \boldsymbol{\alpha} - \boldsymbol{\alpha}), \end{aligned}$$

on hem utilitzat que el pullback i la derivada exterior commuten. En conseqüència, la 1−forma $\phi_t^* \boldsymbol{\alpha} - \boldsymbol{\alpha}$ és tancada. A més, es pot demostrar que la forma $\phi_t^* \boldsymbol{\alpha} - \boldsymbol{\alpha}$ és exacta. Vegeu l'Apèndix A. És a dir, existeix una funció p_t tal que $\phi_t^* \boldsymbol{\alpha} - \boldsymbol{\alpha} = \mathrm{d}p_t$. Direm aleshores que ϕ_t és exactament simplectica o, equivalentment, un simplectomorfisme exacte. L'estructura que tenim per al nostre sistema dinàmic té implicacions sobre el tipus de solucions que podem esperar i com les podem calcular. En aquest treball, ens ocuparem principalment d'objectes invariants per als quals la geometria ens ajudarà a demostrar la seva existència i en el seu càlcul. Ara mostrarem algunes implicacions de naturalesa geomètrica en la dinàmica.

Considerem una funció Hamiltoniana analítica real $H : U \to \mathbb{R}$, on $U \subset \mathbb{R}^{2n}$ és un conjunt obert, i sigui $X : U \to \mathbb{R}^{2n}$ el camp vectorial Hamiltonià construït a partir de (0.2). Introduïm el claudàtor de Poisson de dues funcions qualssevol $f, g : U \to \mathbb{R}$ com

$$\{f, g\} = \boldsymbol{\omega}(X_f, X_g),$$

on X_f i X_g són els camps vectorials Hamiltonians associats a les funcions f i g. En coordenades canòniques (q, p), el claudàtor de Poisson es pot expressar com

$$\{f, g\} = \sum_{i=1}^{n} \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}.$$

D'aquesta expressió veiem que la variació d'una funció f al llarg de les corbes integrals del camp vectorial Hamiltonià es dóna per

$$\frac{\mathrm{d}}{\mathrm{d}t} f \circ \phi_t = \left(\sum_{i=1}^{n} \frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial p_i}\dot{p}_i \right) \circ \phi_t$$

$$= \left(\sum_{i=1}^{n} \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right) \circ \phi_t$$

$$= \{H, f\} \circ \phi_t, \tag{0.4}$$

Conseqüentment, si el claudàtor de Poisson $\{f, H\}$ és nul, f és una integral de moviment per al sistema dinàmic definit per H. Cal notar que aquesta conclusió també és vàlida per a una forma simplèctica arbitrària $\boldsymbol{\omega}$ en una formulació sense coordenades com segueix

$$\{f, H\} = \boldsymbol{\omega}(X_f, X_H) = (\iota_{X_f}\boldsymbol{\omega})(X_H) = -\mathrm{df}(X_H) = -X_H(f) = -\mathcal{L}_{X_H} f,$$

on hem utilitzat (0.2). Per tant, si $\{f, H\} = 0$, f és una integral de moviment per al flux Hamiltonià donat per H. Les implicacions de l'existència d'integrals de moviment en un sistema Hamiltonià són clares a partir del següent teorema ben conegut [Lio55, Arn89]

Theorem 0.1 (Liouville-Arnold). *Suposem que tenim una col · lecció de n funcions $\{H_i\}_{i=0}^{n-1}$ en involució sobre una varietat de $2n$ dimensions, és a dir,*

$$\{H_i, H_j\} \equiv 0 \qquad \textit{per } i, j = 0, \dots n-1,$$

i considerem el conjunt de nivells

$$M_f = \{z : H_i = c_i \quad \textit{per } i = 0, \dots, n-1\}.$$

Suposem que $\mathrm{d}H_i$ *són linealment independents a cada punt de M_f. Llavors:*

1.) M_f és una varietat suau, invariant sota el flux Hamiltonià de H_0.

2.) Si la varietat M_f és compacta i conexa llavors és difeomorfa a un tor $n-$dimensional.

3.) El flux de H_0 sobre M_f és quasi-periòdic.

4.) Les equacions canòniques amb Hamiltonià H_0 es poden integrar per quadratures.

Un corol · lari del teorema anterior és que si M_f és difeomorfa a $\mathbb{T}^n$, existeix una transformació simplèctica a coordenades d'*acció-angle* $(\theta, I) = \psi(q, p)$ de manera que el nou Hamiltonianà $\mathcal{H}(\theta, I) = H_0 \circ \psi(p, q)$ adopta la forma $\mathcal{H}(\theta, I) = h(I)$, per a $\theta \in \mathbb{T}^n$ i $I \in \mathbb{R}^n$. Conseqüentment, la dinàmica governada per $\mathcal{H}$ és trivialment donada per

$$\dot{\theta} = \frac{\partial h}{\partial I}, \quad \dot{I} = 0. \tag{0.5}$$

Les accions I són constants i el moviment en $\mathbb{T}^n$ és una rotació rígida amb freqüència $\omega = \partial_I h(I)$. Si ω és racionalment independent, llavors les trajectòries són denses en $\mathbb{T}^n$. En resum, si s'aplica el teorema de Liouville-Arnold, tot l'espai de fases està foliat per tors invariants i diem que el sistema és integrable en el sentit de Liouville-Arnold. Un exemple clàssic d'un sistema integrable és el problema de dos cossos, també conegut com el *problema de Kepler*—que és de fet *super integrable*, ja que hi ha cinc integrals primeres independents per a un espai de fase de sis dimensions. Ara ens preguntarem sobre sistemes que es poden escriure com una pertorbació d'un de integrable. Això inclou, per exemple, els models restringits dels tres cossos, que de fet van ser un dels principals motius per estudiar aquesta classe de sistemes.

Suposem que tenim un sistema Hamiltonià integrable amb Hamiltonià analític real $h = h(I)$. Quan el Hamiltonianà integrable és pertorbat com

$$\tilde{\mathcal{H}}(\theta, I) = h(I) + \varepsilon f(\theta, I, \varepsilon), \tag{0.6}$$

per algun ε petit i f analíticia real, les solucions del sistema governat per h continuen estant presents d'alguna manera per al sistema governat per $\tilde{\mathcal{H}}$? Estudiar els sistemes Hamiltonians de la forma (0.6) és considerat per molts com el problema fonamental de la dinàmica—que va ser expressat per primera vegada per Poincaré. Aquests sistemes són anomenats sistemes Hamiltonians gairebé integrables i són especialment adequats per a l'aplicació de tècniques de teoria de pertorbacions.

Una nova revolució en l'estudi d'aquests problemes va sorgir d'una competició científica organitzada en nom del Òscar II rei de Noruega i Suècia sobre l'estabilitat del nostre sistema solar. El problema era en realitat demostrar la convergència de les sèries formals de la solució del problema dels n cossos que va estudiar Weierstrass. El premi va ser atorgat a Poincaré pel seu tractat [Poi90]—no per demostrar la convergència de les sèries, sinó pel seu extens i revolucionari treball que es resumeix en demostrar que el problema dels tres cossos no és integrable en el sentit clàssic. De fet, hi va haver un error en el treball presentat a la competició, però aquest només es va descobrir i corregir després d'atorgar el premi. La versió corregida contenia noves idees que Poincaré va continuar desenvolupant en els famosos volums de *Les Méthodes Nouvelles de la Mécanique céleste*. Per tal d'obtenir una visió històrica, vegeu [Dum14].

L'enfocament clàssic per demostrar la integrabilitat consisteix en realitzar canvis canònics (simplèctics) de variables per transformar el Hamiltonià gairebé integrable $\tilde{\mathcal{H}}$ a la forma $\mathcal{H} = h(J)$, on J són les noves accions. Aleshores, la dinàmica és de la forma (0.5) i el nostre Hamiltonià transformat és integrable. Ara il · lustrarem per què aquest procediment formal no es pot esperar que convergeixi sota condicions genèriques. En el temps de Poincaré, aquest procediment es basava en funcions generadores de transformacions canòniques amb variables mixtes, però il · lustrarem les obstruccions amb l'enfocament modern equivalent de les sèries de Lie. Considerem el Hamiltonià (0.6) i suposem que tenim un altre Hamiltonià analític real g. No és difícil mostrar que el Hamiltonià transformat $\mathcal{H} = \tilde{\mathcal{H}} \circ \phi_t^g$, on ϕ^g és el flux per al Hamiltonià g, i per tant un simplectomorfisme, adopta la forma

$$\mathcal{H} = \tilde{\mathcal{H}} + \{g, \tilde{\mathcal{H}}\}t + \{g, \{g, \tilde{\mathcal{H}}\}\}\frac{t^2}{2!} + \{g, \{g, \{g, \tilde{\mathcal{H}}\}\}\}\frac{t^3}{3!} + \dots \tag{0.7}$$

on hem utilitzat (0.4). Com que la perturbació és de mida ε, buscarem un Hamiltonià εg tal que en les noves coordenades $(\varphi, J) = \phi_t^{\varepsilon g}(\theta, I)$, $\mathcal{H} = \tilde{\mathcal{H}} \circ \phi_t^{\varepsilon g}$ sigui integrable a a primer ordre en ε. Observem que a partir de (0.7) obtenim

$$\mathcal{H} = h + \varepsilon\left(f + \{g, h\}t\right) + \mathcal{O}(\varepsilon^2). \tag{0.8}$$

Podem agafar ara $t = 1$, és a dir, el nostre canvi simplèctic de coordenades és l'aplicació de temps-1 del Hamiltonià εg, i requerir que g satisfaci l'equació homològica

$$f + \{g, h\} = 0,$$

de manera que $\mathcal{H}$ sigui integrable fins a l'ordre un en ε. Escrivim ara f i g com la següent sèrie de Fourier

$$f(\varphi, J) = \sum_{k\in\mathbb{Z}^n} \hat{f}_k(J)e^{2\pi i k\cdot\varphi}, \quad g(\varphi, J) = \sum_{k\in\mathbb{Z}^n} \hat{g}_k(J)e^{2\pi i k\cdot\varphi},$$

on $\boldsymbol{i}$ és la unitat imaginària. Aleshores, si utilitzem les expressions anteriors per a f i g, la solució formal a l'equació homològica (0.8) ve donada per

$$\hat{g}_k = \frac{-\hat{f}_k}{2\pi \boldsymbol{i} k \cdot \omega}, \tag{0.9}$$

on recordem que $\omega = \omega(J) = \frac{\partial h}{\partial J}(J)$ depèn de les accions J. Anomenem $\omega(J)$ l'apliacació de freqüències. Repetiríem aleshores el procés de normalització per "empènyer" els termes d'ordre ε^2 als termes d'ordre ε^k per a $k > 2$ i esperaríem eliminar tota dependència de l'Hamiltonià transformat en els angles després d'un nombre infinit de transformacions.

Veiem immediatament que si volem tenir alguna esperança que el procediment formal convergeixi, necessitem que $f_0 = 0$, és a dir, necessitem que la mitjana de f sobre $\mathbb{T}^n$ sigui zero. Suposarem això sense perdre generalitat. Observem que sempre podem considerar $f = \langle f \rangle + \tilde{f}$, on $\langle \cdot \rangle$ denota la mitjana sobre $\mathbb{T}^n$ i $\langle \tilde{f} \rangle = 0$, i reemplaçar h pel Hamiltonià integrable $h + \varepsilon \langle f \rangle$. També necessitem la condició de no ressonància $k \cdot \omega \neq 0$ per evitar divisors iguals a zero. A més, si inspeccionem (0.9) més detingudament, veiem que fins i tot per a ω no ressonant el denominador a (0.9) pot ser arbitràriament a prop de zero. Aquest és el problema dels *petits divisors.* Aquests petits divisors constitueixen una obstrucció a la integrabilitat clàssica, ja que les normes dels coeficients de Fourier (0.9) poden créixer per a conjunts oberts de condicions inicials en les accions, fent que la sèrie (0.9) no convergeixi uniformement a una funció analítica real.

El problema dels petits divisors va ser resolt per primera vegada per Siegel en [Sie42] per al que s'anomena problema de linearització o simplement el problema de Siegel. En la seva prova, Siegel va utilitzar condicions diofàntiques per controlar quantitativament els petits divisors que apareixen en el problema. Parlant en termes generals, el vector de freqüència $\omega \in \mathbb{R}^d$ és diofàntic si els nombres racionals l'aproximen malament de manera quantitativa. Volem subratllar que normalment hi ha dos tipus de condicions diofàntiques:

$$\mathcal{D}_{cont}(\gamma, \tau) := \{\omega \in \mathbb{R}^d : |k \cdot \omega| \geq \gamma |k|^{-\tau}\} \quad \forall k \in \mathbb{Z}^d \setminus \{0\}, \tag{0.10}$$
$$\mathcal{D}_{disc}(\gamma, \tau) := \{\omega \in \mathbb{R}^d : |k \cdot \omega - n| \geq \gamma |k|^{-\tau}\} \quad \forall k \in \mathbb{Z}^d \setminus \{0\},\ n \in \mathbb{Z}, \tag{0.11}$$

ambdues per a algun $\gamma > 0$ i $\tau \geq d - 1$ per el conjunt $\mathcal{D}_{cont}$ i $\tau \geq d$ per el conjunt $\mathcal{D}_{disc}$. La condició (0.10) apareix freqüentment quan treballem amb sistemes dinàmics continus, mentre que la condició (0.11) és típica del cas discret. Observem que, si suposem que $\omega(J)$ a (0.9) és de la forma $\mathcal{D}_{cont}(\gamma, \tau)$, aleshores del decaïment exponencial de $\hat{f}_k$ per l'analiticitat de f resulta que la sèrie (0.9) convergeix absolutament—per a J tal que $\omega(J) \in \mathcal{D}_{cont}(\gamma, \tau)$. Podríem dir que la convergència té lloc en el conjunt d'accions

$$\mathcal{D}^J(\gamma, \tau) := \{J \in \mathbb{R}^n : |k \cdot \omega(J)| \geq \gamma |k|^{-\tau}\} \quad \forall k \in \mathbb{Z}^d \setminus \{0\}.$$

En el que segueix, requerirem que $\omega(J)$ sigui no degenerat, és a dir, $\det |\frac{\partial \omega}{\partial p}| \neq 0$, de manera que l'aplicació de freqüències sigui un difeomorfisme local. Subratllem que els conjunts $\mathcal{D}(\gamma, \tau)$ de (0.22)-(0.23) són de tipus Cantor i, com que $\omega(J)$ és un difeomorfisme, també ho és, $\mathcal{D}^J(\gamma, \tau)$.

Recordem que la convergència de la sèrie (0.9) és només el primer pas per transformar el Hamiltonià (0.6) en forma integrable. La manera de realitzar el procediment infinit va ser traçada per Kolmogorov [Kol54] i provada pel seu alumne Arnold [Arn63a] per a

Hamiltonians analítics. L'enfoc seguit per Kolmogorov en [Kol54] consisteix a transformar el Hamiltonià en una forma normal de Kolmogorov

$$K := E + \omega \cdot J + Q(\varphi, J),$$

on $E \in \mathbb{R}$ i $Q = \mathcal{O}(|J|^2)$ per a la qual $\mathbb{T}^n \times \{0\}$ és un tor invariant. Cal notar que Kolmogorov considerava ω fixa, a diferència de considerar-la una aplicació sobre l'espai d'acció. Val la pena esmentar que Arnold va provar el Teorema de Kolmogorov seguint un enfoc lleugerament diferent. Concretament, va construir funcions generadores per transformar el Hamiltonià (0.6), després de k canvis canònics de coordenades, en

$$\mathcal{H}_k(\varphi, J) = h_k(J) + \varepsilon^{2^k} f_k(\varphi, J).$$

El terme ε^{2^k} mostra la convergència quadràtica dels mètodes iteratius de tipus Newton, que permeten superar els petits divisors a cada pas del procediment. El cas de la diferenciabilitat finita va ser provat per Moser utilitzant les tècniques de suavitzat de Nash per superar la pèrdua de derivades en el que ara es coneixen com a mètodes iteratius Nash-Moser [Mos62]. Aquests treballs van donar lloc al que avui es coneix com la teoria KAM, que rep el nom dels seus fundadors Kolmogorov, Arnold i Moser.

Recordem que originalment el desenvolupament de la teoria KAM va ser motivat per problemes de mecànica celeste. No obstant això, un ingredient clau en les proves de Kolmogorov, Arnold i Moser és la suposició que la pertorbació és prou petita. En [Arn62, Arn63b], Arnold va provar la persistència del moviment quasi-periodic en el Problema Pla de Tres Cossos per a una raó prou petita del semieix major. No obstant això, com va assenyalar Hénon [Hén66], aquests resultats, tot i ser de gran importància teòrica, no tenen aplicabilitat als sistemes planetaris realistes a causa de la necessitat que la pertorbació sigui extremadament petita, de l'ordre de 10^{-333} per al treball d'Arnold. Això és una sorpresa, ja que les expansions formals efectuades feia temps pels astrònoms, on van trobar divisors petits, van ser útils per aproximar trajectòries. El fet que els tors invariants persisteixin per a valors grans del paràmetre de la pertorbació va ser fortament suggerit posteriorment per experiments numèrics amb l'arribada de l'ordinador. Potser el treball més famós en aquesta direcció és el de l'aplicació estàndard de Chirikov [Chi79], per la qual, utilitzant el criteri de Greene desenvolupat en [Gre75], no existeix cap cercle invariant amb el nombre d'or com a nombre de rotació per a $\varepsilon > \varepsilon_c = 0.97163540324$. Aquest valor de ε està fora del regim pertorbatiu de la teoria KAM clàssica. No obstant això, durant les darreres dues dècades hi ha hagut una quantitat significativa de treball en la teoria KAM no pertorbativa, que ha permès provar—amb tècniques assistides per ordinador—que per a $\varepsilon = 0.9716$ existeix un cercle invariant d'or [FHL17]. Ara ens endinsarem en aquest enfocament modern de la teoria KAM.

Un avenç important va sorgir precisament per deixar de banda la suposició que la pertorbació és prou petita. Això es va fer a [dlL01, dlLGJV05] amb l'objectiu de desenvolupar el què inicialment es va anomenar teoria KAM sense variables d'acció-angle. Els autors van obtenir teoremes per a l'existència de tors invariants de dimensió màximal en sistemes Hamiltonians. L'existència es va obtenir a partir de la convergència d'un procediment iteratiu en una escala adequada d'espais de Banach. El seu enfocament és conceptualment molt diferent del de la teoria KAM clàssica. Concretament, en lloc de transformar el Hamiltonià, l'estratègia és buscar directament una incrustació (embedding) d'un tor que satisfaci una equació funcional triada per tal d'establir la invariància del tor i per tant anomenada *equació d'invariància.* Il · lustrem aquest enfocament per a un sistema dinàmic discret sobre $\mathbb{R}^{2n}$ (o algun subconjunt obert d'ell) i una aplicació $F : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$. Volem trobar una incrustació $K : \mathbb{T}^n \to \mathbb{R}^{2n}$ tal que

$$F \circ K - K \circ \mathcal{R} = 0, \tag{0.12}$$

on l'aplicació $\mathcal{R} : \mathbb{T}^n \to \mathbb{T}^n$ representa la dinàmica interna en el tor $\mathcal{K} := K(\mathbb{T}^n)$ donada per F. Si l'aplicació F té estructura addicional, com ara que F sigui un simplectomorfisme exacte, podem utilitzar la geometria per resoldre (0.12) de manera eficient. Observem que (0.12) demana que la incrustació K conjugui la dinàmica sobre $\mathcal{K}$ donada per F amb la dinàmica donada per $\mathcal{R}$ sobre $\mathbb{T}^n$.

En comptes de ser una petita pertorbació d'un sistema integrable, el que és necessari en aquest enfocament és que la incrustació satisfaci (0.12) amb un error prou petit. Anomenarem K la parametrització del tor. El punt de vista es coneix generalment com a *el mètode de parametrització* i s'estén molt més enllà de la teoria KAM i de les aplicacions simplèctiques exactes. Una observació important sobre els tipus de teoremes que es poden construir amb aquest enfocament és que segueixen una formulació *a posteriori*: si hi ha una solució aproximada de l'equació d'invariància que compleix algunes condicions de no degeneració, llavors existeix una solució exacta propera. L'existència es conclou a partir d'assegurar que es compleixen certes condicions, les quals es poden comprovar a partir d'alguns dades a priori en un nombre finit de càlculs. En combinació amb tècniques numèriques rigoroses, és possible realitzar proves assistides per ordinador d'existència, com les de l'aplicació estàndard de Chirikov mencionades abans.

L'objectiu d'aquesta tesi és contribuir al punt de vista modern de la teoria KAM sota el mètode de parametrització. Concretament, obtenim teoremes per a l'existència de tors invariants de dimensió baixa i parcialment hiperbòlics i dissenyem algoritmes per al seu càlcul que són adequats per a realitzar proves assistides per ordinador. Ens centrem en sistemes Hamiltonians dependents del temps on la dependència és periòdica o quasi-periòdica, motivats per les pertorbacions del Problema Restringit Circular de Tres Cossos (CRTBP).

L'estructura d'aquesta monografia és la següent. El Capítol 1 proporciona prerequisits i context. En el Capítol 2, dissenyem un mètode de parametrització basat en esquemes KAM per a calcular tors invariants parcialment hiperbòlics i els seus fibrats invariants en sistemes Hamiltonians quasi-periòdics. Busquem tors i fibrats invariants sota aplicacions flux temps T per T adequades, que permeten reduir la dimensió del tors a calcular en un. El recompte d'operacions per manipular funcions creix exponencialment amb el nombre de variables de la parametrització. Per tant, la reducció permesa pels mètodes d'aplicacions de flux és avantatjosa computacionalment. Si la parametrització s'aproxima amb N coeficients de Fourier, el pas iteratiu requereix un emmagatzematge de $\mathcal{O}(N)$ i $\mathcal{O}(N \log N)$ operacions, en lloc dels mètodes de Newton obtinguts directament a partir de la discretització de les equacions d'invarància, que requereixen un emmagatzematge de $\mathcal{O}(N^2)$ i $\mathcal{O}(N^3)$ operacions. Aquest guany en eficiència prové de les propietats geomètriques de l'espai de fase (és a dir, geometria simplèctica), els sistemes (és a dir, simplecticitat exacta), els tors (és a dir, isotropia, Lagrangianitat), així com les propietats dinàmiques (és a dir, reducibilitat). En particular, la reducibilitat de la linealització de la dinàmica al voltant del tor a una matriu triangular per blocs es coneix habitualment com a reducció automàtica, i és una propietat important tant en teoria com en aplicacions. Els algoritmes s'han implementat i aplicat en el Problema Restringit El · líptic de Tres Cossos per calcular un gran conjunt de tors invariants 3-dimensional no resonants amb els seus fibrats invariants. Els continguts d'aquest capítol han estat publicats a [FMHM24b].

Al Capítol 3, ens basem als mètodes descrits en el Capítol 2 i obtenim un teorema a posteriori per a tors invariants parcialment hiperbòlics i els seus fibrats invariants de rang 1 en sistemes Hamiltonians quasi-periòdics. L'enfocament emprat permet que el teorema sigui aplicat a sistemes Hamiltonians autònoms, periòdics i quasi-periòdics i constitueix la demostració de la convergència dels mètodes d'aplicació de flux de [HM21] i del Capítol 2. A més, obtenim simultàniament els fibrats estables i inestables, la qual cosa proporciona una imatge geomètrica clara de l'espai tangent al tor. La demostració es fonamenta en propietats geomètriques de caire simplèctic que es compleixen aproximadament quan les parametritzacions satisfan les equacions d'invariància de manera aproximada. Hem demostrat lemes geomètrics sobre aquestes propietats que controlen l'error en el procediment iteratiu KAM. L'error nou en les equacions d'invariància es controla mitjançant constants explícites que requereixen un tractament cuidadós de la pèrdua d'analiticitat amb cada pas iteratiu. La demostració conclou obtenint condicions per a la convergència del procediment iteratiu KAM. El teorema a posteriori obtingut és adequat per a realitzar proves assistides per ordinador. Un preprint està disponible a [FMHM24a].

Els tors parcialment hiperbòlics mencionats al capítol 2 tenen varietats invariants associades, les trajectòries de les quals s'apropen exponencialment, ja sigui temps endavant o temps endarrere, al tor. Aquestes varietats s'anomenen *whiskers* ("bigotis"). Al Capítol 4, hem desenvolupat esquemes KAM per a calcular expansions de Fourier-Taylor d'ordre alt dels whiskers en sistemes Hamiltonians autònoms i quasi-periòdics. A diferència dels mètodes que calculen ordre a ordre, que calculen terme a terme els tors i els seus fibrats abans de calcular els whiskers, l'enfocament seguit a aquesta tesi calcula simultàniament tant el tor com els whiskers utilitzant el mateix mètode iteratiu KAM. Aquest marc unificat millora l'eficiència del càlcul dels whiskers, tot duplicant el nombre de termes correctes a l'expansió amb cada iteració. Els continguts d'aquest capítol s'han recollit en el preprint [FMHdlLM24].

Finalment, a l'Annex, incloem un lema sobre la funció primitiva dels simplectomorfismes exactes, algunes proves sobre mitjanes necessàries per al teorema KAM a posteriori, i algunes notes sobre deformacions simplèctiques per fibra que poden tenir interès per si mateixes. Concloem aquesta monografia amb taules que contenen les constants explícites que apareixen en el teorema KAM.

Introduction

Since the Newtonian revolution, many physical problems can be formulated as solutions of differential equations that define *continuous* dynamical systems. Let us introduce a simple setting for such problems. Let $\mathcal{M}$ be some smooth manifold and $X : \mathcal{M} \to T\mathcal{M}$ be some real-analytic vector field. We will be concerned with solutions of ordinary differential equations given by

$$\dot{z} = X(z), \tag{0.13}$$

for $z \in \mathcal{M}$. The flow of (0.13), $\phi : \mathbb{R} \times \mathcal{M} \to \mathcal{M}$, defines a (local) one parameter group of smooth diffeomorphisms $\{\phi_t\}_{t\in\mathbb{R}}$ that dictate the dynamics on $\mathcal{M}$ under X. Another perspective to study the evolution of a system consists on studying the orbits of a map acting on $\mathcal{M}$ as

$$\begin{aligned} F : \quad \mathcal{M} &\longrightarrow \mathcal{M} \\ z &\longmapsto F(z), \end{aligned}$$

where the orbits are given by the iterates of F. Such a dynamical system is said to be *discrete*. It is possible to construct discrete dynamical systems from continuous ones in a straightforward manner. For instance, one could consider the time-1 map ϕ_1 or the Poincaré first return map. Constructing compatible vector fields from maps is a much more subtle question. We emphasize that albeit this thesis concerns mainly non-autonomous Hamiltonian systems, as of now we shall concentrate in *autonomous* systems for the sake of clarity. We simply recall that any time-dependent system can be autonomized by extending the phase space.

We could certainly ask for additional structure on $\mathcal{M}, X$, and F. For instance, in Hamiltonian mechanics, $\mathcal{M}$ is a smooth manifold equipped with a symplectic form $\boldsymbol{\omega}$ which is a differential $2-$form that is closed, i.e., $\mathrm{d}\boldsymbol{\omega} = 0$, and non-degenerate. The non-degeneracy condition implies that $\mathcal{M}$ needs to be even dimensional. We also need a smooth function $H : \mathcal{M} \to \mathbb{R}$, the Hamiltonian. We can then define its associated vector field $X : \mathcal{M} \to T\mathcal{M}$ implicitly from

$$\iota_X \boldsymbol{\omega} = -\mathrm{d}H, \tag{0.14}$$

where $\iota_X \boldsymbol{\omega} = \boldsymbol{\omega}(X, \cdot)$ is the contraction of the symplectic form with the vector field X.

Due to a theorem of Darboux, there always exist canonical coordinates (q, p) that locally transform the symplectic form $\boldsymbol{\omega}$ into the standard symplectic form

$$\boldsymbol{\omega}_o = \mathrm{d}p \wedge \mathrm{d}q = \sum_{i=1}^{d} \mathrm{d}p_i \wedge \mathrm{d}q_i,$$

where $\wedge$ is the wedge product. In these coordinates, the equations of motion given by (0.14) read

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}.$$

These are the classical Hamilton's equations.

From Cartan's magic formula we can immediately see that the flow of X preserves the symplectic structure since

$$\mathcal{L}_X \boldsymbol{\omega} = \mathrm{d}\iota_X \boldsymbol{\omega} + \iota_X \mathrm{d}\boldsymbol{\omega} = 0,$$

where $\mathcal{L}_X$ is the Lie derivative with respect to the vector field X and d is the exterior derivative. The first term is zero because $\iota_X \boldsymbol{\omega}$ is an exact $1-$form and the second one is zero because $\boldsymbol{\omega}$ is closed by assumption. Then for any $t \in \mathbb{R}$, ϕ_t is a symplectomorphism and $\{\phi_t\}_{t\in\mathbb{R}}$ is a one-parameter group of symplectomorphisms. That ϕ_t preserves the symplectic form can also be written as

$$\phi_t^* \boldsymbol{\omega} = \boldsymbol{\omega}, \tag{0.15}$$

where $\phi_t^* \boldsymbol{\omega}$ denotes the pullback of the symplectic form by the diffeomorphism ϕ_t, which we already know is a symplectomorphism.

If we further assume $\boldsymbol{\omega}$ is exact, i.e., there exists a $1-$form $\boldsymbol{\alpha}$ such that $\boldsymbol{\omega} = \mathrm{d}\boldsymbol{\alpha}$, then we observe that

$$\begin{aligned} 0 =& \phi_t^* \mathrm{d}\boldsymbol{\alpha} - \mathrm{d}\boldsymbol{\alpha} \\ =& \mathrm{d}(\phi_t^* \boldsymbol{\alpha} - \boldsymbol{\alpha}), \end{aligned}$$

where we used that the pullback and the exterior derivative commute. Consequently, the $1-$form $\phi_t^* \boldsymbol{\alpha} - \boldsymbol{\alpha}$ is closed. Furthermore, the form $\phi_t^* \boldsymbol{\alpha} - \boldsymbol{\alpha}$ can be shown to be exact, See Appendix A. That is, there exists a function p_t such that $\phi_t^* \boldsymbol{\alpha} - \boldsymbol{\alpha} = \mathrm{d}p_t$. We shall say then that ϕ_t is exact symplectic or, equivalently, an exact symplectomorphism. The structure that we have for our dynamical system has implications to the type of solutions that we can expect and on how we can compute them. In this work, we will mostly be concerned with invariant objects for which the geometry will aid us in proving their existence and in their computation. We shall now show some implications in dynamics of geometric nature.

Consider a real-analytic Hamiltonian function $H : U \to \mathbb{R}$, where $U \subset \mathbb{R}^{2n}$ is some open set and let $X : U \to \mathbb{R}^{2n}$ be the Hamiltonian vector field constructed from (0.14). Let us introduce the Poisson bracket of any two functions $f, g : U \to \mathbb{R}$ as

$$\{f, g\} = \boldsymbol{\omega}(X_f, X_g),$$

where X_f and X_g are Hamiltonian vector fields associated to the functions f and g. In canonical coordinates (q, p), the Poisson bracket is expressed as

$$\{f, g\} = \sum_{i=1}^{n} \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i}.$$

From this expression we see that the variation of a function f along integral curves of the Hamiltonian vector field is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} f \circ \phi_t = \left(\sum_{i=1}^{n} \frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial p_i}\dot{p}_i \right) \circ \phi_t$$

$$= \left(\sum_{i=1}^{n} \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right) \circ \phi_t$$
$$= \{H, f\} \circ \phi_t, \tag{0.16}$$

Consequently, if the Poisson bracket $\{f, H\}$ vanishes f is an integral of motion for the dynamics system defined by H. Note that this conclusion also holds for an arbitrary symplectic form $\boldsymbol{\omega}$ in a coordinate-free formulation as follows

$$\{f, H\} = \boldsymbol{\omega}(X_f, X_H) = (\iota_{X_f}\boldsymbol{\omega})(X_H) = -\mathrm{d}f(X_H) = -X_H(f) = -\mathcal{L}_{X_H} f,$$

where we have used (0.14). Therefore, if $\{f, H\} = 0$, f is an integral of motion for the Hamiltonian flow given by H. The implications of the existence of integrals of motion in a Hamiltonian system are clear from the following well-known theorem [Lio55, Arn89]

Theorem 0.2 (Liouville-Arnold)**.** *Suppose we have a collection of n functions $\{H_i\}_{i=0}^{n-1}$ in involution on a $2n$-dimensional manifold, i.e.,*

$$\{H_i, H_j\} \equiv 0 \qquad \textit{for } i, j = 0, \dots n-1$$

and consider the level set

$$M_f = \{z : H_i = c_i \quad \textit{for } i = 0, \dots, n-1\}.$$

Assume that $\mathrm{d}H_i$ are linearly independent at every point of M_f. Then:

1.) M_f is a smooth manifold, invariant under the Hamiltonian flow of H_0.

2.) If the manifold M_f is compact and connected, then it is diffeomorphic to the $n-$dimensional torus.

3.) The flow of H_0 on M_f is quasi-periodic.

4.) The canonical equations with Hamiltonian H_0 can be integrated by quadratures.

A corollary of the previous theorem is that if M_f is diffeomorphic to $\mathbb{T}^n$, there exist a symplectic transformation into *action-angle* coordinates $(\theta, I) = \psi(q, p)$ such that the new Hamiltonian $\mathcal{H}(\theta, I) = H_0 \circ \psi(p, q)$ adopts the form $\mathcal{H}(\theta, I) = h(I)$, for $\theta \in \mathbb{T}^n$ and $I \in \mathbb{R}^n$. Consequently, the dynamics governed by $\mathcal{H}$ are trivially given by

$$\dot{\theta} = \frac{\partial h}{\partial I}, \quad \dot{I} = 0. \tag{0.17}$$

The actions I are constant and the motion in $\mathbb{T}^n$ is a rigid rotation with frequency $\omega = \partial_I h(I)$. If ω is rationally independent, then trajectories are dense in $\mathbb{T}^n$. In summary, if the Liouville-Arnold theorem applies, the entire phase space is foliated by invariant tori and we say the system is integrable in the sense of Liouville-Arnold. A classical example of an integrable system is the two-body problem also known as the *Kepler problem*—which is in fact *super integrable* since there are five independent first integrals for a six dimensional phase space. Let us now wonder about problems which can be written as a perturbation of integrable problems. This includes, for instance, restricted three body problems, which were in fact one of the main reasons to study these problems.

Suppose we have an integrable Hamiltonian system with real-analytic Hamiltonian $h = h(I)$. When the integrable Hamiltonian is perturbed as

$$\tilde{\mathcal{H}}(\theta, I) = h(I) + \varepsilon f(\theta, I, \varepsilon), \tag{0.18}$$

for some ε small and f real-analytic, are solutions to the system governed by h somehow still present for the system governed by $\tilde{\mathcal{H}}$? Studying Hamiltonian systems of the form (0.18) is considered by many the fundamental problem of dynamics—which was first stated by Poincaré. Such systems are called nearly-integrable Hamiltonian systems and are very suitable for the aplication of perturbation theory techniques.

A new revolution in the study of such problems came from a scientific competition organized in the name of Oscar II king of Norway and Sweden on the stability of our solar system. The problem was actually to prove the convergence of formal series solution of the $n-$body problem that were studied by Weierstrass. The price was awarded to Poincaré from his treatise [Poi90]—not for proving the convergence of the series but for his extensive revolutionary work which summarizes as proving that the three-body problem is not integrable in the classical sense. There was actually a mistake in the work presented to the competition but this was only unveiled, and corrected, after the award. The corrected version contained new ideas that Poincaré continued to develop in the famous volumes of *Les Méthodes Nouvelles de la Mécanique céleste*. For a historical overview see [Dum14].

The classical approach to prove integrability consists in performing canonical (symplectic) changes of variables to transform the nearly-integrable Hamiltonian $\tilde{\mathcal{H}}$ into the form $\mathcal{H} = h(J)$, where J are the new actions. Then, the dynamics is of the form (0.17) and our transformed Hamiltonian is integrable. We shall illustrate now why this formal procedure cannot be expected to converge under generic conditions. In the time of Poincaré, this procedure was based on mixed variables generating functions of canonical transformations, but we will illustrate the obstructions with the more modern equivalent approach of Lie series. Consider the Hamiltonian (0.18) and suppose we have some other real-analytic Hamiltonian g. Then, it is not hard to show that the transformed Hamiltonian $\mathcal{H} = \tilde{\mathcal{H}} \circ \phi_t^g$, where ϕ^g is the flow for the Hamiltonian g, and hence a symplectomorphism, adopts the form

$$\mathcal{H} = \tilde{\mathcal{H}} + \{g, \tilde{\mathcal{H}}\}t + \{g, \{g, \tilde{\mathcal{H}}\}\}\frac{t^2}{2!} + \{g, \{g, \{g, \tilde{\mathcal{H}}\}\}\}\frac{t^3}{3!} + \dots \tag{0.19}$$

where we used (0.16). Since the perturbation is of size ε, we will look for a Hamiltonian εg such that, in the new coordinates, $(\varphi, J) = \phi_t^{\varepsilon g}(\theta, I)$, $\mathcal{H} = \tilde{\mathcal{H}} \circ \phi_t^{\varepsilon g}$ is integrable at order one in ε. Observe that from (0.19) we obtain

$$\mathcal{H} = h + \varepsilon\left(f + \{g, h\}t\right) + \mathcal{O}(\varepsilon^2). \tag{0.20}$$

We can then take $t = 1$, i.e., our symplectic change of coordinates is the time-1 map of the Hamiltonian εg, and require that g satisfies the homological equation

$$f + \{g, h\} = 0,$$

so $\mathcal{H}$ is integrable up to order one in ε. Let us now write f and g as the following Fourier series

$$f(\varphi, J) = \sum_{k \in \mathbb{Z}^n} \hat{f}_k(J) e^{2\pi i k\cdot\varphi}, \quad g(\varphi, J) = \sum_{k \in \mathbb{Z}^n} \hat{g}_k(J) e^{2\pi i k\cdot\varphi},$$

where $\boldsymbol{i}$ is the imaginary unit. Then, if we use the previous expressions for f and g, the formal solution to the homological equation (0.20) is given by

$$\hat{g}_k = \frac{-\hat{f}_k}{2\pi \boldsymbol{i} k \cdot \omega}, \tag{0.21}$$

where recall that $\omega = \omega(J) = \frac{\partial h}{\partial J}(J)$ depends on the actions J. We call $\omega(J)$ the frequency map. We would then repeat the normalization process to "push" the terms of order ε^2 into terms of order ε^k for $k > 2$ and hope to remove all dependence of the transformed Hamiltonian on the angles after an infinite number of transformations.

We immediately see that if we want to have any hope on the formal procedure to converge, we need $f_0 = 0$, i.e., we need the average of f on $\mathbb{T}^n$ to be zero. We shall assume this without loss of generality. Note that we can always consider $f = \langle f \rangle + \tilde{f}$, where $\langle \cdot \rangle$ denotes the average in $\mathbb{T}^n$ and $\langle \tilde{f} \rangle = 0$, and replace h by the integrable Hamiltonian $h + \varepsilon \langle f \rangle$. We also need the non-resonance condition $k \cdot \omega \neq 0$ to avoid zero divisors. Furthermore, if we inspect (0.21) further, even if ω is non-resonant the denominator in (0.21) can be arbitrarily close to zero. This is the so called *small divisors* problem. These small divisors constitute an obstruction to classical integrability since the norms of the Fourier coefficients (0.21) can grow, for open sets of initial conditions in the actions, resulting in the series (0.21) to not converge uniformly to some real-analytic function.

The small divisors problem was first solved by Siegel in [Sie42] for the so called linearization problem or simply the Siegel problem. In his proof, Siegel used Diophantine conditions to control quantitatively the small divisors appearing in the problem. Roughly speaking, the frequency vector $\omega \in \mathbb{R}^d$ is Diophantine if it is badly approximated by rational numbers in a quantitative way. Let us emphasize that there are typically two types of Diophantine conditions:

$$\mathcal{D}_{cont}(\gamma, \tau) := \{\omega \in \mathbb{R}^d : |k \cdot \omega| \geq \gamma |k|^{-\tau}\} \quad \forall k \in \mathbb{Z}^d \setminus \{0\}, \tag{0.22}$$

$$\mathcal{D}_{disc}(\gamma, \tau) := \{\omega \in \mathbb{R}^d : |k \cdot \omega - n| \geq \gamma |k|^{-\tau}\} \quad \forall k \in \mathbb{Z}^d \setminus \{0\},\ n \in \mathbb{Z}, \tag{0.23}$$

both for some $\gamma > 0$ and $\tau \geq d - 1$ for the set $\mathcal{D}_{cont}$ and $\tau \geq d$ for the set $\mathcal{D}_{disc}$. Condition (0.22) frequently appears when working with continuous dynamical systems whereas condition (0.23) is typical in the discrete case. Observe that if we assume $\omega(J)$ in (0.21) to be in $\mathcal{D}_{cont}(\gamma, \tau)$, then the exponential decay of $\hat{f}_k$ from the analyticity of f results in the series (0.21) to converge absolutely—for J such that $\omega(J) \in \mathcal{D}_{cont}(\gamma, \tau)$. We could say that convergence takes place in the set of actions

$$\mathcal{D}^J(\gamma, \tau) := \{J \in \mathbb{R}^n : |k \cdot \omega(J)| \geq \gamma |k|^{-\tau}\} \quad \forall k \in \mathbb{Z}^d \setminus \{0\}.$$

In what follows, we will require than $\omega(J)$ is non-degenerate, i.e., $\det |\frac{\partial \omega}{\partial p}| \neq 0$, so the frequency map is a local diffeomorphism. We emphasize that the sets $\mathcal{D}(\gamma, \tau)$ from (0.22)-(0.23) are Cantor-like and, since $\omega(J)$ is a diffeomorphism, so is $\mathcal{D}^J(\gamma, \tau)$.

Recall that the convergence of the series (0.21) is only the first step to transform the Hamiltonian (0.18) into integrable form. How to carry the infinite procedure was first outlined by Kolmogorov [Kol54] and proved by his student Arnold [Arn63a] for analytic Hamiltonians. The approach followed by Kolmogorov in [Kol54] consists in transforming the Hamiltonian into a Kolmogorov normal form

$$K := E + \omega \cdot J + Q(\varphi, J),$$

where $E \in \mathbb{R}$ and $Q = \mathcal{O}(|J|^2)$ for which $\mathbb{T}^n \times \{0\}$ is an invariant torus. Note that Kolmogorov considered ω to be fixed as opposed to considering it a map over the action space. It is worth mentioning that Arnold proved the Theorem of Kolmogorov following a slightly different approach. Namely, he constructed generating functions to transform the Hamiltonian (0.18), after k canonical changes of coordinates, into

$$\mathcal{H}_k(\varphi, J) = h_k(J) + \varepsilon^{2^k} f_k(\varphi, J).$$

The term ε^{2^k} shows the quadratic convergence of Newton-like iterative methods which allows to overcame the small divisors at every step of the procedure. The case of finite differentiability was proved by Moser using the smoothing techniques of Nash to overcome the loss of derivatives in what are now called Nash-Moser iterative methods [Mos62]. These works gave birth to what is today known as KAM theory after their founders Kolmogorov, Arnold, and Moser.

Let us recall that originally KAM theory was motivated by problems in celestial mechanics. However, a key ingredient in the proofs of Kolmogorov, Arnold, and Moser is the assumption that the perturbation is small enough. In [Arn62, Arn63b], Arnold proved the persistence of quasi-periodic motion in the Planar Three-Body Problem for a small enough ratio of the semi-major axis. However, as pointed out by Hénon [Hén66], these results, albeit of significant theoretical importance, lack the applicability to realistic planetary systems due to requiring an extremely small size for the perturbation—which is of order 10^{-333} for the work or Arnold. This comes as a surprise since formal procedures carried out long ago by astronomers, where they encountered small divisors, proved to be useful in approximating trajectories. The fact that invariant tori persist for large values of the perturbation parameter was subsequently strongly suggested by numerical experiments with the advent of the computer. Perhaps the most famous work in this direction is that of the Chirikov standard map [Chi79], for which, using the Greene criterion developed in [Gre75], no golden invariant circle exist for $\varepsilon > \varepsilon_c = 0.97163540324$. Such value of ε is beyond the perturbative regime of classical KAM theory. Nonetheless, over the last two decades, there has been a significant amount of work in non-perturbative KAM theory that allowed to prove—with computer-assisted techniques—that for $\varepsilon = 0.9716$ there exist a golden invariant circle [FHL17]. We shall now delve into this modern approach to KAM theory.

A significant breakthrough came precisely from dropping the assumption on the perturbation being small enough. This was done in [dlL01, dlLGJV05] for what was originally called KAM theory without action-angle variables. The authors obtained theorems for the existence of invariant *maximal* dimensional tori in Hamiltonian systems where existence is obtained from the convergence of an iterative procedure in a suitable scale of Banach spaces. Their approach is conceptually very different from that of classical KAM theory. Namely, instead of transforming the Hamiltonian, the strategy is to look directly for an embedding of a torus satisfying a functional equation. This equation implies invariance of the manifold parameterized by the embedding. Let us illustrate this approach for a discrete dynamical system on $\mathbb{R}^{2n}$ (or some open subset of it) and a map $F : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$. We want to find an embedding $K : \mathbb{T}^n \to \mathbb{R}^{2n}$ such that

$$F \circ K - K \circ \mathcal{R} = 0, \tag{0.24}$$

where the map $\mathcal{R} : \mathbb{T}^n \to \mathbb{T}^n$ represents the internal dynamics in the torus $\mathcal{K} := K(\mathbb{T}^n)$ under F. If the map F has additional structure—such as F being an exact symplectomorphism—we can use geometry to solve (0.24) efficiently. Observe that (0.24) asks that the embedding K conjugates the dynamics on $\mathcal{K}$ under F to the dynamics $\mathcal{R}$ on $\mathbb{T}^n$.

In exchange of the small enough perturbation, what is necessary in this approach is that the embedding satisfies (0.24) with an error that is small enough. We shall call K the parameterization of the torus. The approach is generally known as *the parameterization method* and extends well beyond KAM theory and exact symplectic maps. An important remark about the kind of theorems that are possible to construct with this approach is that they follow an *a-posteriori* formulation: if there is an approximate solution of the invariance equation satisfying some non-degeneracy conditions, then there exists a true solution nearby. Concluding existence follows from ensuring that certain conditions hold, which can be checked starting from some a-priori data in a finite number of computations. In combination with rigorous numerical techniques, it is possible to perform computer-assisted proofs of existence like those mentioned previously of the Chirikov standard map.

The objective of this thesis is to contribute to the modern approach on KAM theory under the parameterization method. In particular, we obtain theorems for the existence for lower-dimensional, partly hyperbolic invariant tori and design algorithms for their computation that are suitable to conduct computer-assisted proofs. Motivated by perturbations of the Circular Restricted Three-Body Problem (CRTBP), we focus in time-dependent Hamiltonian systems where the dependency is periodic or quasi-periodic.

The structure of this monograph is the following. Chapter 1 provides the necessary background material and the setting. In Chapter 2, we design a parameterization method based on KAM schemes to compute partly hyperbolic invariant tori and their invariant bundles in quasi-periodic Hamiltonian systems. We look for tori and bundles invariant under suitable time$-T$ maps that allow the reduction of the torus dimension to be computed by one. The operation count to manipulate functions grows exponentially with the number of variables of the parameterization. Therefore, the reduction allowed by flow map methods is computationally advantageous. If the parameterization is approximated with N Fourier coefficients, the iterative step requires $\mathcal{O}(N)$ storage and $\mathcal{O}(N \log N)$ operations as opposed to Newton methods ontained from directly discretizing the invariance equations, that require $\mathcal{O}(N^2)$ storage and $\mathcal{O}(N^3)$ operations. This gain in efficiency comes from the geometric properties of the phase space (i.e. symplectic geometry), the systems (i.e. exact symplecticity), the tori (i.e. isotropy, Lagrangianity), as well as from dynamical properties (i.e. reducibility). In particular, the reducibility of the linearized dynamics around the torus to a block triangular matrix is commonly known as automatic reducibility and it is an important property both in theory and applications. The algorithms have been implemented and applied in the Elliptic Restricted Three Body problem to compute a large set of non-resonant $3-$dimensional invariant tori with their invariant bundles. The contents of this chapter have been published in [FMHM24b].

In Chapter 3, we build on the methods described in Chapter 2 and obtain an a-posteriori theorem for partly hyperbolic invariant tori and their rank-1 invariant bundles in quasi-periodic Hamiltonian systems. The approach followed allows the theorem to be applied to autonomous, periodic, and quasi-periodic Hamiltonian systems and constitutes the proof of the convergence of the flow map methods of [HM21] and Chapter 2. Additionally, we obtain both the stable and unstable bundles simultaneously, which provides a clear geometrical picture of the tangent space on the torus. The proof relies on geometric properties of symplectic flavor that hold approximately when the parameterizations satisfy their invariance equations approximately. We obtain geometric lemmas for such properties that control the error in the KAM iterative procedure. The new error in the invariance equations is controlled by explicit constants which require the careful treatment of the analyticity loss with each iterative step. The proof concludes by obtaining conditions for convergence of the KAM iterative procedure. The obtained a-posteriori theorem is suitable to conduct computer-assisted proofs. A preprint is available in [FMHM24a].

Partly hyperbolic invariant tori have associated invariant manifolds whose trajectories approach exponentially, either in the future or in the past, the torus—these invariant manifolds are known as *whiskers*. In Chapter 4, we develop KAM schemes to compute high order Fourier-Taylor expansions of the whiskers in autonomous and quasi-periodic Hamiltonian systems. As opposed to the order-by-order methods, which compute the torus and its bundles before computing the whiskers term by term, the followed approach simultaneously computes both the torus and whiskers using the same KAM iterative method. This unified framework enhances the efficiency of whisker computation by doubling the number of correct terms in the expansion with each iteration. The contents of this chapter have been gathered in the preprint [FMHdlLM24].

Lastly, in the Appendix, we include a lemma on the primitive function of exact symplectomorphisms, some proofs on averages necessary for the a-posteriori KAM theorem, and some notes on fiberwise symplectic deformations that could be of independent interest. We conclude this monograph with tables containing the explicit constants appearing in the KAM theorem.

Chapter 1

Preliminaries and setting

In this chapter, we provide the necessary background material and introduce the notation used throughout the thesis. The reader may wish to proceed directly to Chapter 2 and return to this chapter for reference as needed.

1.1 Basic notation

We assume all objects to be sufficiently smooth, even real-analytic. Furthermore, we consider the vector spaces $\mathbb{R}^m$ and $\mathbb{C}^m$ equipped with the norm

$$|v| = \max_{i=1,\dots,m} |v_i|,$$

which induces the following norm in the spaces of $n_1 \times n_2$ matrices with components in $\mathbb{R}$ or $\mathbb{C}$

$$|M| = \max_{i=1,\dots,n_1} \sum_{j=1,\dots,n_2} |M_{ij}|.$$

We denote by $M^\top$ the transpose of the matrix M. Also, the $n \times n$ identity and zero matrices are denoted by I_n and $\mathcal{O}_n$, respectively. Meanwhile, the $n-$dimensional zero vector is denoted by 0_n. When it is clear from the context, empty blocks denote zero matrices of suitable dimensions. Let $\mathbb{T}^\ell = \mathbb{R}^\ell/\mathbb{Z}^\ell$ be the standard $\ell-$dimensional torus. For $\alpha \in \mathbb{R}^\ell$, we define the rotation map $\mathcal{R}_\alpha : \mathbb{T}^\ell \to \mathbb{T}^\ell$ as

$$\mathcal{R}_\alpha(\varphi) = (\varphi + \alpha) \mod 1.$$

Similarly, for $(\omega, \alpha) \in \mathbb{R}^{d+\ell}$, we define the rotation map $\mathcal{R}_{\omega\alpha} : \mathbb{T}^{d+\ell} \to \mathbb{T}^{d+\ell}$ as

$$\mathcal{R}_{\omega\alpha}(\theta, \varphi) = (\theta + \omega, \varphi + \alpha) \mod 1.$$

We will also use the notation $\bar{\theta} := \theta + \omega$ and $\bar{\varphi} := \varphi + \alpha$.

1.2 Symplectic setting

We assume we have an exact symplectic form $\boldsymbol{\omega}$, defined on an open set $U \subset \mathbb{R}^{2n}$, with matrix representation $\Omega : U \to \mathbb{R}^{2n\times 2n}$. We also assume we have an analytic quasi-periodic Hamiltonian $H : U \times \mathbb{T}^\ell \to \mathbb{R}$ with frequencies $\hat{\alpha} \in \mathbb{R}^\ell$. We can then construct

$X : U \times \mathbb{T}^\ell \to \mathbb{R}^{2n}$ as

$$X(z,\varphi) := \Omega(z)^{-1}\mathrm{D}_z H(z,\varphi)^\top,$$

and on $U \times \mathbb{T}^\ell$, we can define the vector field $\tilde{X} : U \times \mathbb{T}^\ell \to \mathbb{R}^{2n} \times \mathbb{R}^\ell$ as follows

$$\begin{pmatrix} \dot{z} \\ \dot{\varphi} \end{pmatrix} = \tilde{X}(z,\varphi) := \begin{pmatrix} X(z,\varphi) \\ \hat{\alpha} \end{pmatrix}. \tag{1.1}$$

Note that when $\varphi \in \mathbb{T}$, our system reduces to the periodic case. We can think of the extended phase space $U \times \mathbb{T}^\ell$ as a bundle with base $\mathbb{T}^\ell$ and of $\tilde{X}$ as a fiberwise Hamiltonian vector field. In each fiber $\mathcal{F}_\varphi = \pi^{-1}(\varphi)$, where π is the bundle projection, we have a symplectic vector structure and a Hamiltonian vector field $X(\cdot,\varphi) : U \to \mathbb{R}^{2n}$. Note that all these vector fields are coupled by the phase equation $\dot{\varphi} = \hat{\alpha}$. We will see our objects inherit this fiberwise structure in different contexts.

Let us denote the flow associated to $\tilde{X}$ by $\tilde{\phi} : D \subset \mathbb{R} \times U \times \mathbb{T}^\ell \to U \times \mathbb{T}^\ell$, where D is the open set domain of definition of the flow. That is, for every (z,φ) let $I_{z,\varphi} \subset \mathbb{R}$ be the maximal interval of existance of $\tilde{\phi}(\cdot, z, \varphi)$, then

$$D = \{(t,z,\varphi) \in \mathbb{R} \times U \times \mathbb{T}^\ell \mid t \in I_{z,\varphi}\}.$$

Notice that $\tilde{\phi}$ adopts the form of a skew product flow

$$\tilde{\phi}(t,z,\varphi) = \begin{pmatrix} \phi(t,z,\varphi) \\ \psi(t,\varphi) \end{pmatrix}, \tag{1.2}$$

where on the torus we have a linear flow with the frequencies of the Hamiltonian, i.e.,

$$\psi(t,\varphi) = \varphi + \hat{\alpha} t \tag{1.3}$$

and the evolution operator ϕ satisfies

$$\frac{\partial}{\partial t}\phi(t,z,\varphi) = X\big(\phi(t,z,\varphi), \psi(t,\varphi)t\big), \quad \phi(0,z,\varphi) = z.$$

From this point onwards, we will adopt the standard notations $\tilde{\phi}_t(z,\varphi) = \tilde{\phi}(t,z,\varphi)$, $\phi_t(z,\varphi) = \phi(t,z,\varphi)$, and $\psi_t(\varphi) = \psi(t,\varphi)$.

Since the symplectic form is exact, $\boldsymbol{\omega} = \mathrm{d}\boldsymbol{\alpha}$, where $\boldsymbol{\alpha}$ is the action 1-form. Then, the matrix representation of $\boldsymbol{\omega}$ is given by

$$\Omega(z) = \mathrm{D}a(z)^\top - \mathrm{D}a(z), \tag{1.4}$$

where $a : U \to \mathbb{R}^{2n}$ and $a(z)^\top$ is the matrix representation of $\boldsymbol{\alpha}$ at z. Also, because $\boldsymbol{\omega}$ is exact symplectic, ϕ_t is a fiberwise exact symplectomorphism.. That is, for each $\varphi \in \mathbb{T}^\ell$ and fixed t, ϕ_t satisfies symplecticity

$$\mathrm{D}_z\phi_t(z,\varphi)^\top \Omega\big(\phi_t(z,\varphi)\big)\mathrm{D}_z\phi_t(z,\varphi) = \Omega(z)$$

and exactness

$$a\big(\phi_t(z,\varphi)\big)^\top \mathrm{D}_z\phi_t(z,\varphi) - a(z)^\top = \mathrm{D}_z p_t(z,\varphi),$$

for some *primitive function* $p_t : U \times \mathbb{T}^\ell \to \mathbb{R}$, see Appendix A for a explicit form of p_t and its properties. The existence of the primitive function for the fiberwise exact symplectomorphism ϕ_t allows certain cancellations that are crucial in the iterative scheme for the computation of parameterizations of invariant tori and bundles.

Let us also assume we have an almost-complex structure $\boldsymbol{J}$ on U, compatible with the symplectic form, with matrix representation $J : U \to \mathbb{R}^{2n\times 2n}$. That is, $\boldsymbol{J}$ is anti-involutive and symplectic—in coordinates, these properties read

$$J(z)^2 = -I_{2n}, \quad J(z)^\top \Omega(z) J(z) = \Omega(z).$$

The almost complex structure and the symplectic form induce a Riemannian metric $\boldsymbol{g}$ with matrix representation $G : U \to \mathbb{R}^{2n\times 2n}$ constructed as $G(z) = -\Omega(z)J(z)$.

A particular example of a compatible triple $(\boldsymbol{\omega}, \boldsymbol{J}, \boldsymbol{g})$ is when $\boldsymbol{\omega}$ is the standard symplectic form, i.e.,

$$\Omega_0(z) = \begin{pmatrix} \mathcal{O}_n & -I_n \\ I_n & \mathcal{O}_n \end{pmatrix}, \tag{1.5}$$

$\boldsymbol{g}$ is the Euclidean metric, and $\boldsymbol{J}$ is given by the standard linear complex structure on $\mathbb{R}^{2n}$ coming from the complex structure on $\mathbb{C}^n$.That is,

$$G_0(z) = \begin{pmatrix} I_n & \mathcal{O}_n \\ \mathcal{O}_n & I_n \end{pmatrix}, \quad J_0(z) = \begin{pmatrix} \mathcal{O}_n & -I_n \\ I_n & \mathcal{O}_n \end{pmatrix}.$$

1.3 Analytic setting

We assume we have a real analytic Hamiltonian. Therefore, we will work with real analytic functions defined in complex neighborhoods of real domains. Let us construct the complex neighborhood of $\mathbb{T}^\ell$ as the following complex strip of width ρ

$$\mathbb{T}^\ell_\rho = \{\theta \in \mathbb{C}^\ell/\mathbb{Z}^\ell : |\mathrm{Im}\ \theta| < \rho\}.$$

Furthermore, assume we have a holomorphic extension $K_0 : \mathbb{T}^d_r \times \mathbb{T}^\ell_r \to \mathbb{C}^{2n}$ of a map $K_0 : \mathbb{T}^d \times \mathbb{T}^\ell \to \mathbb{R}^{2n}$, for some $d > 0$ and $r > 0$. Let us then construct the sets $\mathcal{U} \subset \mathbb{C}^{2n}$, $\tilde{\mathcal{U}} \subset \mathbb{C}^{2n} \times \mathbb{T}^\ell_r$, and $\tilde{\mathcal{U}}_\varphi \subset \mathbb{C}^{2n}$ as

$$\begin{aligned} \mathcal{U} &:= \{z \in \mathbb{C}^{2n} : \exists(\theta,\varphi) \in \mathbb{T}^d_r \times \mathbb{T}^\ell_r : |z - K_0(\theta,\varphi)| < R\}, \\ \tilde{\mathcal{U}} &: = \{(z,\varphi) \in \mathbb{C}^{2n} \times \mathbb{T}^\ell_r : \exists\theta \in \mathbb{T}^d_r : |z - K_0(\theta,\varphi)| < R\} \\ \tilde{\mathcal{U}}_\varphi &: = \{z \in \mathbb{C}^{2n} : \exists\theta \in \mathbb{T}^d_r : |z - K_0(\theta,\varphi)| < R\}, \end{aligned}$$

for some $R > 0$. For analytic functions $f : \mathcal{U} \to \mathbb{C}$, we consider the norm

$$\|f\|_{\mathcal{U}} = \sup_{z\in\mathcal{U}} |f(z)|, \tag{1.6}$$

that extends naturally for matrix-valued maps $M : \mathcal{U} \to \mathbb{C}^{n_1\times n_2}$, with M_{ij} analytic, as

$$\|M\|_{\mathcal{U}} = \max_{i=1,\ldots,n_1} \sum_{j=1,\ldots,n_2} \|M_{ij}\|_{\mathcal{U}}. \tag{1.7}$$

For their derivatives, we define the norms

$$\|\mathrm{D}^k f\|_{\mathcal{U}} = \sum_{l_1,\ldots,l_k} \left\| \frac{\partial^k f}{\partial z_{l_1} \ldots \partial z_{l_k}} \right\|_{\mathcal{U}}, \quad \|\mathrm{D}^k M\|_{\mathcal{U}} = \max_{i=1,\ldots,n_1} \sum_{j=1,\ldots,n_2} \|\mathrm{D}^k M_{ij}\|_{\mathcal{U}}, \tag{1.8}$$

where each index $l_1, \dots, l_k$ goes from 1 to $2n$. Furthermore, the action of the $k-$order derivative on $v_1, \dots, v_k \in \mathbb{C}^{n_2}$ satisfies the bounds

$$\|\mathrm{D}^k M(z)[v_1, \dots, v_k]\|_{\mathcal{U}} \leq \|\mathrm{D}^k M\|_{\mathcal{U}} \, |v_1| \dots |v_k| \, ,$$

where

$$\begin{aligned}\left(\mathrm{D}^k M(z)[v_1, \dots, v_k]\right)_{ij} &= \mathrm{D}^k M_{ij}(z)[v_1, \dots, v_k] \\ &= \sum_{l_1, \dots, l_k} \frac{\partial^k M_{ij}}{\partial_{z_{l_1}}, \dots, \partial_{z_{l_k}}}(z) \big(v_1\big)_{l_1}, \dots, \big(v_k\big)_{l_k}.\end{aligned}$$

Lastly, for analytic functions $f : \tilde{\mathcal{U}} \to \mathbb{C}$ we consider the norm

$$\|f\|_{\tilde{\mathcal{U}}} = \sup_{\varphi \in \mathbb{T}^\ell_r} \|f(\cdot, \varphi)\|_{\tilde{\mathcal{U}}_\varphi}$$

and the analogous norm to (1.7) for matrix-valued maps.

1.3.1 Spaces of quasi-periodic real-analytic functions

Let us denote by $\mathcal{A}(\mathbb{T}^{d+\ell}_\rho)$ the Banach space of real-analytic functions $f : \mathbb{T}^{d+\ell}_\rho \to \mathbb{C}$—such that $f(\mathbb{T}^{d+\ell}) \subset \mathbb{R}$—that can be continuously extended to the closure $\bar{\mathbb{T}}^{d+\ell}_\rho$ and endowed with the norm

$$\|f\|_\rho = \sup_{(\theta,\varphi) \in \mathbb{T}^{d+\ell}_\rho} |f(\theta, \varphi)|.$$

We will also use the notation $\mathcal{A}(\mathbb{T}^{d+\ell}_\rho)^m$ for the analogous space for functions $f : \mathbb{T}^{d+\ell}_\rho \to \mathbb{C}^m$, and the natural extension of the norm for functions $M : \mathbb{T}^{d+\ell}_\rho \to \mathbb{C}^{n_1 \times n_2}$

$$\|M\|_\rho = \max_{i=1,\dots,n_1} \sum_{j=1,\dots,n_2} \|M_{ij}\|_\rho.$$

For the derivatives of quasi-periodic real-analytic functions, we use Cauchy estimates. For any $\delta \in (0, \rho)$, the derivative operator $\partial_i : \mathcal{A}(\mathbb{T}^{d+\ell}_\rho) \to \mathcal{A}(\mathbb{T}^{d+\ell}_{\rho-\delta})$ satisfies the bounds

$$\|\partial_i f\|_{\rho-\delta} \leq \frac{1}{\delta}\|f\|_\rho, \quad \|\mathrm{D}_\theta f\|_{\rho-\delta} \leq \frac{d}{\delta}\|f\|_\rho, \quad \|\mathrm{D}_\varphi f\|_{\rho-\delta} \leq \frac{\ell}{\delta}\|f\|_\rho.$$

Lastly, for functions $f : \tilde{\mathcal{U}} \to \mathbb{C}$ and $M : \tilde{\mathcal{U}} \to \mathbb{C}^{n_1 \times n_2}$, we define the analogous norms to (1.6), (1.7), and (1.8).

1.4 Function representations

For every function $f : \mathbb{T}^{d+\ell} \to \mathbb{R}$, we consider their Fourier series

$$f(\theta, \varphi) = \sum_{(j,k) \in \mathbb{Z}^{d+\ell}} \hat{f}_{jk} e^{i 2\pi (j\theta + k\varphi)}, \tag{1.9}$$

where

$$\hat{f}_{00} = \int_{\mathbb{T}^{d+\ell}} f(\theta, \varphi) d\theta d\varphi$$

is the average of f that we denote as $\langle f \rangle$.

Remark 1.1. Note that, for analytic f, the Fourier coefficients go to zero exponentially fast when

$$|(j,k)|_1 := \sum_{u=1}^{d} |j_u| + \sum_{v=1}^{\ell} |k_v|$$

goes to infinity.

In a computer, we can only work with a truncation of (1.9). That is, we consider *Fourier representations* of the function f given by

$$f \sim \sum_{(j,k)\in\mathcal{I}} \hat{f}_{jk} e^{i2\pi(j\theta+k\varphi)},$$

where $\mathcal{I}$ is some finite multi-index set with elements in $\mathbb{Z}^{d+\ell}$. A Fourier representation of f is given by the finite set of coefficients $\{\hat{f}_{jk}\}_{(j,k)\in\mathcal{I}}$.

As it is usual in the numerical implementation of the parameterization method in KAM-like contexts (see [HCF+16] for an overview), it is convenient to work also with a grid representation of the function f. For some multi-index set I with elements in $\mathbb{N}^{d+\ell}$, we also consider representations of f of the following form

$$f \sim \{f_{jk}\}_{(j,k)\in\mathrm{I}},$$

where $f_{jk} = f(\theta_j, \varphi_k)$ and $\{(\theta_j, \varphi_k)\}_{(j,k)\in\mathrm{I}}$ is an equally spaced grid of values in $\mathbb{T}^{d+\ell}$ indexed by I. Let N_F be the cardinality of the set $\{f_{jk}\}_{(j,k)\in\mathrm{I}}$, i.e., the number of elements of the grid of $\mathbb{T}^{d+\ell}$. Then, the $\{\hat{f}_{jk}\}_{(j,k)\in\mathcal{I}}$ coefficients can be obtained from the set $\{f_{jk}\}_{(j,k)\in\mathrm{I}}$ through the Discrete Fourier Transform (DFT) and vice-versa through the inverse DFT—at a computational cost of $O(N_F \log N_F)$ operations. We refer to the set of values $\{f_{jk}\}_{(j,k)\in\mathrm{I}}$ as a *grid representation* of the function f.

We will also encounter functions defined in $\mathbb{T}^{d+\ell} \times \mathbb{R}$. A natural representation for these functions is as Fourier-Taylor series. Consider the formal power series

$$f(\theta, \varphi, s) = \sum_{i=0}^{\infty} f_i(\theta, \varphi) s^i, \tag{1.10}$$

where every $f_i : \mathbb{T}^{d+\ell} \to \mathbb{R}$ is 1-periodic in each variable and can be expanded as a Fourier series of the form (1.9). Naturally, we can only work in a computer with truncations of (1.10). Let N be some truncation order. We consider *Fourier-Taylor* representations of f of the following form

$$f \sim \sum_{i=0}^{N} f_i(\theta, \varphi) s^i \sim \sum_{i=0}^{N} \sum_{(j,k)\in\mathcal{I}} \hat{f}_{ijk} e^{i2\pi(j\theta+k\varphi)}. \tag{1.11}$$

Such representation is given by the finite set of coefficients $\{\hat{f}_{ijk}\}_{(j,k)\in\mathcal{I}}^{0\le i\le N}$.

As it is usual in the numerical implementation of the parameterization method in KAM-like contexts (see [HCF+16] for an overview), it is convenient to work with a grid representation of the f_i functions. For some multi-index set I with elements in $\mathbb{N}^{d+\ell}$, we also consider representations of f of the following form

$$f \sim \{f_{ijk}\}_{(j,k)\in\mathrm{I}}^{0\le i\le N},$$

where $f_{ijk} = f_i(\theta_j, \varphi_k)$ and $\{(\theta_j, \varphi_k)\}_{(j,k)\in \mathrm{I}}$ is an equally spaced grid of values in $\mathbb{T}^{d+\ell}$. We refer to the set of coefficients $\{f_{ijk}\}_{(j,k)\in \mathrm{I}}^{0\le i\le N}$ as a *grid-Taylor* representation of f. This representation is better suited, for instance, for the computation of product of functions with variables in $\mathbb{T}^{d+\ell}\times\mathbb{R}$ and for the (numerical) application of the flow and its differential to such functions which can be done using the Jet Transport technique. See the Section 1.6.

Observe that function representations are not unique since we can consider different orders of the Fourier-Taylor series as well as different number of Fourier coefficients for the f_i functions.

1.5 Cohomological equations

We will encounter cohomological equations in $\mathbb{T}^{d+\ell}$ and in $\mathbb{T}^{d+\ell}\times\mathbb{R}$. We dedicate this section to such type of equations that naturally arise in KAM theory. Albeit the theory is well known, see e.g. [Rüs75], we tailor the exposition to be well-suited for our purposes. Let us include the definition of Diophantine vectors for the sake of completeness.

Definition 1.1. The rotation vector $(\omega, \alpha) \in \mathbb{T}^{d+\ell}$ is Diophantine if there exists $\gamma > 0$ and $\tau \ge d+\ell$ such that for all $n\in\mathbb{Z}$ and $(j,k)\in\mathbb{Z}^d\times\mathbb{Z}^\ell\setminus\{0\}$

$$|j\cdot\omega + k\cdot\alpha - n| \ge \frac{\gamma}{(|j|_1+|k|_1)^\tau},$$

where $|\cdot|_1$ is the ℓ^1-norm.

1.5.1 Cohomological equations in $\mathbb{T}^{d+\ell}$

On the torus, we find two types of cohomological equations: small divisors and non-small divisors cohomological equations.

Non-small divisors cohomological equations. For $a, b\in\mathbb{R}$ such that $|a|\ne|b|$, $\omega\in\mathbb{R}^d$, and $\alpha\in\mathbb{R}^\ell$, we define the cohomological operator $\mathfrak{L}^{ab}_{\omega\alpha} : \mathcal{A}(\mathbb{T}^{d+\ell}_\rho)\to\mathcal{A}(\mathbb{T}^{d+\ell}_\rho)$ acting on functions $\xi : \mathbb{T}^{d+\ell}_\rho\to\mathbb{C}$ as

$$\mathfrak{L}^{ab}_{\omega\alpha}\xi := a\xi - b\xi\circ\mathcal{R}_{\omega\alpha}.$$

We also define the Rüssmann operator $\mathfrak{R}^{ab}_{\omega\alpha} : \mathcal{A}(\mathbb{T}^{d+\ell}_\rho)\to\mathcal{A}(\mathbb{T}^{d+\ell}_\rho)$ acting on functions $\eta : \mathbb{T}^{d+\ell}_\rho\to\mathbb{C}$ as the solution of the cohomological equation

$$\mathfrak{L}^{ab}_{\omega\alpha}\xi = \eta. \tag{1.12}$$

That is, $\xi = \mathfrak{R}^{ab}_{\omega\alpha}\eta$ solves (1.12). Expanding ξ and η in their Fourier series, we obtain the converging expansion

$$\left(\mathfrak{R}^{ab}_{\omega\alpha}\eta\right)(\theta,\varphi) = \sum_{(j,k)\in\mathbb{Z}^{d+\ell}}\hat{\xi}_{jk}e^{i2\pi(j\theta+k\varphi)},\quad \hat{\xi}_{jk} = \frac{\hat{\eta}_{jk}}{a-be^{i2\pi(j\omega+k\alpha)}}.$$

Additionally, we have the estimate

$$\|\xi\|_\rho \le \frac{1}{\big||a|-|b|\big|}\|\eta\|_\rho. \tag{1.13}$$

We will also use the notation $\mathfrak{L}^{ab}_{\omega\alpha} = \mathfrak{L}^{ab}$ and $\mathfrak{R}^{ab}_{\omega\alpha} = \mathfrak{R}^{ab}$, when it is clear from the context.

Small divisors cohomological equations. We also define the cohomological operator $\mathfrak{L}_{\omega\alpha} : \mathcal{A}(\mathbb{T}^{d+\ell}_\rho) \to \mathcal{A}(\mathbb{T}^{d+\ell}_\rho)$ as

$$\mathfrak{L}_{\omega\alpha}\xi := \xi - \xi \circ \mathcal{R}_{\omega\alpha},$$

and for any $\delta \in (0, \rho]$, we define the Rüssmann operator $\mathfrak{R}_{\omega\alpha} : \mathcal{A}(\mathbb{T}^{d+\ell}_\rho) \to \mathcal{A}(\mathbb{T}^{d+\ell}_{\rho-\delta})$ as the zero-average solution of

$$\mathfrak{L}_{\omega\alpha}\xi = \eta - \langle\eta\rangle. \tag{1.14}$$

That is $\xi = \mathfrak{R}_{\omega\alpha}\eta$ solves (1.14) with $\langle\xi\rangle = 0$. Again, expanding ξ and η in their Fourier series, we obtain the formal expansion

$$(\mathfrak{R}_{\omega\alpha}\eta)(\theta, \varphi) = \sum_{(j,k)\in\mathbb{Z}^{d+\ell}} \hat{\xi}_{jk} e^{i2\pi(j\theta + k\varphi)}, \quad \hat{\xi}_{jk} = \frac{\hat{\eta}_{jk}}{1 - e^{i2\pi(j\omega + k\alpha)}},$$

where the series is uniformly convergent in a narrower strip if (ω, α) is Diophantine. Moreover, we have the estimate

$$\|\xi\|_{\rho-\delta} \leq \frac{c_{\mathfrak{R}}}{\gamma\delta^\tau}\|\eta\|_\rho, \tag{1.15}$$

where $c_{\mathfrak{R}}$ is the Rüssmann constant. See [Rüs75] for details. We will also use the notations $\mathfrak{L}_{\omega\alpha} = \mathfrak{L}$ and $\mathfrak{R}_{\omega\alpha} = \mathfrak{R}$.

1.5.2 Cohomological equations in $\mathbb{T}^{d+\ell} \times \mathbb{R}$

For $a, b \in \mathbb{R}$ and for functions $\xi, \eta : \mathbb{T}^{d+\ell} \times \mathbb{R} \to \mathbb{R}$, we will encounter the following functional equations

$$a\xi - b\xi \circ \mathcal{R}^\lambda_{\omega\alpha} = \eta, \tag{1.16}$$

where

$$\begin{aligned} \mathcal{R}^\lambda_{\omega\alpha} : \quad \mathbb{T}^{d+\ell} \times \mathbb{R} &\longrightarrow \mathbb{T}^{d+\ell} \times \mathbb{R} \\ (\theta, \varphi, s) &\longmapsto (\theta + \omega, \varphi + \alpha, \lambda s), \end{aligned}$$

for $\lambda \in \mathbb{R}/\{1, -1\}$. If we use Fourier-Taylor representations of the form (1.11) for ξ and η, we obtain—for each order i—the cohomological equation

$$a\xi_i - b\lambda^i \xi_i \circ \mathcal{R}_{\omega\alpha} = \eta_i. \tag{1.17}$$

These equations can be solved as described in Section 1.5.1. Note that for $|a| = |b\lambda^i|$ we have small divisors cohomological equations for the term of order i of the Fourier-Taylor representation of ξ.

1.6 Jet transport

Suppose that for some $W : \mathbb{T}^{d+\ell} \times \mathbb{R} \to \mathbb{R}^{2n}$, we want to compute the flow and its differential on W. That is, we want to compute $\phi_T \circ (W, \mathrm{id})$ and $\mathrm{D}_z\phi_T \circ (W, \mathrm{id})$. More specifically, we want to compute the coefficients $\phi_i, \Phi_i : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$ of the following Fourier-Taylor representations of such objects

$$\phi_T \circ (W, \mathrm{id}) \sim \sum_{i=0}^{N} \phi_i(\theta, \varphi) s^i, \tag{1.18}$$

$$(\mathrm{D}\phi_T \circ (W, \mathrm{id}))\, u \sim \sum_{i=0}^{N} \Phi_i(\theta, \varphi) s^i, \tag{1.19}$$

for some generic $u : \mathbb{T}^{d+\ell} \times \mathbb{R} \to \mathbb{R}^{2n}$. Such computations can be performed with *Jet Transport*—which is a technique to propagate the jet of derivatives of a function. This is done by replacing the floating point arithmetic of a numerical integrator with a (truncated) formal power series arithmetic. In this Section we briefly cover the adaptation of the Jet transport technique to compute the representations (1.18) and (1.19). For details see e.g., [Pér15, GJJC$^+$23].

More specifically, assume we have grid-Taylor representations of W and u, i.e, we have $\{W_{ijk}\}_{(j,k)\in \mathrm{I}}^{0\le i\le N}$ and $\{u_{ijk}\}_{(j,k)\in \mathrm{I}}^{0\le i\le N}$. Recall from Section 1.4 that for the Fourier-Taylor representations

$$W \sim \sum_{i=0}^{N} W_i(\theta, \varphi) s^i, \qquad u \sim \sum_{i=0}^{N} u_i(\theta, \varphi) s^i,$$

we have

$$W_{ijk} = W_i(\theta_j, \varphi_k), \qquad u_{ijk} = u_i(\theta_j, \varphi_k),$$

where $W_{ijk}, u_{ijk} \in \mathbb{R}^{2n}$, and $\{\theta_j, \varphi_k\}_{(j,k)\in \mathrm{I}}$ an equally spaced grid of $\mathbb{T}^{d+\ell}$ indexed by I. Let us define the truncated Taylor series

$$W_{jk}(s) := \sum_{i=0}^{N} W_{ijk} s^i, \qquad u_{jk}(s) := \sum_{i=0}^{N} u_{ijk} s^i.$$

We can then propagate $W_{jk}(s)$ and $u_{jk}(s)$ to obtain the following representations

$$\phi_T \circ (W_{jk}, \mathrm{id}) \sim \sum_{i=0}^{N} \phi_{ijk} s^i,$$

$$(\mathrm{D}\phi_T \circ (W_{jk}, \mathrm{id}))\, u_{jk} \sim \sum_{i=0}^{N} \Phi_{ijk} s^i,$$

for all $(j,k) \in \mathrm{I}$. Therefore, we obtain $\{\phi_{ijk}\}_{(j,k)\in \mathrm{I}}^{0\le i\le N}$ and $\{\Phi_{ijk}\}_{(j,k)\in \mathrm{I}}^{0\le i\le N}$. That is, we obtain grid-Taylor representations for (1.18) and (1.19). We can recover a Fourier-Taylor representation with the inverse DFT. For some implementation details, see Section 4.4.

Chapter 2

Invariant tori and invariant bundles

The study of invariant manifolds constitutes the center piece in understanding dynamical systems. It is a rather natural first approach—and often the only hope—to unveil the qualitative behavior of a time-evolving system. Besides the intrinsic interest of invariant manifolds, such structures have found their "real-world" analogues in celestial mechanics, astrodynamics and mission design, plasma physics, semi-classical quantum theory, magnetohydrodynamics, neuroscience, and the list goes on. In particular, celestial mechanics has a long-held tradition in considering such objects, especially periodic orbits and invariant tori carrying quasi-periodic motion. It is in fact one of the main fields that promoted their rigorous and numerical study.

Astronomers used perturbative techniques for centuries in the form of formal series of dubious convergence due to the existence of small divisors problems. As already mentioned in the Introduction, progress had to wait until the pioneering works [Kol54, Arn63a, Mos62] that gave birth to the celebrated KAM theory and the subsequent proofs on the persistence of invariant tori under small enough perturbations of integrable systems. In later works, the KAM ideas were carried out far from the perturbative regime, without the need of action-angle variables [dlLGJV05], and rigorous results and algorithms were developed for hyperbolic invariant tori and their whiskers [HdlL06a, HdlL06b, HdlL07]. The methodology is based on the solution of functional equations in the spirit of the parametrization method introduced in [CFdlL03a, CFdlL03b, CFdlL05] for invariant manifolds of fixed points. Rather rapidly, the parameterization method lead to a plethora of rigorous results [CCdlL13, CH17b, FdlLS09, GHdlL14, HdlLS12, LV11] and numerical explorations [CCGdlL22, CCH21, CH17a, GHdlL22, HM21, KAdlL22], in different contexts, to name a few. See [HCF$^+$16] for a survey.

In this chapter, we design flow map parameterization methods building on [HM21] to compute non-resonant partially hyperbolic invariant tori and their invariant bundles in quasi-periodic Hamiltonian systems. Time-dependent Hamiltonians appear naturally in astrodynamics and celestial mechanics as improvements of the CRTBP. There is a hierarchy of models of increasing complexity that provide a closer behavior to the real solar system dynamics, which is generally accepted to be given by the Newtonian attraction of the celestial bodies described according to the JPL ephemeris [DTT17]. The improvements of the CRTBP include the ERTBP, the Quasi Bicircular Problem of [And98], and the frequency models of [GMM02], to mention a few. These improved models enable the consideration of advanced space missions concepts and can serve as a seed to compute bounded motion for several decades in the JPL ephemeris model [And03]. The application of the parameterization method ideas to non-autonomous complex models in astrodynamics and celestial mechanics is our main motivation.

Flow maps methods allow the reduction of the torus dimension to be computed by one [GM01, HM21]. Keeping in mind that the operation count to manipulate functions grows exponentially with the number of variables of the parameterization, the reduction allowed by flow map methods is computationally advantageous. This comes at the expense of numerical integration, which can be easily parallelized. A similar argument can be made for using the parameterization method instead of following a normal form approach. Normal forms require to manipulate functions of the same number of variables as the dimension of the phase space. Instead, the parameterization method requires to manipulate functions with as many variables as the dimension of the invariant manifold. Furthermore, the algorithms are very efficient—if the parameterization is approximated with either N sample values in a regular grid or N Fourier coefficients, the Newton-like step requires $\mathcal{O}(N)$ storage and $\mathcal{O}(N \log N)$ operations.

This efficiency comes from the geometrical properties of the phase space (i.e. symplectic geometry), the systems (i.e. exact symplecticity), the tori (i.e. isotropy, Lagrangianity), as well as from dynamical properties (i.e. reducibility). In particular, the reducibility of the linearized dynamics around the torus to a block triangular matrix is commonly known as automatic reducibility and it is an important property both in theory [CCdlL13, CH14, dlLGJV05, FdlLS09, GHdlL14, HdlL06a, HdlL06b, HdlLS12, LV11] and applications [CCGdlL22, HM21, KAdlL22]. All these properties lead to the presence of "magic cancellations" and allow the solution of the so-called small divisors equations, which appear naturally in the algorithms and are the hallmark of KAM theory. Proving these cancellations in our context has required the introduction of new geometrical constructs such as fiberwise symplectic deformations. The technical details can be found in the Appendices B, C, and D.

Finally, a standard practice when working with non-autonomous systems is to consider an extended phase space by defining extra angle variables and conjugated fictitious variables to make the system autonomous. Although this is mathematically equivalent, this incurs in increasing the dimension of the phase space which leads to less efficient algorithms. We avoid this practice by considering appropriate functional equations. Another standard practice to deal with non-autonomous systems, in particular periodic, is to use flow map methods for a time-T map that make the discrete system autonomous, i.e., choosing for T the period of the system. Instead, we take T as one of the internal periods of the torus sought for. This enables continuation from an autonomous approximation of quasi-periodic models starting directly from tori of the autonomous approximation computed via flow map methods. Additionally, with such approach, we can construct *a-posteriori* KAM theorems that work for both autonomous and quasi-periodic Hamiltonian systems. See Chapter 3.

2.1 Invariance equations for invariant tori

We are interested in quasi-periodic solutions for the system given by (1.1), with frequencies $\hat{\omega} \in \mathbb{R}^{d+1}$ and $\hat{\alpha} \in \mathbb{R}^{\ell}$—which correspond to the frequencies of the Hamiltonian. We will refer to $\hat{\omega}$ and to $\hat{\alpha}$ as internal and external frequencies, respectively. Geometrically speaking, quasi-periodic solutions lie in $(d+\ell+1)$-dimensional tori $\tilde{\mathcal{K}} \subset U \times \mathbb{T}^{\ell}$. In the light of the parameterization method, we consider suitable parameterizations $\tilde{K} : \mathbb{T}^{d+\ell+1} \to U \times \mathbb{T}^{\ell}$ that conjugate the dynamics in $\tilde{\mathcal{K}}$ to a linear flow in $\mathbb{T}^{d+\ell+1}$ with frequency $(\hat{\omega}, \hat{\alpha})$. In particular, the frequencies $\hat{\omega}$ and $\hat{\alpha}$ need to be, at least, non-resonant or ergodic. That is,

$$k \cdot \hat{\omega} + j \cdot \hat{\alpha} \neq 0 \quad \text{for} \quad (k, j) \in \mathbb{Z}^{d+1} \times \mathbb{Z}^{\ell} \setminus \{0\},$$

Then, for the range of $\tilde{K}$ to be an invariant torus, the parameterization is required to satisfy the functional equation

$$\tilde{\phi}_t \circ \tilde{K}(\hat{\theta}, \varphi) - \tilde{K}(\hat{\theta} + \hat{\omega} t, \varphi + \hat{\alpha} t) = 0, \tag{2.1}$$

where $\hat{\theta} \in \mathbb{T}^{d+1}, \varphi \in \mathbb{T}^{\ell}$, and $t \in \mathbb{R}$.

Remark 2.1. By differentiating (2.1) with respect to t, we obtain the following vector field version of the invariance equation

$$\tilde{X} \circ \tilde{K} = \mathrm{D}_{\hat{\theta}} \tilde{K} \hat{\omega} + \mathrm{D}_{\varphi} \tilde{K} \hat{\alpha}.$$

Recall that the extended phase space, $U \times \mathbb{T}^{\ell}$, has a bundle structure with $\mathbb{T}^{\ell}$ as the base space. Because of this bundle structure, we consider parameterizations for $\tilde{\mathcal{K}}$ of the form

$$\tilde{K}(\hat{\theta}, \varphi) = \begin{pmatrix} \hat{K}(\hat{\theta}, \varphi) \\ \varphi \end{pmatrix},$$

where $\hat{K} : \mathbb{T}^{d+\ell+1} \to U$ parameterizes a $(d+\ell+1)-$dimensional invariant torus $\hat{\mathcal{K}} \subset U$. Then, for $\tilde{K}$ to satisfy (2.1), it suffices that $\hat{K}$ satisfies

$$\phi_t\Big(\hat{K}(\hat{\theta}, \varphi), \varphi\Big) - \hat{K}(\hat{\theta} + \hat{\omega} t, \varphi + \hat{\alpha} t) = 0, \tag{2.2}$$

for all $t \in \mathbb{R}$. From a computational point of view, the cost to compute parameterizations rapidly increases with the dimension of the torus. We therefore follow the trick from [GM01, HM21] and look for a parameterization $K : \mathbb{T}^{d+\ell} \to U$ of a $(d+\ell)$-dimensional torus $\mathcal{K} \subset \hat{\mathcal{K}}$, invariant under time-$T$ maps where T is the period associated to one of the internal frequencies of $\hat{\mathcal{K}}$. Note that if the frequencies $\hat{\omega}$ are fixed, then so is T.

Let us first define $\hat{\omega} =: \frac{1}{T}(\omega, 1)$ and $\hat{\alpha} =: \frac{1}{T}\alpha$, where $\omega \in \mathbb{R}^d$ and $\alpha \in \mathbb{R}^{\ell}$. In what follows, we will require of (ω, α) stronger non-resonance conditions. We will assume (ω, α) to be Diophantine, see Definition 1.1. We then assume $\hat{\theta} = (\theta, \theta_d)$, for some fixed $\theta_d \in \mathbb{T}$ and with $\theta \in \mathbb{T}^d$, and consider a parameterization for $\mathcal{K}$ of the form $K(\theta, \varphi) = \hat{K}(\theta, \theta_d, \varphi)$. If we look at the invariance equation (2.2) for $t = T$ we observe that

$$\phi_T\Big(\hat{K}(\theta, \theta_d, \varphi), \varphi\Big) - \hat{K}(\theta + \omega, \theta_d + 1, \varphi + \alpha) = 0,$$

from where we obtain the following invariance equation for K

$$\phi_T \circ \Big(K, \mathrm{id}\Big) - K \circ \mathcal{R}_{\omega\alpha} = 0, \tag{2.3}$$

where the map $\mathcal{R}_{\omega\alpha}$ defines the dynamics in $\mathcal{K}$ as

$$\begin{aligned}\mathcal{R}_{\omega\alpha}: \quad \mathbb{T}^d \times \mathbb{T}^\ell &\longrightarrow \mathbb{T}^d \times \mathbb{T}^\ell \\ (\theta, \varphi) &\longmapsto \big(\theta + \omega, \varphi + \alpha\big)\end{aligned}$$

and (K, id) stands for the following bundle map over the identity

$$\begin{aligned}(K, \mathrm{id}): \quad \mathbb{T}^d \times \mathbb{T}^\ell &\longrightarrow U \times \mathbb{T}^\ell \\ (\theta, \varphi) &\longmapsto \big(K(\theta, \varphi), \varphi\big).\end{aligned}$$

We refer to ω and α as the internal and external rotation vectors, respectively. We can recover a parameterization for $\hat{\mathcal{K}}$, and consequently for $\tilde{\mathcal{K}}$, from the parameterization of $\mathcal{K}$ as

$$\hat{K}(\hat{\theta}, \varphi) = \phi_{\theta_d T}\Big(K(\theta - \theta_d\omega, \varphi - \theta_d\alpha), \varphi - \theta_d\alpha\Big). \tag{2.4}$$

In what follows, we refer to $\mathcal{K}$ as the generator of $\hat{\mathcal{K}}$ and to T as the flying time of $\mathcal{K}$. We emphasize that, although we began by considering a $(d + \ell + 1)-$dimensional invariant torus $\tilde{\mathcal{K}}$ living in the extended phase space $U \times \mathbb{T}^\ell$, this torus is completely determined by $(d + \ell)-$dimensional tori $\mathcal{K}$ living in U. Consequently, with this formulation, we not only manage to reduce the dimension of the phase space but also of the invariant tori to be computed.

Remark 2.2. For invariant tori $\hat{\mathcal{K}}$ and $\mathcal{K}$, the parameterizations $\hat{K}$ and K are not unique. For $\hat{a} \in \mathbb{R}^{d+1}$, if $\hat{K}$ satisfies Eq. (2.2), $\hat{K}(\hat{\theta} + \hat{a}, \varphi)$ is also a solution parameterizing the same torus. Similarly, for $a \in \mathbb{R}^d$ and $\tau \in \mathbb{R}$, if K satisfies Eq. (2.3), $\phi_\tau\Big(K(\theta + a, \varphi), \varphi\Big)$ is also a solution generating the same torus as K. Consequently, Eq. (2.2) determines $\hat{K}$ up to a $(d + 1)-$dimensional phase shift and Eq. (2.3) determines K up to a d-dimensional phase shift and up to a translation within the torus $\hat{\mathcal{K}}$. Hence, for both $\hat{\mathcal{K}}$ and $\mathcal{K}$, we have $(d + 1)$ degrees of freedom in their parameterizations.

Remark 2.3. The frequency and the rotation vectors $\hat{\omega}$ and ω, respectively, are defined up to unimodular matrices. To the frequency vector $\hat{A}\hat{\omega}$, with $\hat{A} \in \mathbb{Z}^{(d+1)\times(d+1)}$, corresponds the reparameterization $(\hat{\theta}, \varphi) \mapsto \hat{K}(\hat{A}^{-1}\hat{\theta}, \varphi)$. Similarly, to the rotation vector $A\omega$, with $A \in \mathbb{Z}^{d\times d}$, corresponds the reparameterization $(\theta, \varphi) \mapsto K(A^{-1}\theta, \varphi)$.

Remark 2.4. When the Hamiltonian is periodic with period T_H, it is somewhat standard to look for tori invariant under time-T_H maps, see e.g. [CJ00]. Let us illustrate our motivation for taking time$-T$ maps, where T is the period associated to one of the internal frequencies, with a simple example commonly found in applications. Assume we have a T_H-periodic Hamiltonian of the following form,

$$H(z, \varphi) = H_0(z) + \varepsilon H_1(z, \varphi),$$

where $\varepsilon \in \mathbb{R}$ is some parameter not necessarily small. Also assume that we have computed a parameterization $K_0 : \mathbb{T}^d \to U$ of a d-dimensional generating torus for the autonomous system given by H_0. It is then natural to consider a parameterization $K_\varepsilon : \mathbb{T}^d \times \mathbb{T} \to U$ of a $d + 1$-dimensional generating torus for the system given by H that can be obtained from K_0 by continuation methods in ε. If we take time-T maps where T is the period associated to one of the frequencies of the torus parameterized by K_0, we can construct K_ε at $\varepsilon = 0$ as

$$K_\varepsilon(\theta, \varphi) = K_0(\theta), \tag{2.5}$$

from where we can directly start the continuation in ε. See Section 2.6. The same argument also holds for the quasi-periodic case.

2.2 Invariance equations for bundles of rank 1

In the present work, we focus on $(d + \ell + 1)$-dimensional partially hyperbolic invariant tori with $d = n - 2$. Our choice is motivated by applications such as quasi-periodic perturbations of the Restricted Three Body Problem—this will be the test case explored in Section 2.8. The results can easily be adapted for Lagrangian tori $(d + 1 = n)$ and for the lower dimensional case $(d + 1 < n - 1)$. In the later case, a complication is that partially hyperbolic invariant tori are not necessarily reducible. Nonetheless, there exist strategies [HdlL06a, HdlL07, HdlLS12]. Following the results from Section 2.1, we will directly consider rank 1 bundles of $\hat{\mathcal{K}}$ and $\mathcal{K}$ instead of bundles of $\tilde{\mathcal{K}}$.

Let us first consider bundles of rank 1, $\hat{\mathcal{W}}$, with base $\hat{\mathcal{K}}$, invariant under the linearization of ϕ_t on $\hat{\mathcal{K}}$, where the dynamics on the fibers is uniformly contracting or expanding. We then look for a parameterization $\hat{W} : \mathbb{T}^{d+\ell+1} \to \mathbb{R}^{2n}$ satisfying

$$\mathrm{D}_z\phi_t\Big(\hat{K}(\hat{\theta}, \varphi), \varphi\Big)\hat{W}(\hat{\theta}, \varphi) = \hat{W}(\hat{\theta} + \hat{\omega}t, \varphi + \hat{\alpha}t)e^{t\chi},$$

with $\chi \in \mathbb{R}$. If $\chi < 0$, $\hat{\mathcal{W}} = \hat{\mathcal{W}}^s$ is the stable bundle and if $\chi > 0$, $\hat{\mathcal{W}} = \hat{\mathcal{W}}^u$ is the unstable bundle. We again reduce the dimension of the object by using time$-T$ maps and look instead for a parameterization $W : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$ of a bundle $\mathcal{W}$ with base $\mathcal{K}$ satisfying

$$\mathrm{D}_z\phi_T \circ \Big(K, \mathrm{id}\Big)W - W \circ \mathcal{R}_{\omega\alpha}\lambda = 0, \tag{2.6}$$

with $\lambda = e^{T\chi}$. We can recover the parameterization of the bundle $\hat{\mathcal{W}}$ as

$$\hat{W}(\hat{\theta}, \varphi) = e^{-\theta_d T\chi}\mathrm{D}_z\phi_{\theta_d T}\Big(K(\theta - \theta_d\omega, \varphi - \theta_d\alpha), \varphi - \theta_d\alpha\Big)W(\theta - \theta_d\omega, \varphi - \theta_d\alpha).$$

Note that the dynamics on the fibers is a uniform contraction when $\lambda < 1$ and a uniform expansion when $\lambda > 1$. We will also refer to $\mathcal{W}$ as the generator of $\hat{\mathcal{W}}$, to λ as the Floquet multiplier of $\mathcal{K}$, and to χ as the Floquet exponent of $\hat{\mathcal{K}}$.

Remark 2.5. Rank 1 stable and unstable bundles are reducible (to constant λ). In general, one could consider higher rank bundles that might not be reducible [HdlL06a, HdlL07, HdlLS12].

Remark 2.6. We are implicitly assuming the bundles are trivial or that can be trivialized (e.g., by using the double covering trick [HdlL07]).

Remark 2.7. The invariant bundle $\hat{\mathcal{W}}$ and $\mathcal{W}$ are the linearization on $\hat{\mathcal{K}}$ and $\mathcal{K}$, respectively, of certain manifolds defined on the annulus: the whiskers. We will come back to these objects in Chapter 4.

2.3 Geometric properties of invariant tori

Until now, we have only considered the geometric properties of the phase space. Nevertheless, invariant tori of exact symplectic maps under non-resonance conditions carry certain geometric properties. Because of these geometric properties, there exists a set of coordinates given by a symplectic frame that reduces the linearization of ϕ_T on $\mathcal{K}$ to upper triangular form constant along the diagonal. The reducibility of the linearized dynamics allows the efficient computation of invariant tori and bundles and is a property known as *automatic reducibility* in a variety of contexts.

For each φ, consider tori $\mathcal{K}_\varphi \subset \mathcal{K}$ with parameterizations $K_\varphi : \mathbb{T}^d \to U$ constructed as $K_\varphi(\theta) = K(\theta, \varphi)$. If we think of $U \times \mathbb{T}^\ell$ as the total space of a bundle with base $\mathbb{T}^\ell$ and projection π, each fiber $\mathcal{F}_\varphi = \pi^{-1}(\varphi)$ contains the torus $\mathcal{K}_\varphi$, see Fig. 2.1 for a sketch. Let

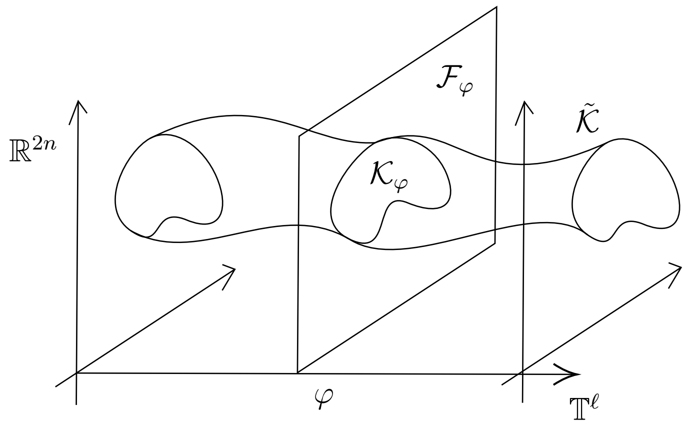

Figure 2.1: Sketch of the torus $\mathcal{K}_\varphi$ within the torus $\tilde{\mathcal{K}}$.

us denote the symplectic form on $\mathcal{K}_\varphi$ as $\Omega_{\mathrm{D}K} : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times 2n}$ and constructed from the pullback of $\boldsymbol{\omega}$ by K_φ as

$$\Omega_{\mathrm{D}K} := (\mathrm{D}K_\varphi)^\top \Omega \circ K_\varphi \mathrm{D}K_\varphi = (\mathrm{D}_\theta K)^\top \Omega \circ K \mathrm{D}_\theta K.$$

Lemma 2.1. *For each $\varphi \in \mathbb{T}^\ell$, $\mathcal{K}_\varphi$ is an isotropic torus.*

Proof. We need to show that for each $\varphi \in \mathbb{T}^\ell$, $\Omega_{\mathrm{D}K} = 0$. That is,

$$(\mathrm{D}_\theta K)^\top \Omega \circ K \mathrm{D}_\theta K = 0 \tag{2.7}$$

for all φ. From the symplecticity of ϕ_T and the invariance of $\mathcal{K}$ we have the identity

$$\Omega \circ K = \Big(\mathrm{D}_z \phi_T(K, \mathrm{id})\Big)^\top \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, \mathrm{D}_z \phi_T(K, \mathrm{id}\Big). \tag{2.8}$$

Also, from the invariance of $\mathcal{K}$, the tangent bundle of $\mathcal{K}_\varphi$, parameterized by the columns of $\mathrm{D}K_\varphi$, is transported by the differential of ϕ_T as

$$\mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\theta K = \mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha}. \tag{2.9}$$

Then, using (2.8) and (2.9), we have

$$(\mathrm{D}_\theta K)^\top \Omega \circ K \mathrm{D}_\theta K = (\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha})^\top \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha},$$

which is constant due to the ergodicity of $(\theta, \varphi) \mapsto (\theta + \omega, \varphi + \alpha)$, coming from the fact that, for fixed (θ, φ), $\{(\theta + j\omega, \varphi + j\alpha)\}_{j\in\mathbb{Z}}$ is dense in $\mathbb{T}^d \times \mathbb{T}^\ell$.

Also, since $\Omega = (\mathrm{D}a)^\top - \mathrm{D}a$, the left-hand side of (2.7) reads

$$\mathrm{D}_\theta \Big(a \circ K\Big)^\top \mathrm{D}_\theta K - (\mathrm{D}_\theta K)^\top \mathrm{D}_\theta \Big(a \circ K\Big) \tag{2.10}$$

and each entry ij of (2.10) can be expressed as

$$\begin{aligned}
&\sum_k \partial_{\theta^i}\Big(a_k\circ K\Big)\partial_{\theta^j}K^k - \partial_{\theta^j}\Big(a_k\circ K\Big)\partial_{\theta^i}K^k \\
=&\sum_k \partial_{\theta^i}\Big(a_k\circ K\partial_{\theta^j}K^k\Big) - a_k\circ K\partial^2_{\theta^i\theta^j}K^k \\
&- \partial_{\theta^j}\Big(a_k\circ K\partial_{\theta^i}K^k\Big) + a_k\circ K\partial^2_{\theta^i\theta^j}K^k \\
=&\sum_k \partial_{\theta^i}\Big(a_k\circ K\partial_{\theta^j}K^k\Big) - \partial_{\theta^j}\Big(a_k\circ K\partial_{\theta^i}K^k\Big).
\end{aligned}$$

Observe that the average of each entry ij of (2.10) is zero since we are taking averages of derivatives with respect to θ. Consequently, since $\Omega_{\mathrm{D}K}$ is constant and with zero average, it is identically zero which concludes the proof. $\square$

We think of $\mathcal{K}$ as a φ-parameterized family of d-dimensional isotropic tori and we say that $\mathcal{K}$ is fiberwise isotropic.

As we will see in Section 2.4, because of the isotropic character of $\mathcal{K}$ and the invariance of $\mathcal{W}$, there exists a set of coordinates given by a symplectic frame $P : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times 2n}$ such that the differential of the flow map on $\mathcal{K}$ reduces to an upper triangular matrix-valued map—constant along the diagonal. That is,

$$(P\circ\mathcal{R}_{\omega\alpha})^{-1}\mathrm{D}_z\phi_T\circ(K,\mathrm{id})P = \left(\begin{array}{c|c} \Lambda & S \\ \hline & \Lambda^{-\top}\end{array}\right), \tag{2.11}$$

where

$$\Lambda := \left(\begin{array}{c|c} I_{n-1} & \\ \hline & \lambda\end{array}\right), \quad \Lambda^{-\top} := \left(\begin{array}{c|c} I_{n-1} & \\ \hline & \lambda^{-1}\end{array}\right), \tag{2.12}$$

and $S : \mathbb{T}^{d+\ell} \to \mathbb{R}^{n\times n}$. The reducibility of the linearized dynamics is commonly known as automatic reducibility and it is an important property both in theory and applications. See e.g. [CCdlL13, CH14, dlLGJV05, FdlLS09, GHdlL14, HdlL06a, HdlL06b, HdlLS12, LV11] for several references on the parameterization method and reducibility in different contexts relatively close to the one of the present paper and [CCGdlL22, HM21, KAdlL22] for recent applications to Celestial Mechanics and Astrodynamics.

2.4 Adapted frames for $\mathcal{K}$

In this section we cover the construction adapted frames P such that the linearized dynamics on the torus reduce to upper triangular form. Furthermore, P can be chosen to be symplectic with respect to the standard symplectic form Ω_0; recall 1.5. That is, P satisfies the following

$$P^\top\Omega\circ KP = \Omega_0. \tag{2.13}$$

Observe that then the inverse of P is straightforward and is given by

$$P^{-1} = -\Omega_0 P^\top\Omega\circ K. \tag{2.14}$$

The symplecticity condition (2.13) shows that the symplectic form restricted to the torus, in the coordinates given by the frame P, reduces to the standard symplectic form. It also suggests to construct the frame $P : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times 2n}$ as the juxtaposition of two subframes $L, N : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times n}$, i.e.,

$$P = \begin{pmatrix} L \; N \end{pmatrix} \tag{2.15}$$

satisfying $\boldsymbol{\omega}(L_i, L_j) = \boldsymbol{\omega}(N_i, N_j) = 0$ and $\boldsymbol{\omega}(N_i, L_j) = \delta_{ij}$, where δ_{ij} is the Kronecker delta. Therefore, P constitutes what is known as a symplectic basis. In fact, any symplectic vector space admits such a basis, see e.g., [CdS01].

The subframe L can be chosen to be tangent to the torus whereas the normal subframe N complements L to a full symplectic basis of $T_{\mathcal{K}}U$, i.e., the tangent bundle of U restricted to $\mathcal{K}$.

2.4.1 Construction of the subframe L

A necessary condition for the frame P to satisfy (2.11) is that L needs to be invariant under the differential of time-T maps on $\mathcal{K}$. Consequently, L is required to satisfy

$$\mathrm{D}_z\phi_T\circ(K, \mathrm{id})L - L\circ\mathcal{R}_{\omega\alpha}\Lambda = 0 \tag{2.16}$$

Note that, according to (2.9) and (2.6), $\mathrm{D}_\theta K$ and W are invariant under the differential of time-T maps and they are therefore suitable to partly generate the subframe L. For autonomous Hamiltonian systems, the Hamiltonian vector field is invariant under the differential of time-T maps and, consequently, this suffices to construct the subframe L, see [HM21]. For periodic and quasi-periodic Hamiltonians, this is no longer the case due to the time dependency. We need to construct a new object invariant under $\mathrm{D}_z\phi_T$—this is the key element to apply the same ideas of flow map parameterization methods for autonomous systems to our setting.

Let us differentiate Eq. (2.3) with respect to φ to obtain

$$\mathrm{D}_z\phi_T\circ(K, \mathrm{id})\mathrm{D}_\varphi K + \mathrm{D}_\varphi\phi_T\circ(K, \mathrm{id}) - \mathrm{D}_\varphi K\circ\mathcal{R}_{\omega\alpha} = 0. \tag{2.17}$$

For the flow $\tilde{\phi}$, let us consider

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\tilde{\phi}_T \circ \tilde{\phi}_t(z, \varphi) \tag{2.18}$$

at $t = 0$ and observe that we have the identity

$$\begin{pmatrix} \mathrm{D}_z\phi_T(z,\varphi) & \mathrm{D}_\varphi\phi_T(z,\varphi) \\ O_{\ell\times 2n} & I_\ell \end{pmatrix} \begin{pmatrix} X(z,\varphi) \\ \hat{\alpha} \end{pmatrix} = \begin{pmatrix} X\Big(\phi_T(z,\varphi), \mathcal{R}_\alpha(\varphi)\Big) \\ \hat{\alpha} \end{pmatrix} \tag{2.19}$$

for any (z, φ). Then, using (2.17) and evaluating (2.19) on the torus, i.e., at $z = K(\theta, \varphi)$, it is easy to show that

$$\mathrm{D}_z\phi_T\circ(K, \mathrm{id})\mathcal{X} - \mathcal{X}\circ\mathcal{R}_{\omega\alpha} = 0, \tag{2.20}$$

where

$$\mathcal{X} := X\circ(K, \mathrm{id}) - \mathrm{D}_\varphi K\hat{\alpha} \tag{2.21}$$

is a geometric object defined on the torus that is invariant under $\mathrm{D}_z\phi_T$. Additionally, at each (θ, φ), $\mathcal{X}(\theta, \varphi)$ complements the column vectors of $\mathrm{D}_\theta K(\theta, \varphi)$, tangent to $\mathcal{K}_\varphi$ at $K(\theta, \varphi)$, to a full basis of the tangent bundle of the range of $\hat{K}(\cdot, \varphi)$ at $K(\hat{\theta}, \varphi)$. Such interpretation follows from differentiating (2.4) with respect to θ_d.

From (2.6),(2.9), and (2.20) it is immediate to see that the subframe L, given by

$$L := \begin{pmatrix} \mathrm{D}_\theta K & \mathcal{X} & W \end{pmatrix} \tag{2.22}$$

satisfies (2.16). Additionally, L has full rank since for each φ, $\mathrm{D}_\theta K$ and $\mathcal{X}$ generate the tangent bundle at $\mathcal{K}_\varphi$ of the $(n-1)$-dimensional torus defined by $\hat{K}(\cdot,\varphi)$ and W generates its stable or the unstable bundle.

Since we require for P to be a symplectic frame, i.e., to satisfy (2.13), this adds another layer of structure to the subframe L. More specifically, the subframe L needs to be a fiberwise Lagrangian subframe. That is, for each $\varphi \in \mathbb{T}^\ell$, $L(\cdot,\varphi)$ generates Lagrangian subspaces on the tangent space of U over the isotropic torus $\mathcal{K}_\varphi$. More specifically, for each $(\theta,\varphi) \in \mathbb{T}^{d+\ell}$, L generates a Lagrangian subspace on $T_{K(\theta,\varphi)}U$. Then, for $\Omega_L : \mathbb{T}^{d+\ell} \to \mathbb{R}^{n\times n}$ constructed as

$$\Omega_L := L^\top \Omega \circ K L = \begin{pmatrix} \Omega_{\mathrm{D}K} & \Omega_{\mathrm{D}K\mathcal{X}} & \Omega_{\mathrm{D}KW} \\ \Omega_{\mathcal{X}\mathrm{D}K} & \Omega_{\mathcal{X}} & \Omega_{\mathcal{X}W} \\ \Omega_{W\mathrm{D}K} & \Omega_{W\mathcal{X}} & \Omega_W \end{pmatrix}, \tag{2.23}$$

where

$$\begin{aligned}
\Omega_{\mathrm{D}K} &= (\mathrm{D}_\theta K)^\top \Omega \circ K \mathrm{D}_\theta K, & \Omega_{\mathrm{D}K\mathcal{X}} &:= (\mathrm{D}_\theta K)^\top \Omega \circ K \mathcal{X}, & \Omega_{\mathrm{D}KW} &:= (\mathrm{D}_\theta K)^\top \Omega \circ K W, \\
\Omega_{\mathcal{X}\mathrm{D}K} &:= \mathcal{X}^\top \Omega \circ K \mathrm{D}_\theta K, & \Omega_{\mathcal{X}} &:= \mathcal{X}^\top \Omega \circ K \mathcal{X}, & \Omega_{\mathcal{X}W} &:= \mathcal{X}^\top \Omega \circ K W, \\
\Omega_{W\mathrm{D}K} &:= W^\top \Omega \circ K \mathrm{D}_\theta K, & \Omega_{W\mathcal{X}} &:= W^\top \Omega \circ K \mathcal{X}, & \Omega_W &:= W^\top \Omega \circ K W,
\end{aligned}$$

L is a fiberwise Lagrangian subframe if and only if $\Omega_L = 0$. Note that because Ω is skew-symmetric, we have $\Omega_{\mathcal{X}\mathrm{D}K} = -\Omega_{\mathrm{D}K\mathcal{X}}{}^\top$, $\Omega_{W\mathrm{D}K} = -\Omega_{\mathrm{D}KW}{}^\top$, and $\Omega_{W\mathcal{X}} = -\Omega_{\mathcal{X}W}{}^\top$.

Proposition 2.2. *The subframe L constructed as in* (2.22) *is a fiberwise Lagrangian subframe.*

Proof. If L is fiberwise Lagrangian, it needs to satisfy

$$\Omega_L = 0.$$

Since Ω is skew-symmetric, it suffices to prove the following:

i) $\Omega_{\mathrm{D}K} = 0$, iii) $\Omega_{\mathrm{D}KW} = 0$, v) $\Omega_{\mathcal{X}W} = 0$,

ii) $\Omega_{\mathrm{D}K\mathcal{X}} = 0$, iv) $\Omega_{\mathcal{X}} = 0$, vi) $\Omega_W = 0$.

From Lemma 2.1, we have that $\mathcal{K}$ is fiberwise isotropic. Hence, i) is immediately satisfied. Additionally, iv) and vi) are the symplectic product of vectors with themselves; hence trivial. Note that vi) is also true for stable and unstable bundles of higher rank. The proof uses a similar contraction argument to the one we use here to prove iii) and v), which is the following. For any positive integer k, let $\mathcal{R}_{\omega\alpha}^{\circ k}$ denote the composition of $\mathcal{R}_{\omega\alpha}$ with itself k times, i.e.,

$$\mathcal{R}_{\omega\alpha}^{\circ k} := \underbrace{\mathcal{R}_{\omega\alpha} \circ \cdots \circ \mathcal{R}_{\omega\alpha}}_{k \text{ times}}.$$

For $k < 0$, we define $\mathcal{R}_{\omega\alpha}^{\circ k}$ as the composition of the inverse of $\mathcal{R}_{\omega\alpha}$ with itself k times. From the symplecticity of ϕ_T and the invariance of $\mathcal{K}$, we obtain

$$\Omega \circ K = \left(\mathrm{D}_z\phi_T \circ (K, \mathrm{id})\right)^\top \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \mathrm{D}_z\phi_T \circ (K, \mathrm{id}). \tag{2.24}$$

If we apply (2.24) recursively k times in iii) and v), and use the invariance of $\mathrm{D}_\theta K$, $\mathcal{X}$, and $\mathcal{W}$, we obtain the identities

$$\begin{aligned} (\mathrm{D}_\theta K)^\top \Omega \circ K W &= (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha}^{\circ k} \Omega \circ K \circ \mathcal{R}_{\omega\alpha}^{\circ k} W \circ \mathcal{R}_{\omega\alpha}^{\circ k} \lambda^k, \\ \mathcal{X}^\top \Omega \circ K W &= \mathcal{X}^\top \circ \mathcal{R}_{\omega\alpha}^{\circ k} \Omega \circ K \circ \mathcal{R}_{\omega\alpha}^{\circ k} W \circ \mathcal{R}_{\omega\alpha}^{\circ k} \lambda^k. \end{aligned}$$

Let us now assume the dynamics in $\mathcal{W}$ is contracting, i.e., $\lambda < 1$. Then by taking $k \to \infty$, we conclude iii) and v) hold. For the case where the dynamics in $\mathcal{W}$ is expanding, we can consider the invariance of $\mathcal{K}$ under ϕ_{-T} and proceed analogously.

For the proof of ii), we substitute (2.24) in $\Omega_{\mathrm{D}K\mathcal{X}}$ and use that $\mathrm{D}_\theta K$ and $\mathcal{X}$ are invariant under $\mathrm{D}_z\phi_T$. Then, we have

$$(\mathrm{D}_\theta K)^\top \Omega \circ K \mathcal{X} = (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \mathcal{X} \circ \mathcal{R}_{\omega\alpha},$$

which is constant by the ergodicity of the rotation $\mathcal{R}_{\omega\alpha}$. Recall that $\langle \cdot \rangle$ denotes the average with respect to θ and φ. Let us then take the average of $\Omega_{\mathrm{D}K\mathcal{X}}$ after expanding $\mathcal{X}$. That is,

$$\langle \Omega_{\mathrm{D}K\mathcal{X}} \rangle = \left\langle (\mathrm{D}_\theta K)^\top \Omega \circ K \Big(X \circ (K, \mathrm{id}) - \mathrm{D}_\varphi K \hat{\alpha} \Big) \right\rangle.$$

From the first term, using that $\Omega(z) X(z, \varphi) = \mathrm{D}_z H(z, \varphi)^\top$, we obtain

$$\left\langle (\mathrm{D}_\theta K)^\top \Big(\mathrm{D}_z H(K, \mathrm{id}) \Big)^\top \right\rangle = \left\langle \Big(\mathrm{D}_\theta \Big(H \circ (K, \mathrm{id}) \Big) \Big)^\top \right\rangle = 0$$

since we are taking averages of derivatives with respect to θ. For the second term, we use that $\Omega = (\mathrm{D}a)^\top - \mathrm{D}a$ and obtain

$$\left\langle (\mathrm{D}_\theta K)^\top \mathrm{D}_\varphi (a \circ K) - \Big(\mathrm{D}_\theta (a \circ K) \Big)^\top \mathrm{D}_\varphi K \right\rangle \hat{\alpha}. \tag{2.25}$$

Then, each term ij in (2.25) is obtained from

$$\begin{aligned} &\left\langle \sum_k \partial_{\theta^i} K^k \partial_{\varphi^j} (a_k \circ K) - \partial_{\theta^i} (a_k \circ K) \partial_{\varphi^j} K^k \right\rangle \\ &= \Big\langle \sum_k \partial_{\varphi^j} (a_k \circ K \partial_{\theta^i} K^k) - a_k \circ K \partial^2_{\theta^i \varphi^j} K^k \\ &\quad - \partial_{\theta^i} (a_k \circ K \partial_{\varphi^j} K^k) + a_k \circ K \partial^2_{\theta^i \varphi^j} K^k \Big\rangle \\ &= \left\langle \sum_k \partial_{\varphi^j} (a_k \circ K \partial_{\theta^i} K^k) - \partial_{\theta^i} (a_k \circ K \partial_{\varphi^j} K^k) \right\rangle = 0 \end{aligned}$$

since we are taking averages of derivatives with respect to θ and φ. Therefore ii) also holds. □

2.4.2 Construction of the subframe N

We now proceed to complement the subframe L with a subframe $N : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times n}$ such that the frame P is symplectic with respect to the standard symplectic form. Note that the symplecticity of P implies the subframe N also needs to be a Lagrangian subframe. There exists several ways to construct the subframe N, see [HCF$^+$16] for adapted frames to Lagrangian tori. We will use the almost complex structure and the Riemannian metric to construct N as

$$N = \hat{N} + LA, \tag{2.26}$$

for certain $A : \mathbb{T}^{d+\ell} \to \mathbb{R}^{n\times n}$, where

$$\begin{aligned} \hat{N} &= J\circ KLB, \\ B &= G_L^{-1}, \\ G_L &= L^\top G\circ KL. \end{aligned} \tag{2.27}$$

Let us emphasize some properties of these matrix-valued maps. A needs to be symmetric to preserve the symplecticity of P and G_L is invertible since K is a fiberwise embedding. That is, for each $\varphi \in \mathbb{T}^\ell$, $K(\cdot, \varphi) : \mathbb{T}^d \to U$ is an embedding. Lastly, both G_L and B are also symmetric. The role of A is to obtain simultaneously the stable and unstable bundles of the invariant torus but its use is not necessary. We can take $A = 0$ and the frame $\hat{P} := (L\ \hat{N})$ is symplectic and reduces the linearized dynamics on the torus to upper triangular form—in fact, this holds for any symmetric A. See Section 3.2.9 and 3.2.10. To obtain both the stable and unstable bundles let us denote the torsion of the normal bundles parameterized by $\hat{N}$ and N, when transported by $\mathrm{D}_z\phi_T$ on $\mathcal{K}$, by the matrix-valued maps $\hat{S}, S : \mathbb{T}^{d+\ell} \to \mathbb{R}^{n\times n}$, i.e.,

$$\begin{aligned} \hat{S} &:= \hat{N}^\top\circ\mathcal{R}_{\omega\alpha}\,\Omega\circ K\circ\mathcal{R}_{\omega\alpha}\,\mathrm{D}_z\phi_T\circ(K,\mathrm{id})\hat{N}, \\ S &:= N^\top\circ\mathcal{R}_{\omega\alpha}\,\Omega\circ K\circ\mathcal{R}_{\omega\alpha}\,\mathrm{D}_z\phi_T\circ(K,\mathrm{id})N, \end{aligned} \tag{2.28}$$

where a straightforward computation reveals

$$S = \hat{S} + \Lambda A - A^\top\circ\mathcal{R}_{\omega\alpha}\Lambda^{-\top}. \tag{2.29}$$

Recall definition (2.12). From (2.11) and (2.12), observe that if we choose A such that S adopts the form

$$S = \left(\begin{array}{c|c} \mathcal{S} & \\ \hline & \end{array}\right), \tag{2.30}$$

where $S : \mathbb{T}^{d+\ell} \to \mathbb{R}^{(n-1)\times(n-1)}$, then the last column of P, that we denote by $\widetilde{W}$, satisfies

$$\mathrm{D}_z\phi_T\circ(K,\mathrm{id})\widetilde{W} - \widetilde{W}\circ\mathcal{R}_{\omega\alpha}\lambda^{-1} = 0. \tag{2.31}$$

Therefore, W and $\widetilde{W}$ parameterize both the stable and unstable bundles of $\mathcal{K}$, respectively.

Let us define splittings for $\hat{S}$ and A into blocks of sizes $(n-1)\times(n-1), (n-1)\times 1, 1\times (n-1), 1\times 1$ as

$$\hat{S} = \begin{pmatrix} \hat{S}_1 & \hat{S}_2 \\ \hat{S}_3 & \hat{S}_4 \end{pmatrix}, \quad A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}. \tag{2.32}$$

Then, using (2.29), condition (2.30) reads

$$\mathcal{S} = \hat{S}_1 + \quad A_1 - A_1^\top\circ\mathcal{R}_{\omega\alpha}, \tag{2.33}$$

$$0 = \hat{S}_2 + \quad A_2 - A_3^\top \circ \mathcal{R}_{\omega\alpha} \lambda^{-1}, \tag{2.34}$$
$$0 = \hat{S}_3 + \lambda A_3 - A_2^\top \circ \mathcal{R}_{\omega\alpha}, \tag{2.35}$$
$$0 = \hat{S}_4 + \lambda A_4 - A_4 \circ \mathcal{R}_{\omega\alpha} \lambda^{-1}, \tag{2.36}$$

where we can take $A_1 = \mathcal{O}_{n-1}$—consequently $\mathcal{S} = \hat{S}_1$—and solve (2.36) as described in Section 1.5.1. Notice that (2.34) and (2.35) are coupled non-small cohomological equations that can be rewritten as

$$0 = \left(\hat{S}_2 \lambda + \hat{S}_3^\top \circ \mathcal{R}_{\omega\alpha} \lambda^{-1}\right) + A_2 \lambda - A_2 \circ R_{\omega\alpha}^{\circ 2} \lambda^{-1}, \tag{2.37}$$
$$0 = \left(\hat{S}_3 + \hat{S}_2^\top \circ \mathcal{R}_{\omega\alpha}\right) + \lambda A_3 - \lambda^{-1} A_3 \circ R_{\omega\alpha}^{\circ 2}, \tag{2.38}$$

where $R_{\omega\alpha}^{\circ 2}$ stands for $\mathcal{R}_{\omega\alpha} \circ \mathcal{R}_{\omega\alpha}$. Then, we can solve (2.37) and (2.38) as non-small divisors cohomological equations. We will also refer to $\mathcal{S}$ as the reduced torsion or the *shear*. Observe that we need A to be symmetric in order for P to be symplectic, but so far we do not know if for a symmetric A (2.29) is solvable for S of the form (2.30).

Proposition 2.3. *The torsion $\hat{S}$ satisfies the symmetry condition*

$$\Lambda^{-1} \hat{S} - \left(\Lambda^{-1} \hat{S}\right)^\top = 0. \tag{2.39}$$

For the proof, see Section 3.2.8. From (2.39), the torsion $\hat{S}$ satisfies $\hat{S}_2 = \hat{S}_3^\top \lambda^{-1}$. Then, since A_2 and A_3 are solutions of (2.37) and (2.38), it follows that $A_2 = A_3^\top$. Consequently, for S as in (2.30) we can find a suitable symmetric A such that (2.29) holds. Using this symmetry, instead of solving (2.37) and (2.38), we can solve

$$0 = \hat{S}_2 + A_2 - A_2 \circ \mathcal{R}_{\omega\alpha} \lambda^{-1} \tag{2.40}$$

as a non-small divisors cohomological equation and take $A_3 = A_2^\top$.

Remark 2.8. From (2.33), it is possible to further reduce $\mathcal{S}$ to a constant matrix $\mathcal{S} = \left\langle \hat{S}_1 \right\rangle$ by solving

$$\mathfrak{L} A_1 = \left\langle \hat{S}_1 \right\rangle - \hat{S}_1.$$

We summarize the construction of adapted frames with the following algorithm.

Algorithm 2.1 Computation of adapted frames

Given (K, W, λ) satisfying (2.3) and (2.6), compute the frame P and the reduced torsion $\mathcal{S}$ by following these steps:

1) Compute $\mathcal{X}$ from (2.21).
2) Compute L from (2.22).
3) Compute $\hat{N}$ from (2.27).
4) Compute $\hat{S}$ from (2.28), set $A_1 = \mathcal{O}_{n-1}$, and $\mathcal{S} = \hat{S}_1$.
5) Compute A_2 by solving (2.40) and take $A_3 = A_2^\top$.
6) Compute A_4 by solving (2.36).
7) Compute N from (2.26).
8) Compute P from (2.15).

2.5 Computation of invariant tori and bundles

Given (K, W, λ) that approximately satisfy equations (2.3) and (2.6), our aim is to improve such approximations and find exact solutions of the invariance equations. Let us define the error in the invariance of $\mathcal{K}$ and $\mathcal{W}$ as $E_K, E_W : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$ given by

$$E_K = \phi_T \circ (K, \mathrm{id}) - K \circ \mathcal{R}_{\omega\alpha}, \tag{2.41}$$

$$E_W = \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) W - W \circ \mathcal{R}_{\omega\alpha} \lambda. \tag{2.42}$$

Note that the geometric properties of Section 2.3 only hold when we have exact solutions of (2.3) and (2.6). Nonetheless, as we will see in Section 3.2.10, since $\mathcal{K}$ and $\mathcal{W}$ are approximately invariant, $\mathcal{K}$ is approximately reducible. That is, there is an error in the reducibility of the linearized dynamics controlled by E_K and E_W as

$$(P \circ \mathcal{R}_{\omega\alpha})^{-1} \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) P = \left(\begin{array}{c|c} \Lambda & S \\ \hline & \Lambda^{-\top} \end{array} \right) + \mathcal{O}(\|E_K\|) + \mathcal{O}(\|E_W\|),$$

for some suitable norm. Also, the frame P is approximately symplectic. Therefore, we have

$$P^{-1} = -\Omega_0 P^\top \Omega \circ K + \mathcal{O}(\|E_K\|) + \mathcal{O}(\|E_W\|),$$

from where we can obtain an approximate inverse.

The strategy is to design an iterative procedure to constructs sequences for K, W, and λ such that for the limiting objects K_∞, W_∞, and λ_∞ the errors satisfy $\|E_{K_\infty}\| = 0$ and $\|E_{W_\infty}\| = 0$. The sequences are obtained recursively by choosing corrections $\Delta K, \Delta W : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$, and $\Delta\lambda \in \mathbb{R}$ such that for $\bar{K} = K + \Delta K$, $\bar{W} = W + \Delta W$, and $\bar{\lambda} = \lambda + \Delta\lambda$, (2.3) and (2.6) vanish at first order. In this section, we describe non-rigorously one step of the iterative scheme—the rigorous description is postponed to Section 3.4.1 and the convergence of the scheme to Section 3.4.2.

2.5.1 One step on the torus

Let us begin with (2.3). We compute the correction of the torus parameterization $\Delta K : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$ in the coordinates given by the frame P. That is, we choose $\Delta K = P\xi_K$, where $\xi_K : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$ satisfies

$$\mathrm{D}_z \phi_T \circ (K, \mathrm{id}) P \xi_K - P \circ \mathcal{R}_{\omega\alpha} \xi_K \circ \mathcal{R}_{\omega\alpha} = -E_K \tag{2.43}$$

up to first order in E_K and E_W. Let us left-multiply the linearized equation (2.43) by the approximate inverse of $P \circ \mathcal{R}_{\omega\alpha}$—that is, $-\Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha}$—to obtain, after neglecting higher order terms,

$$\left(\begin{array}{c|c} \Lambda & S \\ \hline & \Lambda^{-\top} \end{array} \right) \xi_K - \xi_K \circ \mathcal{R}_{\omega\alpha} = \eta_K, \tag{2.44}$$

where

$$\eta_K = \Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} E_K. \tag{2.45}$$

We can now write ξ_K and η_K into $(n-1) \times 1 \times (n-1) \times 1$ components as

$$\xi_K = \begin{pmatrix} \xi_K^1 \\ \xi_K^2 \\ \xi_K^3 \\ \xi_K^4 \end{pmatrix}, \quad \eta_K = \begin{pmatrix} \eta_K^1 \\ \eta_K^2 \\ \eta_K^3 \\ \eta_K^4 \end{pmatrix},$$

and rewrite system (2.44) as follows

$$\xi_K^1 + \mathcal{S}\xi_K^3 - \xi_K^1 \circ \mathcal{R}_{\omega\alpha} = \eta_K^1, \tag{2.46}$$

$$\lambda\xi_K^2 - \xi_K^2 \circ \mathcal{R}_{\omega\alpha} = \eta_K^2, \tag{2.47}$$

$$\xi_K^3 - \xi_K^3 \circ \mathcal{R}_{\omega\alpha} = \eta_K^3 - \langle \eta_K^3 \rangle, \tag{2.48}$$

$$\lambda^{-1}\xi_K^4 - \xi_K^4 \circ \mathcal{R}_{\omega\alpha} = \eta_K^4. \tag{2.49}$$

Observe that we have subtracted $\langle \eta_K^3 \rangle$ in (2.48) in order to overcome the obstruction in the small divisors cohomological equation—in Appendix B, we prove that $\langle \eta_K^3 \rangle$ is quadratically small in E_K so its contribution in (2.48) is of second order. Then, we can obtain ξ_K^i for $i = 2, 3, 4$ as

$$\xi_K^2 = \mathfrak{R}^{\lambda 1}\eta_K^2, \tag{2.50}$$

$$\xi_K^3 = \mathfrak{R}\eta_K^3 + \left\langle \xi_K^3 \right\rangle, \tag{2.51}$$

$$\xi_K^4 = \mathfrak{R}^{\lambda^{-1}1}\eta_K^4, \tag{2.52}$$

where $\langle \xi_K^3 \rangle$ is free. Recall the definition of the Rüssmann operators from Section 1.5. We can then require

$$\left\langle \xi_K^3 \right\rangle = \langle \mathcal{S} \rangle^{-1} \left\langle \eta_K^1 - \mathcal{S}\mathfrak{R}\eta_K^3 \right\rangle \tag{2.53}$$

given the non-degeneracy condition $\det \langle \mathcal{S} \rangle \neq 0$. Then

$$\xi_K^1 = \mathfrak{R}\left(\eta_K^1 - \mathcal{S}\left(\mathfrak{R}\eta_K^3 + \left\langle \xi_K^3 \right\rangle\right)\right) + \left\langle \xi_K^1 \right\rangle, \tag{2.54}$$

where $\langle \xi_K^1 \rangle$ is free and we can simply take as zero. This freedom has to do with the freedom to choose the phase of the generating torus $\mathcal{K}$, see Remark 2.2.

The iterative step on the torus is summarized with the following algorithm.

Algorithm 2.2 Iterative step on the torus

Let (K, W, λ) satisfy equations (2.3) and (2.6) approximately. Obtain the corrected generating torus by following these steps:

1) Compute P and $\mathcal{S}$ by following Algorithm 2.1.
2) Compute the error E_K from (2.41).
3) Compute η_K, the right-hand side of the cohomological equations given in (2.44), from (2.45).
4) Solve (2.47) and (2.49) as non-small divisors cohomological equations and obtain ξ_K^2 and ξ_K^4 from (2.50) and (2.52).
5) Solve (2.48) as small divisors cohomological equations in order to obtain its zero-average solution $\mathfrak{R}\eta_K^3$.
6) Compute $\langle \xi_K^3 \rangle$ from (2.53) and ξ_K^3 from (2.51).
7) Set $\langle \xi_K^1 \rangle = 0$ and solve (2.46) as small divisors cohomological equations to obtain ξ_K^1 from (2.54).
8) Set $\bar{K} \leftarrow K + P\xi_K$.

2.5.2 One step on the bundle

For the parameterization of the invariant bundle, we proceed in an analogous manner and add corrections $\Delta W : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$ and $\Delta\lambda \in \mathbb{R}$ to the parameterization of the generating bundle and to the Floquet multiplier, respectively. We compute as well the correction in the coordinates given by the frame P. That is, we take $\Delta W = P\xi_W$ where $\xi_W : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$ satisfies

$$\mathrm{D}_z\phi_T \circ (K, \mathrm{id}) P\xi_W - P \circ \mathcal{R}_{\omega\alpha} \xi_W \circ \mathcal{R}_{\omega\alpha} \lambda - W \circ \mathcal{R}_{\omega\alpha} \Delta\lambda = -\widetilde{E_W}, \tag{2.55}$$

and

$$\widetilde{E_W} := E_W + \int_0^1 \mathrm{D}_z^2 \phi_T \circ \Big(K + s\Delta K, \mathrm{id}\Big)[\Delta K, W] ds. \tag{2.56}$$

We emphasize that albeit we compute ΔK in the previous section, we linearize around K and not $\bar{K} = K + \Delta K$. This allows us to use the same frame P to correct the torus and the bundle.

Let us make some comments on possible variations. When linearizing (2.6), it is possible to consider

$$\widetilde{E_W} = E_W + \mathrm{D}_z^2 \phi_T \circ (K, \mathrm{id})[\Delta K, W] \tag{2.57}$$

which can be easier to compute than (2.56) in practical applications. However, for the proof on convergence, (2.57) requires upper bounds on the norm of $\mathrm{D}_z^3\phi_T$, whereas (2.56) requires bounds only of $\mathrm{D}_z^2\phi_T$. It is also possible to compute first $\bar{K}$ and consider (2.6) around $\bar{K}$. Then, we can obtain ξ_W and $\Delta\lambda$ from

$$\mathrm{D}_z\phi_T \circ (\bar{K}, \mathrm{id}) \bar{P}\xi_W - \bar{P} \circ \mathcal{R}_{\omega\alpha} \xi_W \circ \mathcal{R}_{\omega\alpha} \lambda - W \circ \mathcal{R}_{\omega\alpha} \Delta\lambda = -E_{\bar{K}W}, \tag{2.58}$$

where $\bar{P}$ is the frame for the corrected parameterization $\bar{K}$ and $E_{\bar{K}W}$ corresponds to (2.42) but on $\bar{K}$ instead of K.

Notice that then we do not have to compute $\mathrm{D}_z^2\phi_T$ but we have to compute two frames: for K and for $\bar{K}$. Computationally, this is an interesting option and it is in fact the approach followed in the numerical explorations of Section 2.8 and in [FMHM24b]. However, for the KAM theorem, it would require to control both frames P and $\bar{P}$.

Regardless of which option we choose, let us left-multiply the linearized equation (2.55) again by the approximate inverse of $P \circ \mathcal{R}_{\omega\alpha}$, i.e., by $-\Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha}$. If instead we would like to follow the approach (2.58), we would proceed analogously with an approximate inverse of $\bar{P} \circ \mathcal{R}_{\omega\alpha}$. Then, after neglecting higher order terms, we obtain

$$\left(\begin{array}{c|c} \Lambda & S \\ \hline & \Lambda^{-\top} \end{array}\right) \xi_W - \xi_W \circ \mathcal{R}_{\omega\alpha} \lambda = \eta_W + e_n \Delta\lambda, \tag{2.59}$$

where

$$\eta_W := \Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \widetilde{E_W}, \tag{2.60}$$

$$e_n := \begin{pmatrix} 0_{n-1} \\ 1 \\ 0_{n-1} \\ 0 \end{pmatrix}.$$

Using analogous splittings for η_W and ξ_W as in Section 2.5.1, system (2.59) reads

$$\xi_W^1 + \mathcal{S}\xi_W^3 - \xi_W^1 \circ \mathcal{R}_{\omega\alpha} \lambda = \eta_W^1, \tag{2.61}$$

$$\lambda\xi_W^2 - \xi_W^2 \circ \mathcal{R}_{\omega\alpha}\lambda = \eta_W^2 + \Delta\lambda, \tag{2.62}$$
$$\xi_W^3 - \xi_W^3 \circ \mathcal{R}_{\omega\alpha}\lambda = \eta_W^3, \tag{2.63}$$
$$\lambda^{-1}\xi_W^4 - \xi_W^4 \circ \mathcal{R}_{\omega\alpha}\lambda = \eta_W^4. \tag{2.64}$$

We can then take $\Delta\lambda = -\langle \eta_W^2 \rangle$ and

$$\xi_W^1 = \mathfrak{R}^{1\lambda}\Big(\eta_W^1 - \mathcal{S}\mathfrak{R}^{1\lambda}\eta_W^3\Big), \tag{2.65}$$
$$\xi_W^2 = \frac{1}{\lambda}\mathfrak{R}\eta_W^2 + \Big\langle \xi_W^2 \Big\rangle, \tag{2.66}$$
$$\xi_W^3 = \mathfrak{R}^{1\lambda}\eta_W^3, \tag{2.67}$$
$$\xi_W^4 = \mathfrak{R}^{\lambda^{-1}\lambda}\eta_W^4, \tag{2.68}$$

where $\langle \xi_W^2 \rangle$ is free and we can take as zero. This freedom is related to the freedom in the length of W, analogous to the underdeterminacy of the lengths of eigenvectors in an eigenvalue problem.

The iterative step on the bundle is summarized with the following algorithm.

Algorithm 2.3 Iterative step on the bundle

Let (K, W, λ) satisfy equations (2.3) and (2.6) approximately. Obtain the corrected generating bundle and Floquet multiplier by following these steps:

1) Choose whether to follow approach (2.55) or (2.58).
2) If approach (2.55) is chosen, take P and $\mathcal{S}$ as computed in Algorithm 2.2. Else, compute $\bar{P}$ and $\bar{\mathcal{S}}$ by following algorithm 2.1 for $\bar{K}$.
3) Compute the error E_W from (2.42).
4) If approach (2.55) is chosen, compute $\widetilde{E_W}$ from (2.56) or (2.57).
5) Compute η_W, using the chosen frame P or $\bar{P}$, from (2.60). For the case (2.58), take $\widetilde{E_W} = E_W$.
6) Solve (2.63), (2.64), and (2.61) as non-small divisors cohomological equations in order to obtain ξ_W^3, ξ_W^4, and ξ_W^1 from (2.67), (2.68), and (2.65).
7) Take $\Delta\lambda = -\langle \eta_W^2 \rangle$.
8) Set $\langle \xi^2 \rangle = 0$ and solve (2.62) as a small divisors cohomological equations to obtain ξ_W^2 from (2.66).
9) Set $\bar{W} \leftarrow W + P\xi_W$ or $\bar{W} \leftarrow W + \bar{P}\xi_W$ and $\bar{\lambda} \leftarrow \lambda + \Delta\lambda$.

2.6 Continuation with respect to parameters of H

Let us assume we have a family of Hamiltonians H_ε that depends analytically on some parameter $\varepsilon \in \mathbb{R}$. Given $(K, W, \lambda)_\varepsilon$ for certain value of ε that satisfies equations (2.3) and (2.6), we want to compute parameterizations of a generating torus and generating bundle (with the corresponding Floquet multiplier) for a different Hamiltonian $H_{\varepsilon'}$. As commonly done in continuation methods, see e.g. [AG03], we will provide a methodology to compute the tangent to the continuation curve with respect to ε from where we obtain a first order approximation in ε of the invariant objects for $H_{\varepsilon'}$.

Let us assume equations (2.3) and (2.6) define implicitly $(K, W, \lambda)_\varepsilon$ as functions of ε. Then, we want to compute $\partial_\varepsilon K, \partial_\varepsilon W$, and $\partial_\varepsilon \lambda$. In the following, for the sake of notation clarity, we will omit the dependency of (K, W, λ) and of ϕ_T on ε. Let us begin by taking Eq. (2.3) and differentiate it with respect to ε to obtain

$$\mathrm{D}_z\phi_T \circ (K, \mathrm{id})\partial_\varepsilon K + \partial_\varepsilon \phi_T \circ (K, \mathrm{id}) - \partial_\varepsilon K \circ \mathcal{R}_{\omega\alpha} = 0,$$

where $\partial_\varepsilon \phi_T$ is the variation of the map ϕ_T with respect to ε that can be computed through variational equations. We now express the derivatives in the frame such that $\partial_\varepsilon K = P\xi_{\partial_\varepsilon K}$. The previous expression then reads

$$\mathrm{D}_z\phi_T \circ (K, \mathrm{id})P\xi_{\partial_\varepsilon K} + \partial_\varepsilon \phi_T \circ (K, \mathrm{id}) - P \circ \mathcal{R}_{\omega\alpha}\xi_{\partial_\varepsilon K} \circ \mathcal{R}_{\omega\alpha} = 0. \tag{2.69}$$

Notice that now we have exact solutions of (2.3) and (2.6) so reducibility holds exactly. Then, after left-multiplication of (2.69) by $(P \circ \mathcal{R}_{\omega\alpha})^{-1}$ we have

$$\left(\begin{array}{c|c} \Lambda & S \\ \hline & \Lambda^{-\top} \end{array}\right) \xi_{\partial_\varepsilon K} - \xi_{\partial_\varepsilon K} \circ \mathcal{R}_{\omega\alpha} = \eta_{\partial_\varepsilon K}. \tag{2.70}$$

where

$$\eta_{\partial_\varepsilon K} = \Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \partial_\varepsilon \phi_T \circ (K, \mathrm{id}). \tag{2.71}$$

The cohomological equations given in (2.70) can be solved as described in Section 2.5.1. Observe that the systems of cohomological equations (2.70) and (2.44) are the same. Then, for (2.70) to be solvable, we encounter again small divisors cohomological equations for $\xi^3_{\partial_\varepsilon K}$ (recall the splittings defined in Section 2.5.1 that require $\langle \eta^3_{\partial_\varepsilon K}\rangle = 0$. Using fiberwise symplectic deformations, see Appendix C, we prove in Appendix D that, for $\eta_{\partial_\varepsilon K}$ defined as in (2.71), $\langle \eta^3_{\partial_\varepsilon K}\rangle = 0$.

In order to obtain $\partial_\varepsilon W$ and $\partial_\varepsilon \lambda$, we differentiate (2.6) with respect to ε to obtain

$$\partial_\varepsilon\Big(\mathrm{D}_z\phi_T \circ (K, \mathrm{id})\Big)W + \mathrm{D}_z\phi_T \circ (K, \mathrm{id})\partial_\varepsilon W - W \circ \mathcal{R}_{\omega\alpha} \partial_\varepsilon \lambda - \partial_\varepsilon W \circ \mathcal{R}_{\omega\alpha} \lambda = 0,$$

which can be rewritten as

$$\mathrm{D}_z\phi_T \circ (K, \mathrm{id})\partial_\varepsilon W - \partial_\varepsilon W \circ \mathcal{R}_{\omega\alpha} \lambda - W \circ \mathcal{R}_{\omega\alpha} \partial_\varepsilon \lambda = -E_{\partial_\varepsilon W}, \tag{2.72}$$

with

$$\begin{aligned} E_{\partial_\varepsilon W} &= \partial_\varepsilon\Big(\mathrm{D}_z\phi_T \circ (K, \mathrm{id})\Big)W \\ &= \partial_\varepsilon \mathrm{D}_z\phi_T \circ (K, \mathrm{id})W + \mathrm{D}^2_z\phi_T \circ (K, \mathrm{id})[W, \partial_\varepsilon K]. \end{aligned} \tag{2.73}$$

Then, by using the frame such that $\partial_\varepsilon W = P\xi_{\partial_\varepsilon W}$, and after left-multiplication by $(P \circ \mathcal{R}_{\omega\alpha})^{-1}$, expression (2.72) reads

$$\left(\begin{array}{c|c} \lambda & S \\ \hline & \lambda^{-\top} \end{array}\right) \xi_{\partial_\varepsilon W} - \xi_{\partial_\varepsilon W} \circ \mathcal{R}_{\omega\alpha} \lambda = \eta_{\partial_\varepsilon W} + e_n \partial_\varepsilon \lambda, \tag{2.74}$$

with

$$\eta_{\partial_\varepsilon W} = \Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K E_{\partial_\varepsilon W}. \tag{2.75}$$

These cohomological equations can be solved as described in Section 2.5.2.

The computation of derivatives of (K, W, λ) with respect to parameters is summarized with the following algorithm.

Algorithm 2.4 Computation of derivatives with respect to parameters

Let $(K, W, \lambda)_\varepsilon$ be implicit functions of ε by equations (2.3) and (2.6). Find $\partial_\varepsilon K$, $\partial_\varepsilon W$, and $\partial_\varepsilon \lambda$ by following these steps:

1) Compute P and $\mathcal{S}$ by following Algorithm 2.1.
2) Compute $\eta_{\partial_\varepsilon K}$, the right-hand side of the cohomological equations given in (2.70), from (2.71).
3) Solve (2.47) and (2.49) with $\eta_{\partial_\varepsilon K}$ as non-small divisors cohomological equations in order to obtain $\xi^2_{\partial_\varepsilon K}$ and $\xi^4_{\partial_\varepsilon K}$.
4) Solve (2.48) with $\eta_{\partial_\varepsilon K}$ as small divisors cohomological equations in order to obtain its zero-average solution $\mathfrak{R}\eta^3_{\partial_\varepsilon K}$.
5) Compute $\left\langle \xi^3_{\partial_\varepsilon K} \right\rangle$ from (2.53) with $\eta_{\partial_\varepsilon K}$ and $\xi^3_{\partial_\varepsilon K} = \mathfrak{R}\eta^3_{\partial_\varepsilon K} + \left\langle \xi^3_{\partial_\varepsilon K} \right\rangle$.
6) Set $\langle \xi^1_{\partial_\varepsilon K} \rangle = 0$ and solve (2.46) with $\eta_{\partial_\varepsilon K}$ as small divisors cohomological equations to obtain $\xi^1_{\partial_\varepsilon K}$.
7) Set $\partial_\varepsilon K \leftarrow P\xi_{\partial_\varepsilon K}$.
8) Compute $E_{\partial_\varepsilon W}$ from (2.73).
9) Compute $\eta_{\partial_\varepsilon W}$ from (2.75).
10) Solve (2.63), (2.64), and (2.61) with $\eta_{\partial_\varepsilon W}$ as non-small divisors cohomological equations in order to obtain $\xi^3_{\partial_\varepsilon W}$, $\xi^4_{\partial_\varepsilon W}$, and $\xi^1_{\partial_\varepsilon W}$, respectively.
11) Take $\partial_\varepsilon \lambda = -\langle \eta^2_{\partial_\varepsilon W} \rangle$.
12) Set $\langle \xi^2_{\partial_\varepsilon W} \rangle = 0$ and solve (2.62) with $\eta_{\partial_\varepsilon W}$ and $\partial_\varepsilon \lambda$ as a small divisors cohomological equations to obtain $\xi^2_{\partial_\varepsilon W} = \frac{1}{\lambda}\mathfrak{R}\eta^2_{\partial_\varepsilon W}$.
13) Set $\partial_\varepsilon W \leftarrow P\xi_{\partial_\varepsilon W}$.

2.7 Comments on implementations

For every function $\zeta : \mathbb{T}^{d+\ell} \to \mathbb{R}$, we use its Fourier representation and their grid representation. The Fourier representation is given in terms of Fourier coefficients $\{\hat{\zeta}_{kj}\} \in \mathbb{C}$, with $k \in \mathbb{Z}^d$ and $j \in \mathbb{Z}^\ell$, and the grid representation is given by the values $\{\zeta_{kj}\}$ of ζ in an equally spaced grid of $\mathbb{T}^{d+\ell}$. We can switch between both representations with the linear one-to-one map provided by the Discrete Fourier Transform (DFT). Operations such as phase shifts, differentiation, or solving cohomological equations can be done efficiently in Fourier space whereas numerical integration or evaluation of vector fields can be done more efficiently in the grid representation. We switch between both representations according to the operations to be performed.

In practice, we can only work with truncated series. Furthermore, the DFT only provides approximate coefficients, that cannot in general be directly identified with the Fourier coefficients. Instead, for each coefficient $\hat{\zeta}_{kj}$, we use the DFT coefficient that best approximates it, see e.g. [Hen79] for details.

The unstable Floquet multiplier of $\mathcal{K}$ can be large, which would compromise the convergence of the method. In the numerical explorations of Section 2.8, instead of solving Eqs. (2.3) and (2.6), we follow a multiple shooting approach. We consider multiple tori $\{\mathcal{K}_i\}_{i=0}^{m-1}$ and bundles $\{\mathcal{W}_i\}_{i=0}^{m-1}$ parameterized by $\{K_i\}_{i=0}^{m-1}$ and $\{W_i\}_{i=0}^{m-1}$ with $K_i : \mathbb{T}^{d+\ell} \to U$ and $W_i : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$ such that

$$\phi_{T_m} \circ (K_i, \mathrm{id}) - K_{i+1} \circ \mathcal{R}_{\omega_m \alpha_m} = 0, \tag{2.76}$$

$$\mathrm{D}_z\phi_{T_m}\circ(K_i,\mathrm{id})W_i - \lambda_m W_{i+1}\circ\mathcal{R}_{\omega_m\alpha_m} = 0 \tag{2.77}$$

for $i = 1, ..., m-1$, where $T_m = T/m$, $\lambda_m = \lambda^{\frac{1}{m}}$, $\omega_m := \omega/m$, and $\alpha_m := \alpha/m$. The subindex in K and W is defined modulo m. We choose m such that both $|\lambda|^{\frac{1}{m}}$ and $|\lambda|^{\frac{-1}{m}}$ are small enough. The methodology described generalizes with some extra work for Eqs. (2.76) and (2.77). For the sake of clarity, we have described the method for $m = 1$. Nonetheless, the generalization from $m = 1$ to $m > 1$ is analogous to the generalization in the autonomous case which can be found in [HM21].

The evaluation of flow maps and the variational equations require numerical integration that can be costly for high-dimensional tori. Numerical integration is easily parallelizable by assigning trajectories to different cores.

To prevent numerical instabilities, we implement a low-pass filter after each iterative step on K and W for the approximate Fourier coefficients of the parameterizations, i.e., at the end of Algorithms 2.2 and 2.3. For simplicity, assume $\mathcal{K}$ is a 2-dimensional torus with a grid of size $N_1 \times N_2$—which is the test case of Section 2.8. The filtering strategy consists of setting to zero the coefficients for $|k| > k_f$ and $|j| > j_f$ for the values $k_f = r_f \cdot N_1$ and $j_f = r_f \cdot N_2$, where $r_f \in \left[\frac{1}{4}, \frac{1}{2}\right)$ is some filtering factor. Because of this filtering strategy, the effective number of approximate Fourier coefficients is not $N_1 \times N_2$.

Since we are working with truncated Fourier series, we need to decide on the number of Fourier coefficients. This number needs to be large enough so the series allow the parameterizations to meet their required error tolerance in the invariance equations. The necessary number of coefficients might change throughout the continuations so we also need a strategy to adjust the grid sizes of the invariant objects. Since our objects are real analytic, their Fourier coefficients decay exponentially fast. Our strategy is based on controlling this decay. In order to do so, we compute the tails of the truncated series. For functions $\zeta : \mathbb{T}^2 \to \mathbb{R}$, we compute the tails in the internal phase θ and the external phase φ as

$$t_\theta(\zeta) = \sum_{|k|>k_t, |j|<j_t} |\hat{\zeta}_{kj}|, \qquad t_\varphi(\zeta) = \sum_{|k|<k_t, |j|>j_t} |\hat{\zeta}_{kj}|,$$

for $k_t = r_t \cdot N_1$, $j_t = r_t \cdot N_2$, where r_t is some tail factor such that $r_t < r_f$. For the case where $\zeta : \mathbb{T}^2 \to \mathbb{R}^{2n}$, we consider t_θ and t_φ to act component-wise.

Furthermore, throughout the continuations on φ, the necessary step size of the continuation procedure needs to be adjusted. We conclude this section with a proposal in Algorithm 2.5 for the continuation of generating tori, bundles, and Floquet multipliers that is an adaptation to our case of Algorithm 3.6.1 in [HM21]. The main difference is that in the algorithm of [HM21], a strategy to reduce the number of Fourier coefficients in the parameterizations is necessary whereas in Algorithm 2.5, we require a strategy to control the decay of the coefficients in the different phases of K and W to adjust the grid sizes of the parameterizations.

Let us represent the parameterizations of the invariant objects and the Floquet multiplier for some parameter ε as $\mathcal{T}_\varepsilon = (\varepsilon, K, W, \lambda)$, for which we consider the norm

$$\|\mathcal{T}_\varepsilon\| := \left(\varepsilon^2 + \lambda^2 + \langle\|K\|_2^2\rangle + \langle\|W\|_2^2\rangle\right)^{\frac{1}{2}},$$

where $\|K\|_2$ stands for the map $(\theta, \varphi) \mapsto \|K(\theta, \varphi)\|_2$ and $\|\cdot\|_2$ denotes the Euclidean norm in $\mathbb{R}^{2n}$. The map $\|W\|_2$ is analogous. Let us estimate the error in the torus and bundle as follows

$$\mathrm{err}_K(\mathcal{T}_\varepsilon) = \max_{\substack{0\le k<N_1\\ 0\le j<N_2}} \|E_K(k/N_1, j/N_2)\|_\infty, \tag{2.78}$$

$$\mathrm{err}_W(\mathcal{T}_\varepsilon) = \max_{\substack{0\le k<N_1\\ 0\le j<N_2}} \|E_W(k/N_1, j/N_2)\|_\infty, \tag{2.79}$$

where $\|\cdot\|_\infty$ stands for the supremum norm in $\mathbb{R}^{2n}$. Then, a proposal for the continuation of tori and their bundles is summarized in the following algorithm.

Algorithm 2.5 Continuation implementation

Let $\mathcal{T}_\varepsilon$ be an approximately invariant torus, bundle, and Floquet multiplier for some value of the parameter ε and some values of the flying time T and rotation vectors ω and α. Let K and W have grid representations of size $N_1 \times N_2$, let $\varepsilon_K, \varepsilon_W, \varepsilon_t$ be tolerances, r_t some tail factor, and $n_{max}, n_\varepsilon, n_{des}, n_t$ be integers. Assume we have a suggested continuation step $\Delta\varepsilon$. Compute a new torus, bundle, and Floquet multiplier $\mathcal{T}_{\varepsilon'}$ for a new value of the parameter ε' and fixed T, ω, and α as follows:

1) Compute $\partial_\varepsilon K, \partial_\varepsilon W$, and $\partial_\varepsilon \lambda$ following algorithm 2.4 and set the continuation direction $\delta \leftarrow (1, \partial_\varepsilon K, \partial_\varepsilon W, \partial_\varepsilon \lambda)$.
2) Set $\Delta\mathcal{T}_\varepsilon \leftarrow \delta/\|\delta\|$ and $\mathcal{T}_{\varepsilon'} \leftarrow \mathcal{T}_\varepsilon + \Delta\varepsilon\Delta\mathcal{T}_\varepsilon$.
3) Perform Newton steps on $\mathcal{T}_{\varepsilon'}$ by following algorithms 2.2 and 2.3 until $\mathrm{err}_K(\mathcal{T}_{\varepsilon'}) < \varepsilon_K$ and $\mathrm{err}_W(\mathcal{T}_{\varepsilon'}) < \varepsilon_W$ or up to n_{max} times.
4) If $\mathrm{err}_K(\mathcal{T}_{\varepsilon'}) < \varepsilon_K$ and $\mathrm{err}_W(\mathcal{T}_{\varepsilon'}) < \varepsilon_W$, let n_{it} be the number of Newton iterations and go to step 8.
5) Go to step 2 with $\Delta\varepsilon \leftarrow \frac{1}{2}\Delta\varepsilon$ up to n_ε times.
6) Compute $\|t_\theta(K)\|_\infty$ and $\|t_\varphi(K)\|_\infty$.
7) If $\|t_\theta(K)\|_\infty > \varepsilon_t$ and $\|t_\varphi(K)\|_\infty > \varepsilon_t$, set $N_1 \leftarrow 2N_1$ and $N_2 \leftarrow 2N_2$. Otherwise, set $N_1 \leftarrow 2N_1$ if $\|t_\theta(K)\|_\infty > \|t_\varphi(K)\|_\infty$ and $N_2 \leftarrow 2N_2$ if $\|t_\varphi(K)\|_\infty > \|t_\theta(K)\|_\infty$. Go to step 1; try up to n_t times.
8) Set $\mathcal{T}_\varepsilon \leftarrow \mathcal{T}_{\varepsilon'}, \Delta\varepsilon \leftarrow \frac{n_{des}}{n_{it}}\Delta\varepsilon$ and go to step 1.

2.8 Numerical experiments: the Elliptic Restricted Three Body Problem

In this section, we apply our method to the periodic Hamiltonian system given by the ERTBP. The strategy we follow consists of taking families of 2D partially hyperbolic invariant tori in the CRTBP and lift them to the elliptic problem as 3D partially hyperbolic invariant tori through continuation in the eccentricity as described in Section 2.6. In practice, we compute parameterizations of 2D generating tori together with their stable, unstable, and center generating bundles.

For the numerical explorations, we have used a machine under Debian 12 with two processors Intel(R) Xeon(R) CPU E5-2630 0 @ 2.30GHz with 6 cores each and two threads per core for a total of 24 threads. The algorithms were written in C, compiled with GCC

10.2.1 and linked against Glibc 2.31, LAPACK 3.9.0, and FFTW 3.3.8. The numerical integration was parallelized using OpenMP 4.5.

2.8.1 Dynamics

The Elliptic Restricted Three Body Problem models the motion of a small body of negligible mass in the gravitational vector field generated by two other massive bodies, known as primaries, moving in elliptical orbits according to two-body dynamics.

Let us denote by m_1 and m_2 the masses of the primary bodies and let us also define the parameter $\mu := \frac{m_2}{m_1+m_2}$. It is possible to define a rotating and pulsating frame after a suitable rescaling in space and time where the primaries of mass m_1 and m_2 are at $(\mu, 0, 0)$ and $(\mu - 1, 0, 0)$, respectively, and their period of revolution is 2π, see [Vic67] for details. The dynamics for the third body are then given by the periodic Hamiltonian

$$\begin{aligned} H(x, p, \varphi) = \frac{1}{2}[(p_1 + x_2)^2 + (p_2 - x_1)^2 + p_3^2 + x_3^2] \\ - \frac{1}{1 + e\cos f}\left[\frac{1}{2}(x_1^2 + x_2^2 + x_3^2) + \frac{1-\mu}{r_1} + \frac{\mu}{r_2}\right], \end{aligned}$$

with $x \in \mathbb{R}^3$ positions, $p \in \mathbb{R}^3$ momenta, $r_1^2 = (x_1-\mu)^2+x_2^2+x_3^2, r_2^2 = (x_1+1-\mu)^2+x_2^2+x_3^2$, and $e \in \mathbb{R}$ the eccentricity of the elliptic orbit of the primaries. The true anomaly $f := 2\pi\varphi$, with $\varphi \in \mathbb{T}$, is the angle that parameterizes the orbit of the primaries and moves according to the frequency $\dot{\varphi} = \hat{\alpha} = 1/2\pi$.

The ERTBP has five fixed points, known as the Lagrange points, denoted by L_i with $i = 1, 2, ..., 5$. The coordinates of these equilibrium points coincide with the coordinates of the fixed points in the circular problem. The collinear solutions L_1, L_2, and L_3 are unstable for any combination of μ and e whereas the stability of the triangular points, L_4 and L_5, depends on these two parameters. For their values in the Sun-Earth system, that is $\mu = 3.040357143 \cdot 10^{-6}$ and $e = 0.01671123$, the triangular Lagrange points are linearly stable (see e.g. [Vic67] for details).

The persistence of the fixed points in the elliptic problem does not generalize to other invariant objects. Periodic solutions only exist in resonance with the frequency of the primaries. Furthermore, the frequency of the primaries, i.e., the external frequency $\hat{\alpha}$, is added to invariant objects such as periodic orbits and invariant tori of the CRTBP increasing the dimension by one of their counterparts in the elliptic case. That way, (non-resonant) periodic orbits survive as 2D tori while the classical Lissajous and quasi-halo orbits become 3D tori—this will be our case study.

2.8.2 From the CRTBP to the ERTBP

As we mentioned previously, we will lift invariant tori and bundles from the CRTBP to the ERTBP. In this section, we include some details on the families of tori that we computed in the circular problem. Note that we can obtain the Hamiltonian of the CRTBP from that of the ERTBP for $e = 0$.

The L_1 point in the CRTBP is of type center×center×saddle; that is

$$\mathrm{Spec}\mathrm{D}X_H(L_1) = \{\boldsymbol{i}2\pi\hat{\omega}_p^0, -\boldsymbol{i}2\pi\hat{\omega}_p^0, \boldsymbol{i}2\pi\hat{\omega}_v^0, -\boldsymbol{i}2\pi\hat{\omega}_v^0, \lambda^0, -\lambda^0\},$$

with a value of the Hamiltonian h^0. Thus, there is a 4-dimensional center manifold generated by the central part of L_1 that contains invariant objects.

The Lyapunov center theorem (see e.g. [MHO09, SM71]) ensures that there exists two 2-dimensional manifolds inside the center manifold filled with families of periodic orbits: the planar and the vertical Lyapunov families. These families present bifurcations that give birth to new families such as the halo family and other more exotic orbits (see e.g. [DDP03]). We can parameterize the orbits in these families in a large neighborhood of the L_1 point by their value of the Hamiltonian.

Besides the 2-dimensional manifolds of periodic orbits, the center manifold is also filled with 2-dimensional tori that exist around periodic orbits with a central part. Let us focus on the tori around vertical Lyapunov orbits. Let $\hat{\omega}_p$ and $\hat{\omega}_v$ be the frequencies of any torus such that $\hat{\omega}_p \to \hat{\omega}_p^0$ and $\hat{\omega}_v \to \hat{\omega}_v^0$ when $h \to h^0$, see Remark 2.3 on the non-uniqueness of the frequencies. Following [HM21], we define the rotation number as

$$\rho := \hat{\omega}_p/\hat{\omega}_v - 1.$$

We can then use the Hamiltonian h and ρ to represent the family of 2-dimensional tori around the vertical Lyapunov orbits. These two parameters uniquely determine each torus. Analogous representations can be obtained for the tori around planar Lyapunov and halo orbits, see [GM01, HM21] for a full description.

For $\mu = 3.040357143 \cdot 10^{-6}$ and $e = 0$, we selected 77 equally spaced values of ρ between 0.03565 and 0.0961 for the Sun-Earth system and "nobilized"[1] them with an absolute tolerance of 1.6×10^{-4}. Then, we performed continuations in the flying time T for each value of $\omega = \rho$, following the methodology from [HM21], and obtained the family of tori around vertical Lyapunov orbits in the CRTBP. The results are gathered in Fig. 2.2 where we plot the rotation number and the Hamiltonian as well as the grid size of the parameterizations in the color map for each of the 8971 tori computed.

2.8.3 Numerical explorations in the ERTBP

In this case study, the generating tori $\mathcal{K}$ are 2-dimensional and the generated tori $\hat{\mathcal{K}}$ are 3-dimensional. To initialize the algorithms, we use the autonomous tori from the CRTBP with a grid size in the internal phase as given in Fig. 2.2. For the initial grid size in the external phase, we took $N_2 = 16$. Since the generating tori in the CRTBP are 1-dimensional, we obtain the 2-dimensional generators by setting the approximate Fourier coefficients $\hat{\xi}_{kj}$ to zero for $|j| > 0$. Equivalently, we can construct the generators as in (2.5). In algorithm 2.5, we use $\varepsilon_K = 10^{-9}, \varepsilon_W = 10^{-5}, \varepsilon_t = 10^{-9}, r_t = 1/5, n_{max} = 6, n_\varepsilon = 3, n_{des} = 4$, and $n_t = 2$. We set a maximum of 2048 Fourier coefficients in each phase and for the multiple shooting approach, we take $m = 4$. We set the continuations to reach the eccentricity of the Sun-Earth system; that is, $e = 0.01671123$. It is worth mentioning that we set

[1]A noble number is one whose continued fraction expansion coefficients are equal to one from a position onward.

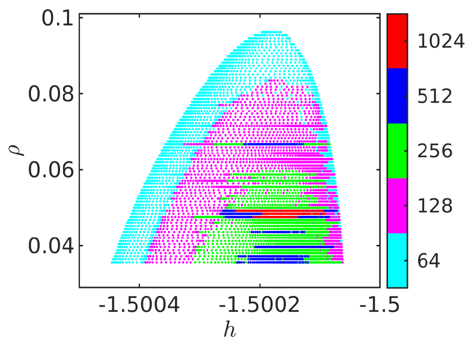

Figure 2.2: Energy-rotation number representation in the CRTBP for the Lissajous family around vertical orbits around L_1 for the Sun-Earth mass parameter. The color map represents the grid size of the parameterizations.

$\varepsilon_K = 10^{-9}$ for such an exhaustive numerical exploration, but we manage to get errors in the invariance equations of the order of 10^{-15} for some tori.

The results of our numerical explorations are gathered in Figures 2.3 and 2.4. We emphasize that these results constitute an improvement with respect to the results presented in [FMHM24b] due to improvements in the continuation strategy. Out of the 8971 tori computed in the CRTBP, 6840 reached the Sun-Earth eccentricity. Each torus in the Figures is labeled by its value of the internal rotation number and the value of the Hamiltonian of the torus in the CRTBP used to initiate the continuations. Note that we use h simply as a label with no dynamical implications. In the color map of Fig. 2.3, we represent the grid size in the internal phase θ (left) and the external phase φ (right) only of the tori that reached the eccentricity of the Sun-Earth system and in 2.4 we represent the value of the unstable rate of expansion of the invariant tori.

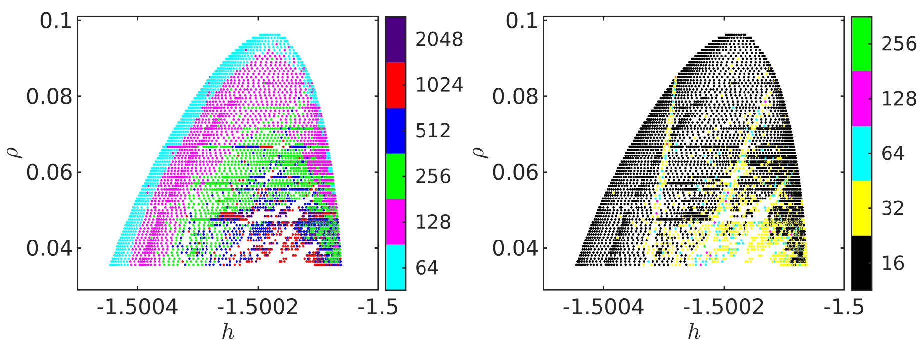

Figure 2.3: Energy-rotation number representation in the ERTBP for the Lissajous family around vertical orbits around L_1 for the Sun-Earth mass parameter. The color map represents the grid size of the internal phase (left) and the external phase (right).

We can observe that not all continuations reach the required value of the eccentricity. Certain empty “lines", where the continuations fail, are present, which suggests the existence of a dynamical barrier. It turns out that these lines correspond to resonances between the frequencies of the 3-dimensional tori, see Section 2.8.4. For values of $\rho < 0.06$, there is also a gap in the energy-rotation number representation. The lack of convergence in this region is related to resonances and also to the existence of homoclinic and heteroclinic tangles. We will come back to this issue in Section 2.8.5.

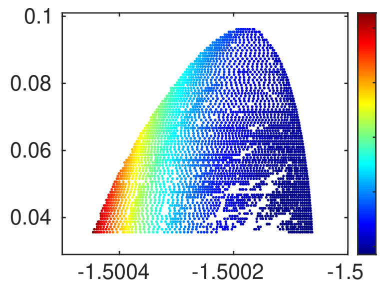

Figure 2.4: Energy-rotation number representation in the ERTBP for the Lissajous family around vertical orbits around L_1 for the Sun-Earth mass parameter. The color map represents the unstable rate of expansion

From the numerical explorations we can see that, generally, tori in the ERTBP do not need a large number of approximate Fourier coefficients in the external phase, see Fig. 2.3 (right), which suggests that our tori are more analytic in the phase φ. Note that φ is essentially time and tori tend to be more analytic in the temporal direction, which is in agreement with our results.

Close to resonances, the number of coefficients needed increases. We also observe an increase in the number of coefficients for the tori surrounding the region where homoclinic and heteroclinic tangles are present. Resonances and homoclinic and heteroclinic tangles are responsible for the breakdown of tori, see [Chi79, dlLO06, OS87]. The fact that more coefficients are necessary for the parameterizations close to such cases reveals that the tori are losing regularity because they are breaking down.

In order to lift the autonomous tori to the elliptic problem, we see that it does not suffice to simply add coefficients in the external phase. When we compare Fig. 2.2 and Fig. 2.3 (left), the number of coefficients in the internal phase changes. Throughout the continuations, it was seen that for a given torus, sometimes it was necessary to increase the number of coefficients in the phase θ and sometimes in the phase φ. It is therefore key to control the decay of the Fourier coefficients as it has been described in Section 2.7.

Lastly, we include in Fig. 2.8.3 some plots in the configuration space of 3-dimensional tori (red) and their 2-dimensional generators (black) for a family with $\rho = 0.071461$. They can be seen as "fattened" with respect to their CRTBP. Such "fattening" is clearly seen in the results from Section 2.8.5.

The behavior of the family of tori for fixed ρ is qualitatively very similar to their autonomous counterparts in the CRTBP, see [HM21]. The family begins with a torus close to a vertical Lyapunov orbit and, for higher energies, tori increase in size and start to bend. Then, they approach a vertical Lyapunov orbit of higher energy where the family collapses.

Remark 2.9. To obtain the results presented in Fig. 2.3, we did several runs of Algorithm 2.5 for different values of the filtering factor for the low-pass filter described in Section 2.7.

To give an idea of computational times, we provide in Table 2.1 details for a few continuations from $e = 0$ to $e = 0.01671123$. That is, a few continuations of tori in the CRTBP to the Sun-Earth ERTBP, see Algorithm 2.5. In the table we include: the initial and final number of Fourier coefficients, N_{F_0} and N_{F_f} respectively, the value of the Hamiltonian

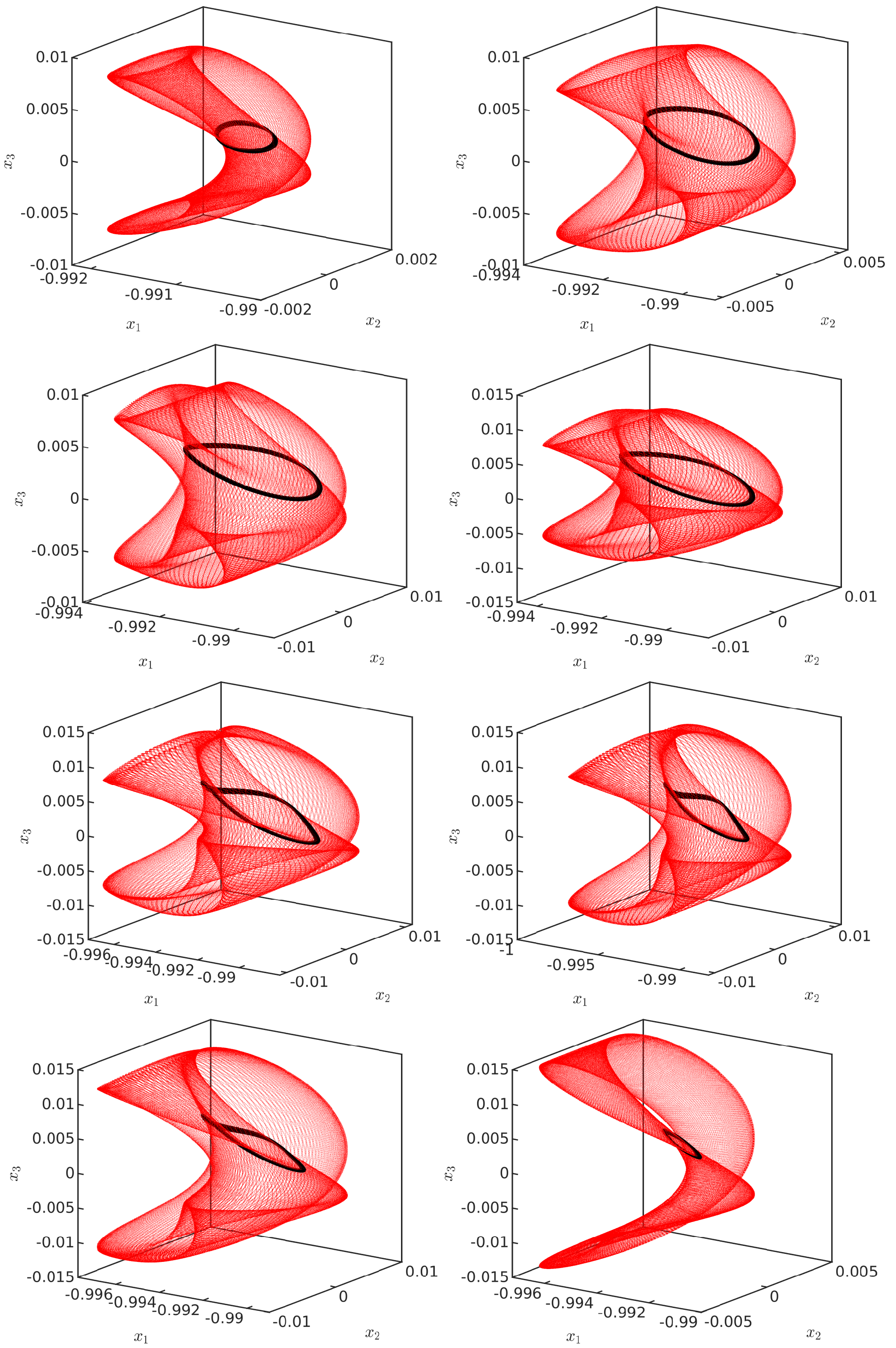

Figure 2.5: Projection onto the configuration space of 3-dimensional generated tori (red) and 2-dimensional generating tori (black) for the family $\rho = 0.071461$.

of the initial autonomous seed h, the rotation number of the invariant curve of the autonomous seed, the number of continuation steps, the norms for the torus and bundle invariance errors, $\|E_K\|$ and $\|E_W\|$ respectively, and the elapsed time. The norms that we use are those in (2.78)-(2.79). For the specifications of the machine that we used see the second paragraph of Section 2.8. Note that when $N_{F_0} \neq N_{F_f}$, the continuations require to increase the number of coefficients because convergence was not possible for N_{F_0}. This incurs naturally in larger computational times. See Algorithm 2.5.

Table 2.1: Computational times for continuations from $e = 0$ to $e = 0.01671123$.

N_{F_0}	N_{F_f}	h	ρ	steps	$\|E_K\|$	$\|E_W\|$	t
$(64, 16)$	$(64, 16)$	-1.50032	0.0747	2	$1.72 \cdot 10^{-12}$	$9.25 \cdot 10^{-10}$	1.81s
$(64, 16)$	$(128, 16)$	-1.50034	0.0507	2	$6.39 \cdot 10^{-13}$	$1.10 \cdot 10^{-10}$	2.94s
$(128, 16)$	$(256, 32)$	-1.50011	0.0547	3	$9.98 \cdot 10^{-11}$	$2.81 \cdot 10^{-8}$	36.43s
$(256, 16)$	$(256, 16)$	-1.50022	0.0411	2	$6.75 \cdot 10^{-11}$	$7.29 \cdot 10^{-9}$	6.63s
$(256, 16)$	$(512, 32)$	-1.50020	0.0460	5	$2.73 \cdot 10^{-11}$	$4.02 \cdot 10^{-7}$	77.31s
$(1024, 16)$	$(1024, 16)$	-1.50024	0.0492	2	$5.90 \cdot 10^{-11}$	$3.92 \cdot 10^{-8}$	33.27s

2.8.4 Resonances

The vector of internal frequencies $\hat{\omega}$ and the vector of external frequencies $\hat{\alpha}$ were assumed to be sufficiently non-resonant; that is, Diophantine. When this assumption does not hold, there is a dynamical obstruction to the existence of invariant tori.

Let $\kappa = (\kappa_1, \kappa_2, \kappa_3)$ be a 3-tuple. The vectors of frequencies $\hat{\omega} = (\frac{\rho}{T}, \frac{1}{T})$ and $\hat{\alpha} = \frac{1}{2\pi}$ are in p-order resonance when $|\kappa|_1 = p$ and

$$\mathcal{R}_\kappa := \kappa_1 \frac{\rho}{T} + \kappa_2 \frac{1}{T} + \kappa_3 \frac{1}{2\pi}, \tag{2.80}$$

becomes zero. The numerical explorations were done for fixed values of ρ but throughout the continuations, the flying time T varies. For certain κ, ρ, and T, $\mathcal{R}_k$ might become small, revealing that the continuations are approaching resonances. Note that for given κ, with $|\kappa|_1 = p$, $\mathcal{R}_k = 0$ defines p-resonant lines. To visualize these resonances in the energy-rotation number representation, we first examine the tori computed in the CRTBP. We look for 3-tuples $\bar{\kappa}$ such that $\mathcal{R}_{\bar{\kappa}} < \varepsilon^{\mathcal{R}}$ up to a maximum order $\bar{p}$. We took $\varepsilon^{\mathcal{R}} = 10^{-4}$ and $\bar{p} = 10$. Once we obtain the 3-tuples $\bar{\kappa}$, we compute the lines $\mathcal{R}_{\bar{\kappa}} = 0$ and obtain $\rho_{\bar{\kappa}} = \rho_{\bar{\kappa}}(T)$. Then, we use the values of h, ρ, and T of the grid of tori computed in the CRTBP to obtain through inverse cubic interpolation, for each line $\mathcal{R}_{\bar{\kappa}} = 0$, the curve $\rho_{\bar{\kappa}} = \rho_{\bar{\kappa}} \circ T(h)$. The results are shown in Fig. 2.6.

It is clear that when lifting the tori from the CRTBP to the ERTBP, the frequencies can become resonant for a large subset of the autonomous tori. We observe that for $h \gtrsim -1.5002$, there is an accumulation of resonant curves in the energy-rotation number representation which explains the gap of Figures 2.3 and 2.4. The accumulation of resonant

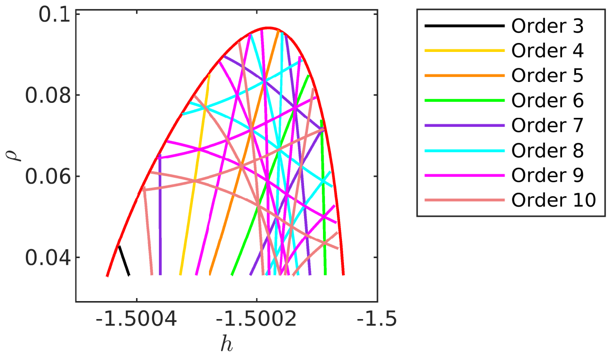

Figure 2.6: Resonant curves up to order 10 for the tori in the ERTBP for the Sun-Earth mass parameters.

curves also explains why, when comparing Fig. 2.2 and Fig 2.3 (left), more coefficients are necessary for the parameterizations of the tori that reached the Sun-Earth eccentricity.

Note the correspondence between the resonant curves and the results from Fig. 2.3—the absence of tori in the region surrounded by tori with $N_2 = 32$ in Fig. 2.2 (right) corresponds to order $4, 6$, and 8 resonances. The fact that even order resonances cause larger gaps of non-existence of invariant tori is likely caused by symmetries but we did not investigate this phenomenon further. Other resonant curves can be appreciated but are more subtle. Due to the presence of resonances, performing continuations for each torus in the CRTBP is a more robust approach than performing them in the ERTBP. When the method tries to compute a resonant torus, it will simply fail and move to the next, whereas if continuations where done in the ERTBP, the method would have to jump through all the resonances it encounters.

2.8.5 Poincaré representation

As we pointed out, the Hamiltonian in the ERTBP is not constant. Therefore, we cannot obtain the isoenergetic sections of the center manifold commonly seen in studies for the CRTBP, see e.g. [JM99, GM01, GJSM01a]. Nonetheless, in this section we show some analogous results. In addition to the numerical results already presented, we have also taken tori in the CRTBP around vertical and halo orbits within a level set of the Hamiltonian, lifted them to the Sun-Earth ERTBP, and computed a Poincaré section with $\Sigma = \{x, p \in \mathbb{R}^3 |\ x_3 = 0, p_3 > 0\}$.

We explore the level set $H = -1.5002$. The results are gathered in Fig. 2.7. On the left, we show (in red) the intersections of the tori of the CRTBP that, when used as seeds of continuations in the eccentricity e, reach the Sun-Earth ERTBP. The intersections of the 3-dimensional tori in the Sun-Earth ERTBP are shown in Fig. 2.7 right. In order to have some reference of where the families of tori end, we show in black the intersections of the last tori computed in the CRTBP around vertical and halo orbits.

We first observe that since the tori in the ERTBP are 3-dimensional, the intersections with Σ are 2-dimensional. This allows us to somewhat see the “fattening” of each torus due to the time dependency of the Hamiltonian.

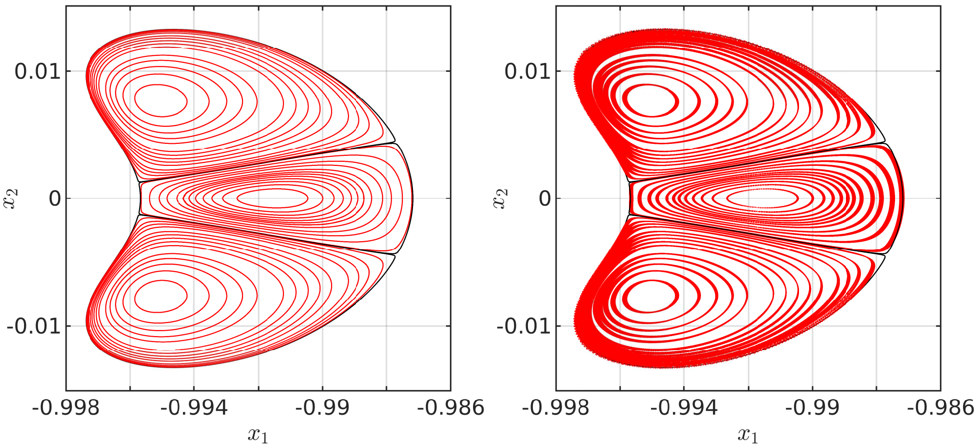

Figure 2.7: Intersections in red of Lissajous and quasi-halo tori with Σ in the CRTBP (left) and in the ERTBP (right). In black, we show the intersections of the last tori computed for each family in the CRTBP. The tori were computed from the level set $H = -1.5002$ in the CRTBP.

In the CRTBP, there is a range of the Hamiltonian where there exist tori that approach double homoclinic connections of planar Lyapunov orbits. For larger energy values, there exist heteroclinic connections between the so called “axial” orbits. These connections act as separatrices between the Lissajous and quasi-halo families. An example of a torus approaching such connections is shown in Fig. 8 of [HM21], where a torus around a vertical orbit approaches two vertically symmetric quasi-halo tori. The existence of transverse homoclinic and heteroclinic points is one known mechanism for breakdown of invariant tori, see [Chi79, dlLO06, OS87], and the tori in the CRTBP that approach connections were found for small values of ρ, see [HM21].

Non-resonant periodic orbits in the CRTBP survive as 2-dimensional tori in the ERTBP, so there might be even more connections than in the CRTBP. Together with the fact that tori in the ERTBP are higher dimensional, suggests that there might be a larger set of tori in the ERTBP that are sufficiently close to homoclinic and heteroclinic connections for the method to fail.

Lastly, in Fig. 2.8, we show in the configuration space plots of the largest Lissajous (red) and quasi-halo (blue) tori, and their projections, computed for Fig. 2.7. That is, the largest Lissajous and quasi-halo tori that reached the Sun-Earth ERTBP from the level set $H = -1.5002$ in the Sun-Earth CRTBP. For easier visualization, of the two symmetric quasi-halo tori of the energy level, only one is shown. We can observe the proximity between the tori in the configuration space (similar plots can be obtained with other projections) and how they almost merge.

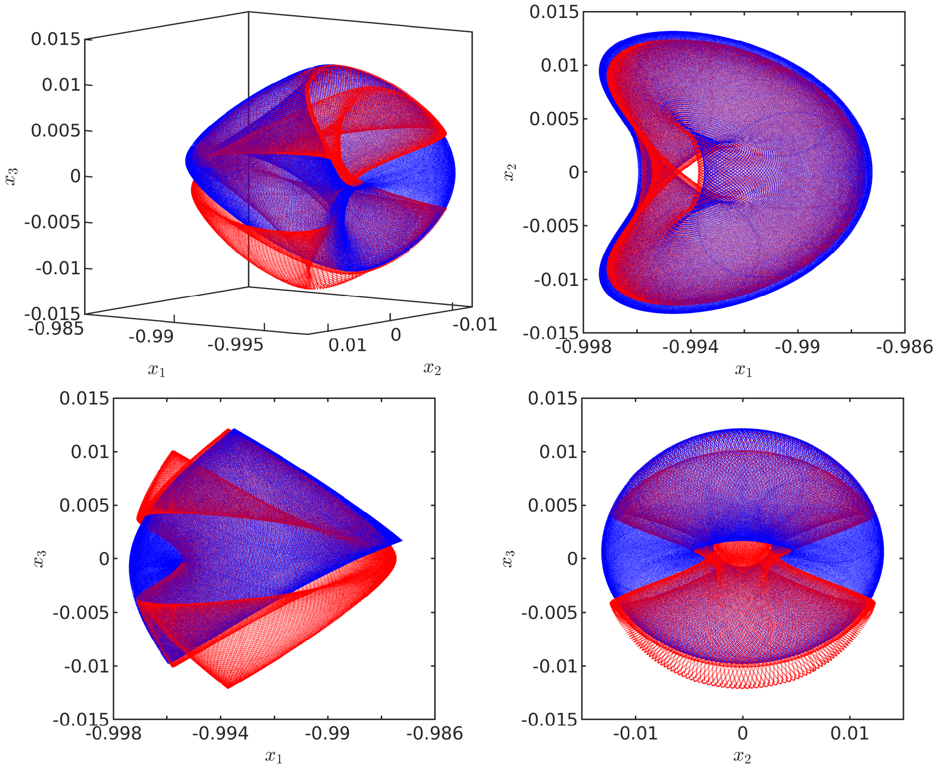

Figure 2.8: Largest Lissajous (red) and quasi-halo (blue) tori, and their projections, computed in the Sun-Earth ERTBP from the level set $H = -1.5002$ in the CRTBP

Chapter 3

A-posteriori KAM theorem

In this chapter, we provide the proof of the convergence of the methods of Chapter 2. That is, we build on [HM21, FMHM24b] and obtain an *a-posteriori* theorem for partly hyperbolic invariant tori and their invariant bundles in quasi-periodic Hamiltonian systems. The a-posteriori formulation summarizes as follows. If there is an approximate solution of the invariance equation satisfying some non-degeneracy conditions, then there is a true solution "nearby". This kind of theorems require some initial data from where a condition for existence can be checked in a finite number of computations. In combination with rigorous numerical techniques, it is possible to perform computer-assisted proofs. This approach led in [FHL17] to the formulation of a general methodology for such kind of proofs—using the parameterization method—for Lagrangian invariant tori. A modified approach that leads to sharp estimates for the condition for existence was introduced in [Vil17] and similar ideas for systems with extra first integrals were developed in [FH24]. For Lagrangian tori in quasi-periodic Hamiltonians, see [CHP25].

As in Chapter 2, we use flow maps related to the natural frequencies of tori to reduce the dimension of the tori by one and do not use an autonomous reformulation by introducing fictitious *action* coordinates. This point of view allows the theorem to be applied to both autonomous and quasi-periodic Hamiltonian systems and constitutes the proof on the convergence of the flow map methods of [HM21, FMHM24b]. Additionally, we obtain both the stable and unstable bundles simultaneously which provides a clear geometrical picture of the tangent space on the torus. The proof relies on geometric properties of symplectic flavour that hold approximately when we have parameterizations that satisfy their invariance equations approximately. We obtain geometric lemmas for such properties that, together with the lemma on quadratically small averages obtained in Appendix B, control the error in the KAM iterative procedure. We emphasize that convergence can still be proved without such lemma but then we would need to use a translated curve theorem and a vanishing lemma. The new error in the invariance equations is controlled by explicit constants that we provide in Appendix E which require the careful treatment of the analyticity loss with each iterative step. The proof concludes by obtaining conditions for convergence of the KAM iterative procedure.

The structure of the chapter is the following. In Section 3.1 we describe partly hyperbolic invariant tori and our main result: the KAM theorem. In Section 3.2 we state and prove the geometric lemmas that we use to control the new errors after one step of the iterative procedure; for which we obtain estimates in Section 3.3. Lastly, in Section 3.4 we conclude the proof of the KAM theorem and in Appendix E we provide tables with all the explicit constants obtained in our estimates.

3.1 Existence of partially hyperbolic invariant tori

In this chapter, we provide the proof on the convergence of the iterative method of Chapter 2—let us recall some key ideas.

We are interested in $(d+\ell+1)-$dimensional tori with frequencies $\hat{\omega} \in \mathbb{R}^{d+1}$ and $\hat{\alpha} \in \mathbb{R}^{\ell}$—invariant for the vector field (1.1)—and in their invariant bundles. Equivalently, for $\hat{\omega} =: \frac{1}{T}(\omega, 1)$ and $\alpha := \hat{\alpha}T$, where $\omega \in \mathbb{T}^d$ and $\alpha \in \mathbb{T}^{\ell}$, we can look for $(d+\ell)-$dimensional tori $\mathcal{K}$, invariant for the flow map ϕ_T, and in bundles $\mathcal{W}$ with base $\mathcal{K}$, invariant under the differential of ϕ_T, see Sections 2.1 and 2.2.

In the spirit of the parameterization method, we are interested in fiberwise embeddings $K : \mathbb{T}^{d+\ell} \to U$ and $W : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$, with rate $\lambda \in \mathbb{R}/\{1, -1\}$, satisfying (2.3) and (2.6). Without loss of generality, we can assume $0 < |\lambda| < 1$. We will refer to K and W as the parameterizations of $\mathcal{K}$ and $\mathcal{W}$, respectively.

It can be shown, see Section 3.2.6 and 3.2.10, that if (K, W, λ) satisfy (2.3) and (2.6), then there exists a set of coordinates given by a frame P such that the differential of the flow map reduces to an upper triangular matrix-valued map—constant along the diagonal. That is, the linearized dynamics in the coordinates given by P reduce to $\Lambda_P : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$ as

$$(P \circ \mathcal{R}_{\omega\alpha})^{-1} \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) P = \Lambda_P, \tag{3.1}$$

where

$$\Lambda_P := \left(\begin{array}{c|c} \Lambda & S \\ \hline & \Lambda^{-\top} \end{array}\right).$$

We can construct the frame P as described in Section 2.4.

The convergence of the iterative process described in Section 2.5 leads to *a-posteriori* KAM theorem on existence of partially hyperbolic invariant tori and its invariant bundles. In what follows, recall the definitions for the sets $\mathcal{U}$ and $\tilde{\mathcal{U}}$ from Section 1.3. We now state the main result.

Theorem 3.1 (KAM theorem for partially hyperbolic tori)**.** *Let $\boldsymbol{\omega} = \mathrm{d}\boldsymbol{\alpha}$ be an exact symplectic form defined in an open set $U \subset \mathbb{R}^{2n}$ and let $\boldsymbol{J}$ be an almost complex structure on U compatible with $\boldsymbol{\omega}$. Let $\boldsymbol{g}$ be the induced metric and $H : U \times \mathbb{T}^{\ell} \to \mathbb{R}$ be a real-analytic quasi-periodic Hamiltonian with frequencies $\hat{\alpha} \in \mathbb{R}^{\ell}$. For $d = n-2$, let $\omega \in \mathbb{R}^d$, $T > 0$, and $\alpha = \hat{\alpha}T$ be such that (ω, α) is Diophantine with constants $\gamma > 0$, $\tau \geq d+\ell$ and suppose we have $K : \mathbb{T}^{d+\ell} \to U$, $W : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n}$, and $|\lambda| < 1$. Assume the following hypothesis.*

H_1 *There exists some $r, R > 0$ such that the geometric objects $\boldsymbol{\omega}$, $\boldsymbol{\alpha}$, $\boldsymbol{J}$, and $\boldsymbol{g}$ can be analytically extended to the complex set $\mathcal{U} \subset \mathbb{C}^{2n}$ containing U. Moreover, for all $t \in [0, T]$ and all $(z, \varphi) \in \tilde{\mathcal{U}}$, $\tilde{\phi}_t(z, \varphi)$ exists. Then, there exist constants such that:*

– *the matrix representations* Ω, a, J, *and* G *satisfy*

$$\begin{array}{cccc} \|\Omega\|_{\mathcal{U}} \leq c_{\Omega}, & \|\mathrm{D}\Omega\|_{\mathcal{U}} \leq c_{\mathrm{D}\Omega}, & \|\mathrm{D}a\|_{\mathcal{U}} \leq c_{\mathrm{D}a}, & \|(\mathrm{D}a)^{\top}\|_{\mathcal{U}} \leq c_{(\mathrm{D}a)^{\top}}, \\ \|\mathrm{D}^2 a\|_{\mathcal{U}} \leq c_{\mathrm{D}^2 a}, & \|J\|_{\mathcal{U}} \leq c_J, & \|J^{\top}\|_{\mathcal{U}} \leq c_{J^{\top}}, & \|\mathrm{D}J\|_{\mathcal{U}} \leq c_{\mathrm{D}J}, \\ \|(\mathrm{D}J)^{\top}\|_{\mathcal{U}} \leq c_{(\mathrm{D}J)^{\top}}, & \|G\|_{\mathcal{U}} \leq c_G, & \|\mathrm{D}G\|_{\mathcal{U}} \leq c_{\mathrm{D}G}. & \end{array}$$

– *The Hamiltonian* H, *the vector field* X, *and the flow map* ϕ_T *satisfy*

$$\begin{array}{cccc} \|X\|_{\tilde{\mathcal{U}}} \leq c_X, & \|X^{\top}\|_{\tilde{\mathcal{U}}} \leq c_{J^{\top}}, & \|\mathrm{D}_z X\|_{\tilde{\mathcal{U}}} \leq c_{\mathrm{D}_z X}, & \|(\mathrm{D}_z X)^{\top}\|_{\tilde{\mathcal{U}}} \leq c_{(\mathrm{D}_z X)^{\top}}, \\ \|\mathrm{D}^2_z H\|_{\tilde{\mathcal{U}}} \leq c_{\mathrm{D}^2_z H}, & \|\mathrm{D}_z \phi_T\|_{\tilde{\mathcal{U}}} \leq c_{\mathrm{D}_z \phi_T}, & \|(\mathrm{D}_z \phi_T)^{\top}\|_{\tilde{\mathcal{U}}} \leq c_{(\mathrm{D}_z \phi_T)^{\top}}, & \|\mathrm{D}^2_z \phi_T\|_{\tilde{\mathcal{U}}} \leq c_{\mathrm{D}^2_z \phi_T}. \end{array}$$

H_2 *There exists* $\rho \in (0, r)$ *such that* $K \in \mathcal{A}(\mathbb{T}^{d+\ell}_{\rho})^{2n}$, $\mathrm{D}K \in \mathcal{A}(\mathbb{T}^{d+\ell}_{\rho})^{2n \times (d+\ell)}$, *and* $W \in \mathcal{A}(\mathbb{T}^{d+\ell}_{\rho})^{2n}$.

– *Moreover, there exist constants such that*

$$\begin{array}{cccc} \|\mathrm{D}_\theta K\|_{\rho} < \sigma_{\mathrm{D}_\theta K}, & \|(\mathrm{D}_\theta K)^{\top}\|_{\rho} < \sigma_{(\mathrm{D}_\theta K)^{\top}}, & \|\mathrm{D}_\varphi K\|_{\rho} < \sigma_{\mathrm{D}_\varphi K}, & \|(\mathrm{D}_\varphi K)^{\top}\|_{\rho} < \sigma_{(\mathrm{D}_\varphi K)^{\top}}, \\ \|W\|_{\rho} < \sigma_W, & \|W^{\top}\|_{\rho} < \sigma_{W^{\top}}, & \|\hat{N}\|_{\rho} < \sigma_{\hat{N}}, & \|\hat{N}^{\top}\|_{\rho} < \sigma_{\hat{N}^{\top}}, \\ \|B\|_{\rho} < \sigma_B. & & & \end{array}$$

– *Also, there exist* $\sigma_\lambda < 1$ *and* $\sigma_{\lambda^{-1}} > 1$ *satisfying*

$$|\lambda| < \sigma_\lambda, \ |\lambda^{-1}| < \sigma_{\lambda^{-1}}.$$

H_3 *There exists a constant* $\sigma_{\langle \mathcal{S} \rangle^{-1}}$ *such that the shear* $\mathcal{S}$ *in* (2.30), *which coincides with the block* $\hat{S}_1$ *in* (2.32), *satisfies the non-degeneracy condition* $\left| \langle \mathcal{S} \rangle^{-1} \right| < \sigma_{\langle \mathcal{S} \rangle^{-1}}$.

Then, for every $\rho_\infty \in (0, \rho)$ *and* $\delta \in (0, \frac{\rho - \rho_\infty}{3})$, *there exist a constant* $\mathfrak{C}$ *such that if the invariance errors*

$$\begin{aligned} E_K &= \phi_T \circ (K, \mathrm{id}) - K \circ \mathcal{R}_{\omega\alpha}, \\ E_W &= \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) W - W \circ \mathcal{R}_{\omega\alpha} \lambda, \end{aligned}$$

satisfy

$$\frac{\mathfrak{C}}{\gamma^2 \delta^{2\tau}} \mathcal{E} < 1, \tag{3.2}$$

where

$$\mathcal{E} := \max \left\{ \frac{\|E_K\|_{\rho}}{\gamma^2 \delta^{2\tau}}, \ \|E_W\|_{\rho} \right\}, \tag{3.3}$$

then there exist λ_∞, *an invariant torus* $\mathcal{K}_\infty = K_\infty(\mathbb{T}^{d+\ell})$ *and an invariant bundle* $\mathcal{W}_\infty = W_\infty(\mathbb{T}^{d+\ell})$ *satisfying:*

i) $K_\infty, W_\infty \in \mathcal{A}(\mathbb{T}^{d+\ell}_{\rho_\infty})^{2n}$ *and* $K_\infty(\mathbb{T}^{d+\ell}_{\rho_\infty}) \subset \mathcal{U}$.

ii) K_∞, W_∞, *and the associated objects* $\hat{N}_\infty, B_\infty$, *and* $\mathcal{S}_\infty$ *satisfy* H_2 *and* H_3 *in the complex strip* $\mathbb{T}^{d+\ell}_{\rho_\infty}$.

iii) *The dynamics on $\mathcal{K}_\infty$ are conjugated to a rigid rotation with rotation vector (ω, α) and on $\mathcal{W}_\infty$ to a constant contraction with rate $|\lambda_\infty|$.*

iv) *The limiting objects are close to the original ones in the following sense*

$$\|K_\infty - K\|_{\rho_\infty} \le \mathfrak{C}_{\Delta K}\mathcal{E}, \tag{3.4}$$

$$\|W_\infty - W\|_{\rho_\infty} \le \frac{\mathfrak{C}_{\Delta W}}{\gamma\delta^\tau}\mathcal{E}, \tag{3.5}$$

$$|\lambda_\infty - \lambda| \le \mathfrak{C}_{\Delta\lambda}\mathcal{E}, \tag{3.6}$$

$$|\lambda_\infty^{-1} - \lambda^{-1}| \le \mathfrak{C}_{\Delta\lambda^{-1}}\mathcal{E}. \tag{3.7}$$

Remark 3.1. In hypothesis H_2, it is not strictly necessary to assume bounds for $\hat{N}$ since we can control them automatically from the bounds on J, L, and B. Nonetheless, we include it in the hypothesis for the sake of flexibility in order apply the theorem to particular problems.

Remark 3.2. Note that since we use the matrix-valued map A to reduce the torsion to the form (2.30), we obtain a parameterization $\widetilde{W}_\infty$ of the complementary invariant bundle to $\mathcal{W}_\infty$—$\widetilde{\mathcal{W}}_\infty$—satisfying (2.31).

Remark 3.3. The theorem applies to autonomous systems with minor modifications—it only requires setting to zero some quantities.

3.2 Error estimates of approximate geometric properties

For the proof of Theorem 3.1, we first need to control certain geometric properties that only hold approximately when the parameterizations of tori and bundles satisfy their invariance equations approximately. In this section we obtain upper bounds for these properties in terms of the error in the invariance equations.

3.2.1 Adapted frames control

Let us begin by controlling the adapted frame we use throughout the iterative scheme. Let

$$C_{\mathcal{X}} := c_X + \sigma_{\mathrm{D}_\varphi K}|\hat{\alpha}|, \qquad C_{\mathcal{X}^\top} := c_{J^\top} + |\hat{\alpha}^\top|\sigma_{(\mathrm{D}_\varphi K)^\top}, \tag{3.8}$$

$$C_L := \sigma_{\mathrm{D}_\theta K} + C_{\mathcal{X}} + \sigma_W, \qquad C_{L^\top} := \max\left\{\sigma_{(\mathrm{D}_\theta K)^\top}, C_{\mathcal{X}^\top}, \sigma_{W^\top}\right\}. \tag{3.9}$$

Then,

$$\|\mathcal{X}\|_\rho < C_{\mathcal{X}}, \qquad \|\mathcal{X}^\top\|_\rho < C_{\mathcal{X}^\top},$$

$$\|L\|_\rho < C_L, \qquad \|L^\top\|_\rho < C_{L^\top}.$$

If we define $C_{\hat{P}} := C_L + \sigma_{\hat{N}}$ and $C_{\hat{P}^\top} := \max\{C_{L^\top},\ \sigma_{\hat{N}^\top}\}$, we can control the frame $\hat{P}$ as

$$\|\hat{P}\|_\rho \le C_{\hat{P}}, \qquad \|\hat{P}^\top\|_\rho \le C_{\hat{P}^\top}. \tag{3.10}$$

Similarly, let $C_{\hat{S}} := \sigma_{\hat{N}^\top} c_\Omega c_{\mathrm{D}_z\phi_T}\sigma_{\hat{N}}$ and $C_{\hat{S}^\top} := \sigma_{\hat{N}^\top} c_\Omega c_{(\mathrm{D}_z\phi_T)^\top}\sigma_{\hat{N}}$. Hence we have the following estimates for $\hat{S}$ and $\hat{S}^\top$

$$\|\hat{S}\|_\rho < C_{\hat{S}}, \qquad \|\hat{S}^\top\|_\rho < C_{\hat{S}^\top}. \tag{3.11}$$

In order to control the frame P, let us first obtain estimates for $\|A\|_\rho$ and $\|A^\top\|_\rho$ as follows

$$\|A\|_\rho \leq \max\{\|A_2\|_\rho,\ \|A_3\|_\rho + \|A_4\|_\rho\},$$
$$\|A^\top\|_\rho \leq \max\left\{\|A_3^\top\|_\rho,\ \|A_2^\top\|_\rho + \|A_4\|_\rho\right\}.$$

Note that A_2 and A_3 satisfy (2.37) and (2.38), respectively. Therefore, we can use estimate (1.13) and obtain

$$\|A_2\|_\rho \leq \frac{1}{|\lambda^{-1}| - |\lambda|}\left(\|\hat{S}_2\|_\rho|\lambda| + \|\hat{S}_3^\top\|_\rho|\lambda^{-1}|\right), \quad \|A_3\|_\rho \leq \frac{1}{|\lambda^{-1}| - |\lambda|}\left(\|\hat{S}_3\|_\rho + \|\hat{S}_2^\top\|_\rho\right),$$
$$\|A_2^\top\|_\rho \leq \frac{1}{|\lambda^{-1}| - |\lambda|}\left(\|\hat{S}_2^\top\|_\rho|\lambda| + \|\hat{S}_3\|_\rho|\lambda^{-1}|\right), \quad \|A_3^\top\|_\rho \leq \frac{1}{|\lambda^{-1}| - |\lambda|}\left(\|\hat{S}_3^\top\|_\rho + \|\hat{S}_2\|_\rho\right).$$

Also, A_4 satisfies (2.36) so

$$\|A_4\|_\rho \leq \frac{1}{|\lambda^{-1}| - |\lambda|}\|\hat{S}_4\|_\rho.$$

Then,

$$\begin{aligned}
\|A\|_\rho &\leq \frac{1}{|\lambda^{-1}| - |\lambda|}\max\left\{\|\hat{S}_2\|_\rho|\lambda| + \|\hat{S}_3^\top\|_\rho|\lambda^{-1}|,\ \|\hat{S}_3\|_\rho + \|\hat{S}_2^\top\|_\rho + \|\hat{S}_4\|_\rho\right\} \\
&\leq \frac{1}{|\lambda^{-1}| - |\lambda|}\max\left\{\|\hat{S}\|_\rho + \|\hat{S}_3^\top\|_\rho|\lambda^{-1}|,\ \|\hat{S}\|_\rho + \|\hat{S}_2^\top\|_\rho|\lambda^{-1}|\right\} \\
&\leq \frac{1}{|\lambda^{-1}| - |\lambda|}\left(\|\hat{S}\|_\rho + \|\hat{S}^\top\|_\rho|\lambda^{-1}|\right) \\
&\leq \frac{1}{1 - |\lambda|^2}\left(\|\hat{S}\|_\rho|\lambda| + \|\hat{S}^\top\|_\rho\right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|A^\top\|_\rho &\leq \frac{1}{|\lambda^{-1}| - |\lambda|}\max\left\{\|\hat{S}_3^\top\|_\rho + \|\hat{S}_2\|_\rho,\ \|\hat{S}_2^\top\|_\rho|\lambda| + \|\hat{S}_3\|_\rho|\lambda^{-1}| + \|\hat{S}_4\|_\rho\right\} \\
&\leq \frac{1}{|\lambda^{-1}| - |\lambda|}\max\left\{\|\hat{S}_3^\top\|_\rho + \|\hat{S}\|_\rho|\lambda^{-1}|,\ \|\hat{S}_2^\top\|_\rho + \|\hat{S}\|_\rho|\lambda^{-1}|\right\} \\
&\leq \frac{1}{|\lambda^{-1}| - |\lambda|}\left(\|\hat{S}^\top\|_\rho + \|\hat{S}\|_\rho|\lambda^{-1}|\right) \\
&\leq \frac{1}{1 - |\lambda|^2}\left(\|\hat{S}^\top\|_\rho|\lambda| + \|\hat{S}\|_\rho\right).
\end{aligned}$$

Finally we obtain

$$\|A\|_\rho < \frac{1}{1 - (\sigma_\lambda)^2}\left(C_{\hat{S}}\sigma_\lambda + C_{\hat{S}^\top}\right) =: C_A, \tag{3.12}$$
$$\|A^\top\|_\rho < \frac{1}{1 - (\sigma_\lambda)^2}\left(C_{\hat{S}^\top}\sigma_\lambda + C_{\hat{S}}\right) =: C_{A^\top}. \tag{3.13}$$

We can now control N and $N^\top$ as

$$\|N\|_\rho \leq \|\hat{N}\|_\rho + \|L\|_\rho\|A\|_\rho < \sigma_{\hat{N}} + C_L C_A =: C_N, \tag{3.14}$$
$$\|N^\top\|_\rho \leq \|\hat{N}^\top\|_\rho + \|A^\top\|_\rho\|L^\top\|_\rho < \sigma_{\hat{N}^\top} + C_{A^\top}C_{L^\top} =: C_{N^\top}, \tag{3.15}$$

and the frame P as

$$\|P\|_\rho < C_L + C_N =: C_P, \tag{3.16}$$
$$\|P^\top\|_\rho < \max\{C_{L^\top},\ C_{N^\top}\} =: C_{P^\top}. \tag{3.17}$$

3.2.2 Approximate invariance of the subframe L

Let $E_L : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times n}$ denote the error in the invariance of the subframe L. That is,

$$E_L := \mathrm{D}_z\phi_T \circ (K, \mathrm{id})L - L \circ \mathcal{R}_{\omega\alpha}\Lambda. \tag{3.18}$$

Also, let us define

$$\begin{aligned}\Delta X :=& X \circ \Big(\phi_T \circ (K, \mathrm{id}), \mathcal{R}_\alpha\Big) - X \circ (K \circ \mathcal{R}_{\omega\alpha}, \mathcal{R}_\alpha)\\ =& \int_0^1 \mathrm{D}_z X \circ \Big(K \circ \mathcal{R}_{\omega\alpha} + sE_K, \mathcal{R}_\alpha\Big) E_K ds,\end{aligned}$$

where the last equality only holds in $\tilde{\mathcal{U}}$ for sufficiently small E_K. More specifically, suppose we want to obtain upper bounds for $\|\Delta X\|_{\rho-\delta}$. Then, we need that for all $(\theta, \varphi) \in \mathbb{T}^d_{\rho-\delta} \times \mathbb{T}^\ell_{\rho-\delta}$

$$\Big(K(\bar\theta, \bar\varphi) + sE_K(\theta, \varphi), \bar\varphi\Big) \in \tilde{\mathcal{U}}.$$

Equivalently, for all $(\theta, \varphi) \in \mathbb{T}^d_{\rho-\delta} \times \mathbb{T}^\ell_{\rho-\delta}$ and all $s \in [0, 1]$, we need some $\theta^* \in \mathbb{T}^d_{\rho-\delta}$ such that

$$\Big|K(\bar\theta, \bar\varphi) + sE_K(\theta, \varphi) - K_0(\theta^*, \bar\varphi)\Big| < R.$$

Observe that if

$$\frac{\|E_K\|_\rho}{R - \|K - K_0\|_\rho} < 1 \tag{3.19}$$

holds, we can take $\theta^* = \bar\theta$ and, consequently, ΔX is well-defined in $\mathbb{T}^d_{\rho-\delta} \times \mathbb{T}^\ell_{\rho-\delta}$. We will encounter the need of hypothesis (3.19) in order to guarantee that analogous norms for various objects are well defined. Therefore, in what follows, we assume (3.19) holds.

Lemma 3.2. *The error in the invariance of the subframe L is given by*

$$E_L = \Big(\mathrm{D}_\theta E_K \;\Big|\; E_{\mathcal{X}} \;\Big|\; E_W\Big), \tag{3.20}$$

where

$$E_{\mathcal{X}} := \Delta X - \mathrm{D}_\varphi E_K \hat\alpha.$$

Additionally, for any $\delta \in (0, \rho)$

$$\|E_L\|_{\rho-\delta} \le \frac{C^K_{E_L}}{\delta}\|E_K\|_\rho + C^W_{E_L}\|E_W\|_\rho, \tag{3.21}$$

$$\|E_L^\top\|_{\rho-\delta} \le \frac{C^K_{E_L^\top}}{\delta}\|E_K\|_\rho + C^W_{E_L^\top}\|E_W\|_\rho. \tag{3.22}$$

Proof. Let us first differentiate (2.41) with respect to both θ and φ to obtain

$$\mathrm{D}_\theta E_K = \mathrm{D}_z\phi_T \circ (K, \mathrm{id})\mathrm{D}_\theta K - \mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha}, \tag{3.23}$$

$$\mathrm{D}_\varphi E_K = \mathrm{D}_z\phi_T \circ (K, \mathrm{id})\mathrm{D}_\varphi K + \mathrm{D}_\varphi\phi_T \circ (K, \mathrm{id}) - \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha}. \tag{3.24}$$

The first term in (3.20) is immediate from (3.23). Also, the last term in (3.20) is precisely definition (2.42). For the second term, we need to show that

$$\mathrm{D}_z\phi_T \circ (K, \mathrm{id})\mathcal{X} - \mathcal{X} \circ \mathcal{R}_{\omega\alpha} = E_{\mathcal{X}}.$$

Let us use identity (2.19) at $z = K(\theta,\varphi)$ to obtain

$$\mathrm{D}_z\phi_T\Big(K(\theta,\varphi),\varphi\Big)X\Big(K(\theta,\varphi),\varphi\Big) + \mathrm{D}_\varphi\phi_T\Big(K(\theta,\varphi),\varphi\Big)\hat{\alpha} = X\Big(\phi_T\Big(K(\theta,\varphi),\varphi\Big),\mathcal{R}_\alpha(\varphi)\Big).$$

Then, after using (3.24) and some manipulation, we finally obtain

$$\mathrm{D}_z\phi_T\circ(K,\mathrm{id})\mathcal{X} - \mathcal{X}\circ\mathcal{R}_{\omega\alpha} = \Delta X - \mathrm{D}_\varphi E_K\hat{\alpha}.$$

For the estimate (3.21), note that for any $\delta \in (0,\rho)$

$$\|E_{\mathcal{X}}\|_{\rho-\delta} \le \|\Delta X\|_{\rho-\delta} + \|\mathrm{D}_\varphi E_K\|_{\rho-\delta}|\hat{\alpha}|$$

and from assumption (3.19), ΔX is well-defined. Then, we can obtain the estimates

$$\|E_{\mathcal{X}}\|_{\rho-\delta} \le (c_{\mathrm{D}_z X}\delta + \ell|\hat{\alpha}|)\,\frac{\|E_K\|_\rho}{\delta} \le C_{E_{\mathcal{X}}}\frac{\|E_K\|_\rho}{\delta}, \tag{3.25}$$

$$\begin{aligned}\|E_{\mathcal{X}}^\top\|_{\rho-\delta} &\le \|(\Delta X)^\top\|_{\rho-\delta} + \left|\hat{\alpha}^\top\right|\,\|(\mathrm{D}_\varphi E_K)^\top\|_{\rho-\delta} \\ &\le 2n\left(c_{(\mathrm{D}_z X)^\top}\delta + \left|\hat{\alpha}^\top\right|\right)\frac{\|E_K\|_\rho}{\delta} \le C_{E_{\mathcal{X}}^\top}\frac{\|E_K\|_\rho}{\delta}.\end{aligned} \tag{3.26}$$

Lastly,

$$\begin{aligned}\|E_L\|_{\rho-\delta} &\le \|\mathrm{D}_\theta E_K\|_{\rho-\delta} + \|E_{\mathcal{X}}\|_{\rho-\delta} + \|E_W\|_{\rho-\delta} \\ &\le \frac{d}{\delta}\|E_K\|_\rho + \frac{C_{E_{\mathcal{X}}}}{\delta}\|E_K\|_\rho + \|E_W\|_\rho \\ &\le \frac{(d + C_{E_{\mathcal{X}}})}{\delta}\|E_K\|_\rho + \|E_W\|_\rho \\ &\le \frac{C_{E_L}^K}{\delta}\|E_K\|_\rho + C_{E_L}^W\|E_W\|_\rho, \\ \|E_L^\top\|_{\rho-\delta} &\le \max\Big\{\|(\mathrm{D}_\theta E_K)^\top\|_{\rho-\delta},\ \|E_{\mathcal{X}}^\top\|_{\rho-\delta},\ \|{E_W}^\top\|_{\rho-\delta}\Big\} \\ &\le \max\left\{\frac{2n}{\delta}\|E_K\|_\rho,\ \frac{C_{E_{\mathcal{X}}^\top}}{\delta}\|E_K\|_\rho,\ 2n\|E_W\|_\rho\right\} \\ &\le \max\Big\{2n,\ C_{E_{\mathcal{X}}^\top}\Big\}\frac{\|E_K\|_\rho}{\delta} + 2n\|E_W\|_\rho \\ &\le \frac{C_{E_L^\top}^K}{\delta}\|E_K\|_\rho + C_{E_L^\top}^W\|E_W\|_\rho.\end{aligned}$$

□

3.2.3 Approximate isotropy of $\mathcal{K}$

Recall that, for each $\varphi \in \mathbb{T}^d$, we denote the coordinate representation of the pullback of the symplectic form $\boldsymbol{\omega}$ by K as

$$\Omega_{\mathrm{D}K} = (\mathrm{D}_\theta K)^\top\Omega\circ K\mathrm{D}_\theta K.$$

Let us also define

$$\Delta\Omega := \Omega\circ\phi_T\circ(K,\mathrm{id}) - \Omega\circ K\circ\mathcal{R}_{\omega\alpha} \tag{3.27}$$

$$= \int_0^1 \mathrm{D}\Omega \circ (K \circ \mathcal{R}_{\omega\alpha} + sE_K) E_K ds,$$

where, similarly as in Section 3.2.2, the last equality only holds in $\mathcal{U}$ for sufficiently small E_K.

Lemma 3.3. *The approximate isotropy of* $\mathcal{K}$ *is given by* $\Omega_{\mathrm{D}K} = \mathfrak{R}(\mathfrak{L}\Omega_{\mathrm{D}K})$, *where*

$$\begin{aligned} \mathfrak{L}\Omega_{\mathrm{D}K} &= \Omega_{\mathrm{D}K} - \Omega_{\mathrm{D}K} \circ \mathcal{R}_{\omega\alpha} \\ &= (\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha})^\top \Delta\Omega \, \mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha} + (\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha})^\top \Omega \circ \phi_T \circ (K, \mathrm{id}) \, \mathrm{D}_\theta E_K \\ &\quad + (\mathrm{D}_\theta E_K)^\top \Omega \circ \phi_T \circ (K, \mathrm{id}) \, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\theta K. \end{aligned} \tag{3.28}$$

Additionally, for any $\delta \in (0, \rho)$

$$\|\Omega_{\mathrm{D}K}\|_{\rho-\delta} \leq \frac{C_{\Omega_{\mathrm{D}K}}}{\gamma \delta^{\tau+1}} \|E_K\|_\rho. \tag{3.29}$$

Proof. From the symplecticity of ϕ_T and (3.23)

$$\begin{aligned} \Omega_{\mathrm{D}K} =& (\mathrm{D}_\theta K)^\top \Big(\mathrm{D}_z \phi_T \circ (K, \mathrm{id})\Big)^\top \Omega \circ \phi_T \circ (K, \mathrm{id}) \, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\theta K \\ =& \Big(\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha} + \mathrm{D}_\theta E_K\Big)^\top \Omega \circ \phi_T \circ (K, \mathrm{id}) \Big(\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha} + \mathrm{D}_\theta E_K\Big) \\ =& (\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha})^\top \Omega \circ \phi_T \circ (K, \mathrm{id}) \, \mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha} + (\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha})^\top \Omega \circ \phi_T \circ (K, \mathrm{id}) \, \mathrm{D}_\theta E_K \\ &+ (\mathrm{D}_\theta E_K)^\top \Omega \circ \phi_T \circ (K, \mathrm{id}) \, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\theta K. \end{aligned}$$

Then, using the previous expression, the definition of $\mathfrak{L}$, and (3.27) we obtain (3.28). We also need to show that $\langle \Omega_{\mathrm{D}K} \rangle = 0$. Observe that since $\Omega = (\mathrm{D}a)^\top - \mathrm{D}a$

$$\Omega_{\mathrm{D}K} = \Big(\mathrm{D}_\theta (a \circ K)\Big)^\top \mathrm{D}_\theta K - \Big(\mathrm{D}_\theta K\Big)^\top \mathrm{D}_\theta (a \circ K)$$

and each entry ij of $\Omega_{\mathrm{D}K}$ reads

$$\begin{aligned} (\Omega_{\mathrm{D}K})_{ij} =& \sum_k \partial_{\theta^i} \Big(a_k \circ K\Big) \partial_{\theta^j} K^k - \partial_{\theta^j} \Big(a_k \circ K\Big) \partial_{\theta^i} K^k \\ =& \sum_k \partial_{\theta^i} \Big(a_k \circ K \partial_{\theta^j} K^k\Big) - a_k \circ K \partial^2_{\theta^i \theta^j} K^k \\ &- \partial_{\theta^j} \Big(a_k \circ K \partial_{\theta^i} K^k\Big) + a_k \circ K \partial^2_{\theta^i \theta^j} K^k \\ =& \sum_k \partial_{\theta^i} \Big(a_k \circ K \partial_{\theta^j} K^k\Big) - \partial_{\theta^j} \Big(a_k \circ K \partial_{\theta^i} K^k\Big). \end{aligned}$$

Then, $\langle \Omega_{\mathrm{D}K} \rangle = 0$ since we are taking averages of derivatives with respect to θ. Therefore, $\Omega_{\mathrm{D}K} = \mathfrak{R}(\mathfrak{L}\Omega_{\mathrm{D}K})$. For the estimate (3.29), note first that—similarly as in Lemma 3.2—$\Delta\Omega$ is well-defined by assumption (3.19). Then, we can obtain the following estimate from (3.28)

$$\begin{aligned} \|\mathfrak{L}\Omega_{\mathrm{D}K}\|_{\rho-\delta} &\leq \Big(\sigma_{(\mathrm{D}_\theta K)^\top} c_{\mathrm{D}\Omega} \sigma_{\mathrm{D}_\theta K} \delta + \sigma_{(\mathrm{D}_\theta K)^\top} c_\Omega d + 2n c_\Omega c_{\mathrm{D}_z \phi_T} \sigma_{\mathrm{D}_\theta K}\Big) \frac{\|E_K\|_\rho}{\delta} \\ &\leq C_{\mathfrak{L}\Omega_{\mathrm{D}K}} \frac{\|E_K\|_\rho}{\delta}, \end{aligned} \tag{3.30}$$

and, using the Rüssmann estimate (1.15), we obtain

$$\|\Omega_{\mathrm{D}K}\|_{\rho-2\delta} \leq \frac{c_{\mathfrak{R}} C_{\mathfrak{L}\Omega_{\mathrm{D}K}}}{\gamma \delta^{\tau+1}} \|E_K\|_\rho \leq \frac{C_{\Omega_{\mathrm{D}K}}}{\gamma \delta^{\tau+1}} \|E_K\|_\rho.$$

□

3.2.4 Approximate Lagrangianity of the subframe L

Recall that the Lagrangianity of the subframe L is given by Ω_L defined in (2.23). Using the symplecticity of ϕ_T, (3.18), and definition (3.27), we obtain that Ω_L is a solution of the following equation

$$\begin{aligned}\Omega_L - \Lambda^\top \Omega_L \circ \mathcal{R}_{\omega\alpha} \Lambda =& \Lambda^\top L^\top \circ \mathcal{R}_{\omega\alpha} \Delta\Omega\, L \circ \mathcal{R}_{\omega\alpha} \Lambda + \Lambda^\top L^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, E_L \\ &+ E_L^\top\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) L.\end{aligned} \tag{3.31}$$

Lemma 3.4. *The approximate Lagrangianity of the frame L, is given by*

$$\begin{aligned}
\Omega_{\mathrm{D}K} &= \mathfrak{R}\Big(\mathfrak{L}\Omega_{\mathrm{D}K}\Big), & \Omega_{\mathrm{D}K\mathcal{X}} &= \mathfrak{R}\Big(\mathfrak{L}\Omega_{\mathrm{D}K\mathcal{X}}\Big), & \Omega_{\mathrm{D}KW} &= \mathfrak{R}^{1\lambda}\Big(\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW}\Big),\\
\Omega_{\mathcal{X}\mathrm{D}K} &= \mathfrak{R}\Big(\mathfrak{L}\Omega_{\mathcal{X}\mathrm{D}K}\Big), & \Omega_{\mathcal{X}} &= 0, & \Omega_{\mathcal{X}W} &= \mathfrak{R}^{1\lambda}\Big(\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W}\Big),\\
\Omega_{W\mathrm{D}K} &= \mathfrak{R}^{1\lambda}\Big(\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K}\Big), & \Omega_{W\mathcal{X}} &= \mathfrak{R}^{1\lambda}\Big(\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}}\Big), & \Omega_W &= 0,
\end{aligned}$$

where

$$\begin{aligned}
\mathfrak{L}\Omega_{\mathrm{D}K\mathcal{X}} &= (\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha})^\top \Delta\Omega\, \mathcal{X} \circ \mathcal{R}_{\omega\alpha} + (\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha})^\top\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, E_{\mathcal{X}}\\
&\quad + \mathrm{D}_\theta E_K{}^\top\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \mathcal{X},\\
\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW} &= (\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha})^\top \Delta\Omega\, W \circ \mathcal{R}_{\omega\alpha} \lambda + (\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha})^\top\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, E_W\\
&\quad + \mathrm{D}_\theta E_K{}^\top\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) W,\\
\mathfrak{L}\Omega_{\mathcal{X}\mathrm{D}K} &= \mathcal{X}^\top \circ \mathcal{R}_{\omega\alpha} \Delta\Omega\, \mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha} + \mathcal{X}^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, \mathrm{D}_\theta E_K\\
&\quad + E_{\mathcal{X}}^\top\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\theta K,\\
\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W} &= \mathcal{X}^\top \circ \mathcal{R}_{\omega\alpha} \Delta\Omega\, W \circ \mathcal{R}_{\omega\alpha} \lambda + \mathcal{X}^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, E_W\\
&\quad + E_{\mathcal{X}}^\top\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) W,\\
\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K} &= \lambda W^\top \circ \mathcal{R}_{\omega\alpha} \Delta\Omega\, \mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha} + \lambda W^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, \mathrm{D}_\theta E_K\\
&\quad + E_W{}^\top\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\theta K,\\
\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}} &= \lambda W^\top \circ \mathcal{R}_{\omega\alpha} \Delta\Omega\, \mathcal{X} \circ \mathcal{R}_{\omega\alpha} + \lambda W^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, E_{\mathcal{X}}\\
&\quad + E_W{}^\top\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \mathcal{X}.
\end{aligned}$$

Additionally, for any $\delta \in (0, \rho/2)$

$$\|\Omega_L\|_{\rho-2\delta} \leq \frac{C^K_{\Omega_L}}{\gamma \delta^{\tau+1}} \|E_K\|_\rho + C^W_{\Omega_L} \|E_W\|_\rho. \tag{3.32}$$

Proof. Note that $\Omega_{\mathrm{D}K}$ is already controlled by Lemma 3.3. Also, the expressions for the cohomological operator applied to each block of Ω_L follow directly from (3.31). Additionally, $\Omega_{\mathcal{X}}$ is trivially zero and, in the case of $1-$dimensional bundles, Ω_W is also trivially zero. We also need to show that $\Omega_{\mathrm{D}K\mathcal{X}}$ has zero average and, consequently, so does $\Omega_{\mathcal{X}\mathrm{D}K}$. Observe that

$$\Omega_{\mathrm{D}K\mathcal{X}} = (\mathrm{D}_\theta K)^\top \Omega \circ K \Big(X \circ (K, \mathrm{id}) - \mathrm{D}_\varphi K \hat{\alpha}\Big).$$

For the first term, if we use that $\Omega(z) X(z, \varphi) = \mathrm{D}_z H(z, \varphi)^\top$, we obtain

$$\Big\langle (\mathrm{D}_\theta K)^\top \Big(\mathrm{D}_z H \circ (K, \mathrm{id})\Big)^\top \Big\rangle = \Big\langle \Big(\mathrm{D}_\theta \Big(H \circ (K, \mathrm{id})\Big)\Big)^\top \Big\rangle = 0,$$

since we are taking averages of derivatives with respect to θ. The second term, using that $\boldsymbol{\omega}$ is exact, reads

$$\left\langle (\mathrm{D}_\theta K)^\top \mathrm{D}_\varphi (a \circ K) - \Big(\mathrm{D}_\theta (a\circ K)\Big)^\top \mathrm{D}_\varphi K \right\rangle \hat{\alpha}.$$

Then, each entry i of $\langle \Omega_{\mathrm{D}K\mathcal{X}} \rangle$ is obtained from

$$\begin{aligned}
\langle \Omega_{\mathrm{D}K\mathcal{X}} \rangle_i &= \sum_j \left\langle \sum_k \partial_{\theta^i} K^k \partial_{\varphi^j} (a_k \circ K) - \partial_{\theta^i} (a_k \circ K) \partial_{\varphi^j} K^k \right\rangle \hat{\alpha}^j \\
&= \sum_j \left\langle \sum_k \partial_{\varphi^j} \Big(a_k \circ K \partial_{\theta^i} K^k \Big) - a_k \circ K \partial^2_{\theta^i \varphi^j} K^k \right. \\
&\quad \left. - \partial_{\theta^i} \Big(a_k \circ K \partial_{\varphi^j} K^k \Big) + a_k \circ K \partial^2_{\theta^i \varphi^j} K^k \right\rangle \hat{\alpha}^j \\
&= 0
\end{aligned}$$

since we are taking averages of derivatives with respect to θ and φ. For estimate (3.32), note first that

$$\begin{aligned}
\|\Omega_L\|_{\rho-2\delta} \leq \max \Big\{ & \|\Omega_{\mathrm{D}K}\|_{\rho-2\delta} + \|\Omega_{\mathrm{D}K\mathcal{X}}\|_{\rho-2\delta} + \|\Omega_{\mathrm{D}KW}\|_{\rho-\delta}, \\
& \|\Omega_{\mathcal{X}\mathrm{D}K}\|_{\rho-2\delta} + \|\Omega_{\mathcal{X}}\|_{\rho-\delta} + \|\Omega_{\mathcal{X}W}\|_{\rho-\delta}, \\
& \|\Omega_{W\mathrm{D}K}\|_{\rho-\delta} + \|\Omega_{W\mathcal{X}}\|_{\rho-\delta} + \|\Omega_{W}\|_{\rho-\delta} \Big\}
\end{aligned} \tag{3.33}$$

and that $\Delta\Omega$ is well-defined by assumption (3.19). We proceed then by obtaining estimates for each block. Let us first estimate

$$\begin{aligned}
\|\mathfrak{L}\Omega_{\mathrm{D}K\mathcal{X}}\|_{\rho-\delta} &\leq \Big(\sigma_{(\mathrm{D}_\theta K)^\top} c_{\mathrm{D}\Omega} C_{\mathcal{X}} \delta + \sigma_{(\mathrm{D}_\theta K)^\top} c_\Omega C_{E_{\mathcal{X}}} + 2n c_\Omega c_{\mathrm{D}_z \phi_T} C_{\mathcal{X}} \Big) \frac{\|E_K\|_\rho}{\delta} \\
&\leq \frac{C_{\mathfrak{L}\Omega_{\mathrm{D}K\mathcal{X}}}}{\delta} \|E_K\|_\rho,
\end{aligned} \tag{3.34}$$

$$\begin{aligned}
\|\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW}\|_{\rho-\delta} &\leq \Big(\sigma_{(\mathrm{D}_\theta K)^\top} c_{\mathrm{D}\Omega} \sigma_W \sigma_\lambda \delta + 2n c_\Omega c_{\mathrm{D}_z \phi_T} \sigma_W \Big) \frac{\|E_K\|_\rho}{\delta} + \Big(\sigma_{(\mathrm{D}_\theta K)^\top} c_\Omega \Big) \|E_W\|_\rho \\
&\leq \frac{C^K_{\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW}}}{\delta} \|E_K\|_\rho + C^W_{\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW}} \|E_W\|_\rho,
\end{aligned} \tag{3.35}$$

$$\begin{aligned}
\|\mathfrak{L}\Omega_{\mathcal{X}\mathrm{D}K}\|_{\rho-\delta} &\leq \Big(C_{\mathcal{X}^\top} c_{\mathrm{D}\Omega} \sigma_{\mathrm{D}_\theta K} \delta + C_{\mathcal{X}^\top} c_\Omega d + C_{E^\top_{\mathcal{X}}} c_\Omega c_{\mathrm{D}_z \phi_T} \sigma_{\mathrm{D}_\theta K} \Big) \frac{\|E_K\|_\rho}{\delta} \\
&\leq \frac{C_{\mathfrak{L}\Omega_{\mathcal{X}\mathrm{D}K}}}{\delta} \|E_K\|_\rho,
\end{aligned} \tag{3.36}$$

$$\begin{aligned}
\|\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W}\|_{\rho-\delta} &\leq \Big(C_{\mathcal{X}^\top} c_{\mathrm{D}\Omega} \sigma_W \sigma_\lambda \delta + C_{E^\top_{\mathcal{X}}} c_\Omega c_{\mathrm{D}_z \phi_T} \sigma_W \Big) \frac{\|E_K\|_\rho}{\delta} + (C_{\mathcal{X}^\top} c_\Omega) \|E_W\|_\rho \\
&\leq \frac{C^K_{\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W}}}{\delta} \|E_K\|_\rho + C^W_{\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W}} \|E_W\|_\rho,
\end{aligned} \tag{3.37}$$

$$\begin{aligned}
\|\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K}\|_{\rho-\delta} &\leq (\sigma_\lambda \sigma_{W^\top} c_{\mathrm{D}\Omega} \sigma_{\mathrm{D}_\theta K} \delta + \sigma_\lambda \sigma_{W^\top} c_\Omega d) \frac{\|E_K\|_\rho}{\delta} + (2n c_\Omega c_{\mathrm{D}_z \phi_T} \sigma_{\mathrm{D}_\theta K}) \|E_W\|_\rho \\
&\leq \frac{C^K_{\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K}}}{\delta} \|E_K\|_\rho + C^W_{\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K}} \|E_W\|_\rho
\end{aligned} \tag{3.38}$$

$$\begin{aligned}
\|\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}}\|_{\rho-\delta} &\leq (\sigma_\lambda \sigma_{W^\top} c_{\mathrm{D}\Omega} C_{\mathcal{X}} \delta + \sigma_\lambda \sigma_{W^\top} c_\Omega C_{E_{\mathcal{X}}}) \frac{\|E_K\|_\rho}{\delta} + (2n c_\Omega c_{\mathrm{D}_z \phi_T} C_{\mathcal{X}}) \|E_W\|_\rho \\
&\leq \frac{C^K_{\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}}}}{\delta} \|E_K\| + C^W_{\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}}} \|E_W\|_\rho,
\end{aligned} \tag{3.39}$$

where we used Lemma 3.2.

We are now in the position to use Rüssmann estimates (1.15) and (1.13) and obtain

$$\|\Omega_{\mathrm{D}K\mathcal{X}}\|_{\rho-2\delta} \leq \frac{c_{\mathfrak{R}} C_{\mathfrak{L}\Omega_{\mathrm{D}K\mathcal{X}}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho \leq \frac{C_{\Omega_{\mathrm{D}K\mathcal{X}}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho, \tag{3.40}$$

$$\begin{aligned}
\|\Omega_{\mathrm{D}KW}\|_{\rho-\delta} &\leq \frac{1}{1-\sigma_\lambda}\left(\frac{C^K_{\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW}}}{\delta}\|E_K\|_\rho + C^W_{\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW}}\|E_W\|_\rho\right) \qquad (3.41)\\
&\leq \frac{C^K_{\Omega_{\mathrm{D}KW}}}{\delta}\|E_K\|_\rho + C^W_{\Omega_{\mathrm{D}KW}}\|E_W\|_\rho,
\end{aligned}$$

$$\|\Omega_{\mathcal{X}\mathrm{D}K}\|_{\rho-2\delta} \leq \frac{c_{\mathfrak{R}} C_{\mathfrak{L}\Omega_{\mathcal{X}\mathrm{D}K}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho \leq \frac{C_{\Omega_{\mathcal{X}\mathrm{D}K}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho, \tag{3.42}$$

$$\begin{aligned}
\|\Omega_{\mathcal{X}W}\|_{\rho-\delta} &\leq \frac{1}{1-\sigma_\lambda}\left(\frac{C^K_{\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W}}}{\delta}\|E_K\|_\rho + C^W_{\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W}}\|E_W\|_\rho\right) \qquad (3.43)\\
&\leq \frac{C^K_{\Omega_{\mathcal{X}W}}}{\delta}\|E_K\|_\rho + C^W_{\Omega_{\mathcal{X}W}}\|E_W\|_\rho,
\end{aligned}$$

$$\begin{aligned}
\|\Omega_{W\mathrm{D}K}\|_{\rho-\delta} &\leq \frac{1}{1-\sigma_\lambda}\left(\frac{C^K_{\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K}}}{\delta}\|E_K\|_\rho + C^W_{\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K}}\|E_W\|_\rho\right) \qquad (3.44)\\
&\leq \frac{C^K_{\Omega_{W\mathrm{D}K}}}{\delta}\|E_K\|_\rho + C^W_{\Omega_{W\mathrm{D}K}}\|E_W\|_\rho,
\end{aligned}$$

$$\begin{aligned}
\|\Omega_{W\mathcal{X}}\|_{\rho-\delta} &\leq \frac{1}{1-\sigma_\lambda}\left(\frac{C^K_{\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}}}}{\delta}\|E_K\|_\rho + C^W_{\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}}}\|E_W\|_\rho\right) \qquad (3.45)\\
&\leq \frac{C^K_{\Omega_{W\mathcal{X}}}}{\delta}\|E_K\|_\rho + C^W_{\Omega_{W\mathcal{X}}}\|E_W\|_\rho.
\end{aligned}$$

Lastly, from (3.33), we obtain

$$\begin{aligned}
\|\Omega_L\|_{\rho-2\delta} \leq \max\Bigg\{ &\left(C_{\Omega_{\mathrm{D}K}} + C_{\Omega_{\mathrm{D}K\mathcal{X}}} + C^K_{\Omega_{\mathrm{D}KW}}\gamma\delta^\tau\right)\frac{\|E_K\|_\rho}{\gamma\delta^{\tau+1}} + C^W_{\Omega_{\mathrm{D}KW}}\|E_W\|_\rho,\\
&\left(C_{\Omega_{\mathcal{X}\mathrm{D}K}} + C^K_{\Omega_{\mathcal{X}W}}\gamma\delta^\tau\right)\frac{\|E_K\|_\rho}{\gamma\delta^{\tau+1}} + C^W_{\Omega_{\mathcal{X}W}}\|E_W\|_\rho,\\
&\left(C^K_{\Omega_{W\mathrm{D}K}} + C^K_{\Omega_{W\mathcal{X}}}\right)\frac{\|E_K\|_\rho}{\delta} + \left(C^W_{\Omega_{W\mathrm{D}K}} + C^W_{\Omega_{W\mathcal{X}}}\right)\|E_W\|_\rho\Bigg\}\\
\leq \max\Bigg\{ &C_{\Omega_{\mathrm{D}K}} + C_{\Omega_{\mathrm{D}K\mathcal{X}}} + C^K_{\Omega_{\mathrm{D}KW}}\gamma\delta^\tau,\ C_{\Omega_{\mathcal{X}\mathrm{D}K}} + C^K_{\Omega_{\mathcal{X}W}}\gamma\delta^\tau,\\
&\left(C^K_{\Omega_{W\mathrm{D}K}} + C^K_{\Omega_{W\mathcal{X}}}\right)\gamma\delta^\tau\Bigg\}\frac{\|E_K\|_\rho}{\gamma\delta^{\tau+1}}\\
&+ \max\left\{C^W_{\Omega_{\mathrm{D}KW}},\ C^W_{\Omega_{\mathcal{X}W}},\ C^W_{\Omega_{W\mathrm{D}K}} + C^W_{\Omega_{W\mathcal{X}}}\right\}\|E_W\|_\rho\\
\leq &\frac{C^K_{\Omega_L}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho + C^W_{\Omega_L}\|E_W\|_\rho.
\end{aligned}$$

□

3.2.5 Approximate symplecticity of the frame $\hat{P}$

Let $E_{\mathrm{symp}\hat{P}} : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times 2n}$ denote the error in the symplecticity of $\hat{P}$. That is,

$$E_{\mathrm{symp}\hat{P}} := \hat{P}^\top \Omega \circ K \hat{P} - \Omega_0 = \begin{pmatrix} L^\top \Omega \circ K L & L^\top \Omega \circ K \hat{N} + I_n \\ \hat{N}^\top \Omega \circ K L - I_n & \hat{N}^\top \Omega \circ K \hat{N} \end{pmatrix}. \tag{3.46}$$

Lemma 3.5. *The symplecticity error $E_{\mathrm{symp}\hat{P}}$ is given by*

$$E_{\mathrm{symp}\hat{P}} = \left(\begin{array}{c|c} \Omega_L & \\ \hline & B^\top \Omega_L B \end{array}\right). \tag{3.47}$$

Additionally, for any $\delta \in (0, \rho/2)$

$$\|E_{\mathrm{symp}\hat{P}}\|_{\rho-2\delta} \leq \frac{C^K_{E_{\mathrm{symp}\hat{P}}}}{\gamma \delta^{\tau+1}} \|E_K\|_\rho + C^W_{E_{\mathrm{symp}\hat{P}}} \|E_W\|_\rho. \tag{3.48}$$

Proof. Let us compute

$$\begin{aligned} L^\top \Omega \circ K \hat{N} &= L^\top \Omega \circ K \Big(\hat{N} + LA \Big) = -I_n, \\ \hat{N}^\top \Omega \circ K L &= I_n, \\ \hat{N}^\top \Omega \circ K \hat{N} &= B^\top \Omega_L B, \end{aligned}$$

where we used (2.27). From these expressions and the definition of $E_{\mathrm{symp}\hat{P}}$, (3.47) follows directly. In order to obtain estimate (3.48) notice that

$$\|E_{\mathrm{symp}\hat{P}}\|_{\rho-2\delta} \leq \|\Omega_L\|_{\rho-2\delta} \max\Big\{1,\ \|B\|^2_\rho\Big\}.$$

We can now use Lemma 3.4 to obtain

$$\begin{aligned} \|E_{\mathrm{symp}\hat{P}}\|_{\rho-2\delta} &\leq \|\Omega_L\|_{\rho-2\delta} \max\Big\{1,\ (\sigma_B)^2\Big\} \\ &\leq \left(\frac{C^K_{\Omega_L}}{\gamma\delta^{\tau+1}} \|E_K\|_\rho + C^W_{\Omega_L} \|E_W\|_\rho \right) \max\Big\{1,\ (\sigma_B)^2\Big\} \\ &\leq \frac{C^K_{E_{\mathrm{symp}\hat{P}}}}{\gamma\delta^{\tau+1}} \|E_K\|_\rho + C^W_{E_{\mathrm{symp}\hat{P}}} \|E_W\|_\rho. \end{aligned}$$

□

3.2.6 Approximate reducibility of the frame $\hat{P}$

Let us define the error in the reducibility of the linearized dynamics in the coordinates given by the frame $\hat{P}$ as $E_{\mathrm{red}P} : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times 2n}$ given by

$$E_{\mathrm{red}\hat{P}} := -\Omega_0 \hat{P}^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \hat{P} - \left(\begin{array}{c|c} \Lambda & \hat{S} \\ \hline & \Lambda^{-\top} \end{array}\right). \tag{3.49}$$

First, let

$$E_{\mathrm{red}\hat{P}} = \begin{pmatrix} E^{11}_{\mathrm{red}\hat{P}} & E^{12}_{\mathrm{red}\hat{P}} \\ E^{21}_{\mathrm{red}\hat{P}} & E^{22}_{\mathrm{red}\hat{P}} \end{pmatrix}$$

where

$$\begin{aligned}
E^{11}_{\mathrm{red}\hat{P}} &= \hat{N}^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) L - \Lambda, \\
E^{12}_{\mathrm{red}\hat{P}} &= \hat{N}^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \hat{N} - \hat{S}, \\
E^{21}_{\mathrm{red}\hat{P}} &= -L^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) L, \\
E^{22}_{\mathrm{red}\hat{P}} &= -L^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \hat{N} - \Lambda^{-\top}.
\end{aligned}$$

Lemma 3.6. *The reducibility error $E_{\mathrm{red}\hat{P}}$, is given by*

$$\begin{aligned}
E^{11}_{\mathrm{red}\hat{P}} &= \hat{N}^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, E_L, \\
E^{12}_{\mathrm{red}\hat{P}} &= \mathcal{O}_n \\
E^{21}_{\mathrm{red}\hat{P}} &= -\Omega_L \circ \mathcal{R}_{\omega\alpha} \Lambda - L^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, E_L, \\
E^{22}_{\mathrm{red}\hat{P}} &= L^\top \circ \mathcal{R}_{\omega\alpha} \Delta\Omega\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \hat{N} \\
&\quad + \Lambda^{-\top} E_L^\top\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \hat{N}.
\end{aligned}$$

Additionally, for any $\delta \in (0, \rho/2)$

$$\|E_{\mathrm{red}\hat{P}}\|_{\rho-2\delta} \leq \frac{C^K_{E_{\mathrm{red}\hat{P}}}}{\gamma \delta^{\tau+1}} \|E_K\|_\rho + C^W_{E_{\mathrm{red}\hat{P}}} \|E_W\|_\rho. \tag{3.50}$$

Proof. First, recall the definition (2.28) for $\hat{S}$. Then, $E^{12}_{\mathrm{red}\hat{P}}$ is identically zero. For $E^{11}_{\mathrm{red}\hat{P}}$, let us use (3.18) to obtain

$$\begin{aligned}
E^{11}_{\mathrm{red}\hat{P}} &= \hat{N}^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \left(L \circ \mathcal{R}_{\omega\alpha} \Lambda + E_L \right) - \Lambda \\
&= \hat{N}^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, E_L
\end{aligned}$$

Similarly, for $E^{21}_{\mathrm{red}\hat{P}}$ we have

$$\begin{aligned}
E^{21}_{\mathrm{red}\hat{P}} &= -L^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) L \\
&= -L^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \left(L \circ \mathcal{R}_{\omega\alpha} \Lambda + E_L \right) \\
&= -\Omega_L \circ \mathcal{R}_{\omega\alpha} \Lambda - L^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, E_L.
\end{aligned}$$

Lastly, using the symplecticity of ϕ_T, and (3.18) we have

$$\begin{aligned}
E^{22}_{\mathrm{red}\hat{P}} &= -\Lambda^{-\top} L^\top (\mathrm{D}_z \phi_T)^\top \circ (K, \mathrm{id}) \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \hat{N} \\
&\quad + \Lambda^{-\top} E_L^\top \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \hat{N} - \Lambda^{-\top} \\
&= \Lambda^{-\top} L^\top (\mathrm{D}_z \phi_T)^\top \circ (K, \mathrm{id}) \Delta\Omega\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \hat{N} \\
&\quad + \Lambda^{-\top} E_L^\top \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \hat{N} \\
&= L^\top \circ \mathcal{R}_{\omega\alpha} \Delta\Omega\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \hat{N} \\
&\quad + \Lambda^{-\top} E_L^\top\, \Omega \circ \phi_T \circ (K, \mathrm{id})\, \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \hat{N}.
\end{aligned}$$

For estimate (3.50), first note that $\Delta\Omega$ is well-defined by assumption (3.19), that

$$\begin{aligned}\|\Lambda\| &= \max\{1,\ |\lambda|\} \le \max\{1,\ \sigma_\lambda\} = 1,\\ \|\Lambda^{-1}\| &= \max\left\{1,\ |\lambda^{-1}|\right\} \le \max\{1,\ \sigma_{\lambda^{-1}}\} = \sigma_{\lambda^{-1}},\end{aligned}$$

and

$$\|E_{\mathrm{red}\hat{P}}\|_{\rho-2\delta} \le \max\left\{\|E^{11}_{\mathrm{red}\hat{P}}\|_{\rho-2\delta},\ \|E^{21}_{\mathrm{red}\hat{P}}\|_{\rho-2\delta} + \|E^{22}_{\mathrm{red}\hat{P}}\|_{\rho-2\delta}\right\}.$$

We proceed then by computing estimates for each block $E^{ij}_{\mathrm{red}\hat{P}}$ as

$$\begin{aligned}\|E^{11}_{\mathrm{red}\hat{P}}\|_{\rho-\delta} &\le \sigma_{\hat{N}^\top} c_\Omega C^K_{E_L} \frac{\|E_K\|_\rho}{\delta} + \sigma_{\hat{N}^\top} c_\Omega C^W_{E_L}\|E_W\|_\rho \qquad (3.51)\\ &\le \frac{C^K_{E^{11}_{\mathrm{red}\hat{P}}}}{\delta}\|E_K\|_\rho + C^W_{E^{11}_{\mathrm{red}\hat{P}}}\|E_W\|_\rho,\\ \|E^{21}_{\mathrm{red}\hat{P}}\|_{\rho-2\delta} &\le \left(C^K_{\Omega_L} + C_{L^\top} c_\Omega C^K_{E_L}\gamma\delta^\tau\right)\frac{\|E_K\|_\rho}{\gamma\delta^{\tau+1}} + \left(C^W_{\Omega_L} + C_{L^\top} c_\Omega C^W_{E_L}\right)\|E_W\|_\rho \qquad (3.52)\\ &\le \frac{C^K_{E^{21}_{\mathrm{red}\hat{P}}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho + C^W_{E^{21}_{\mathrm{red}\hat{P}}}\|E_W\|_\rho,\\ \|E^{22}_{\mathrm{red}\hat{P}}\|_{\rho-\delta} &\le \left(C_{L^\top} c_{\mathrm{D}\Omega} c_{\mathrm{D}_z\phi_T}\sigma_{\hat{N}}\delta + \sigma_{\lambda^{-1}} C^K_{E_L^\top} c_\Omega c_{\mathrm{D}_z\phi_T}\sigma_{\hat{N}}\right)\frac{\|E_K\|_\rho}{\delta} \qquad (3.53)\\ &\quad + \sigma_{\lambda^{-1}} C^W_{E_L^\top} c_\Omega c_{\mathrm{D}_z\phi_T}\sigma_{\hat{N}}\|E_W\|_\rho\\ &\le \frac{C^K_{E^{22}_{\mathrm{red}\hat{P}}}}{\delta}\|E_K\|_\rho + C^W_{E^{22}_{\mathrm{red}\hat{P}}}\|E_W\|_\rho,\end{aligned}$$

where we used Lemmas 3.2 and 3.4. Finally,

$$\begin{aligned}\|E_{\mathrm{red}\hat{P}}\|_{\rho-2\delta} &\le \max\left\{C^K_{E^{11}_{\mathrm{red}\hat{P}}}\frac{\|E_K\|_\rho}{\delta} + C^W_{E^{11}_{\mathrm{red}\hat{P}}}\|E_W\|_\rho,\ \left(C^K_{E^{21}_{\mathrm{red}\hat{P}}} + C^K_{E^{22}_{\mathrm{red}\hat{P}}}\gamma\delta^\tau\right)\frac{\|E_K\|_\rho}{\gamma\delta^{\tau+1}}\right.\\ &\qquad \left. + \left(C^W_{E^{21}_{\mathrm{red}\hat{P}}} + C^W_{E^{22}_{\mathrm{red}\hat{P}}}\right)\|E_W\|_\rho\right\}\\ &\le \max\left\{C^K_{E^{11}_{\mathrm{red}\hat{P}}}\gamma\delta^\tau,\ C^K_{E^{21}_{\mathrm{red}\hat{P}}} + C^K_{E^{22}_{\mathrm{red}\hat{P}}}\gamma\delta^\tau\right\}\frac{\|E_K\|_\rho}{\gamma\delta^{\tau+1}}\\ &\qquad + \max\left\{C^W_{E^{11}_{\mathrm{red}\hat{P}}},\ C^W_{E^{21}_{\mathrm{red}\hat{P}}} + C^W_{E^{22}_{\mathrm{red}\hat{P}}}\right\}\|E_W\|_\rho\\ &\le \frac{C^K_{E_{\mathrm{red}\hat{P}}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho + C^W_{E_{\mathrm{red}\hat{P}}}\|E_W\|_\rho.\end{aligned}$$

□

3.2.7 Approximate invertibility of $\hat{P}$

Let $E_{\mathrm{inv}\hat{P}} : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times 2n}$ be the error in the invertibility of the frame $\hat{P}$ under the assumption that $\hat{P}$ is symplectic. That is,

$$E_{\mathrm{inv}\hat{P}} := \hat{P}\left(-\Omega_0\hat{P}^\top\Omega\circ K\right) - I_{2n}. \qquad (3.54)$$

In what follows, we will need certain matrix to be sufficiently close to the identity. In particular, we will need $I_{2n} - \Omega_0 E_{\mathrm{symp}\hat{P}}$ to be invertible in $\mathbb{T}^{d+\ell}_{\rho-2\delta}$. This can be ensured if

$$\|E_{\mathrm{symp}\hat{P}}\|_{\rho-2\delta} \le \hat{\nu} < 1. \qquad (3.55)$$

For instance, if (3.55) holds, then $\hat{P}^{-1}$ is invertible with inverse

$$\hat{P}^{-1} = -\left(I_{2n} - \Omega_0 E_{\mathrm{symp}\hat{P}}\right)^{-1} \Omega_0 \hat{P}^{\top}\Omega \circ K.$$

In the remaining of this section, we will assume (3.55) holds.

Lemma 3.7. *The error in the invertibility of $\hat{P}$ is given by*

$$E_{\mathrm{inv}\hat{P}} = \hat{P}\Omega_0 E_{\mathrm{symp}\hat{P}} \left(I - \Omega_0 E_{\mathrm{symp}\hat{P}}\right)^{-1} \Omega_0 \hat{P}^{\top}\Omega \circ K. \tag{3.56}$$

Additionally, for any $\delta \in (0, \rho/2)$

$$\|E_{\mathrm{inv}\hat{P}}\|_{\rho-2\delta} \leq \frac{C^K_{E_{\mathrm{inv}\hat{P}}}}{\gamma\delta^{\tau+1}} \|E_K\|_{\rho} + C^W_{E_{\mathrm{inv}\hat{P}}} \|E_W\|_{\rho}. \tag{3.57}$$

Proof. Note that

$$-\Omega_0 \hat{P}^{\top}\Omega \circ K \hat{P} = I - \Omega_0 E_{\mathrm{symp}\hat{P}}.$$

Therefore, we obtain

$$\begin{aligned}\hat{P}\left(-\Omega_0 \hat{P}^{\top}\Omega \circ K\right) &= \hat{P}\left(I - \Omega_0 E_{\mathrm{symp}\hat{P}}\right)\hat{P}^{-1} \\ &= I + \hat{P}\Omega_0 E_{\mathrm{symp}\hat{P}} \left(I - \Omega_0 E_{\mathrm{symp}\hat{P}}\right)^{-1} \Omega_0 \hat{P}^{\top}\Omega \circ K\end{aligned}$$

and

$$E_{\mathrm{inv}\hat{P}} = \hat{P}\Omega_0 E_{\mathrm{symp}\hat{P}} \left(I - \Omega_0 E_{\mathrm{symp}\hat{P}}\right)^{-1} \Omega_0 \hat{P}^{\top}\Omega \circ K. \tag{3.58}$$

For estimate (3.57), since $\|E_{\mathrm{symp}\hat{P}}\|_{\rho-2\delta} \leq \hat{\nu} < 1$ by assumption,

$$\left\|\left(I_{2n} - \Omega_0 E_{\mathrm{symp}\hat{P}}\right)^{-1}\right\|_{\rho-2\delta} \leq \sum_{k=0}^{\infty} \left\|E_{\mathrm{symp}\hat{P}}\right\|^k_{\rho-2\delta} \leq \frac{1}{1 - \|E_{\mathrm{symp}\hat{P}}\|_{\rho-2\delta}} \leq \frac{1}{1-\hat{\nu}}.$$

Then, from (3.58) and Lemma 3.5, we obtain

$$\begin{aligned}\|E_{\mathrm{inv}\hat{P}}\|_{\rho-2\delta} &\leq \frac{1}{1-\hat{\nu}} C_P C_{P^{\top}} c_{\Omega} \|E_{\mathrm{symp}\hat{P}}\|_{\rho-2\delta} \\ &\leq \frac{C^K_{E_{\mathrm{inv}\hat{P}}}}{\gamma\delta^{\tau+1}} \|E_K\|_{\rho} + C^W_{E_{\mathrm{inv}\hat{P}}} \|E_W\|_{\rho}.\end{aligned}$$

□

3.2.8 Approximate symmetries of $\Lambda^{-1}\hat{S}$ and A

The use of the matrix-valued map A to reduce the torsion is equivalent to applying the following transformation to the frame $\hat{P}$

$$P = \hat{P} \left(\begin{array}{c|c} I_n & A \\ \hline & I_n \end{array}\right). \tag{3.59}$$

This transformation is symplectic—and consequently so is P—if A is symmetric. Recall that according to remark 2.3, A is symmetric if the symmetry condition (2.39) holds, which only holds approximately when K and W satisfy (2.3) and (2.6) approximately. In order to control the approximate geometric properties of the frame P, we need to control the error in the symmetry of A which in turn is controlled by the error in the symmetry condition (2.39). Let us define the following errors in the symmetry of $\Lambda^{-1}\hat{S}$ and A $E_{\mathrm{sym}\Lambda^{-1}\hat{S}}, E_{\mathrm{sym}A} : \mathbb{T}^{d+\ell} \to \mathbb{R}^{n\times n}$ be defined as

$$\begin{aligned} E_{\mathrm{sym}\Lambda^{-1}\hat{S}} &= \Lambda^{-1}\hat{S} - \left(\Lambda^{-1}\hat{S}\right)^{\top}, \\ E_{\mathrm{sym}A} &= A - A^{\top}. \end{aligned}$$

Lemma 3.8. *For any $\delta \in (0, \rho/2)$ the torsion $\hat{S}$ satisfies*

$$\|E_{\mathrm{sym}\Lambda^{-1}\hat{S}}\|_{\rho-2\delta} \leq \frac{C^K_{\Lambda^{-1}\hat{S}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho + C^W_{\Lambda^{-1}\hat{S}}\|E_W\|_\rho. \tag{3.60}$$

Proof. Let us inspect the symplecticity of the reduced dynamics under the frame $\hat{P}$ with respect to Ω_0, i.e., the symplecticity of the following matrix-valued map

$$\Lambda_{\hat{P}} := \left(\begin{array}{c|c} \Lambda & \hat{S} \\ \hline & \Lambda^{-\top} \end{array}\right).$$

A straightforward computation reveals

$$\begin{aligned} E_{\mathrm{symp}\Lambda_{\hat{P}}} :&= \Lambda_{\hat{P}}^{\top}\Omega_0\Lambda_{\hat{P}} - \Omega_0 \\ &= \left(\begin{array}{c|c} & \\ \hline & \Lambda^{-1}\hat{S} - \left(\Lambda^{-1}\hat{S}\right)^{\top} \end{array}\right). \end{aligned} \tag{3.61}$$

Then, from (3.49) and (3.54) we obtain

$$\begin{aligned} E_{\mathrm{symp}\Lambda_{\hat{P}}} =& -\hat{P}^{\top}(\mathrm{D}_z\phi_T)^{\top}\circ(K,\mathrm{id})\Omega\circ K\circ\mathcal{R}_{\omega\alpha}\hat{P}\circ\mathcal{R}_{\omega\alpha}\Omega_0\hat{P}^{\top}\circ\mathcal{R}_{\omega\alpha}\Omega\circ K\circ\mathcal{R}_{\omega\alpha}\mathrm{D}_z\phi_T\circ(K,\mathrm{id})\hat{P} \\ &- \hat{P}^{\top}(\mathrm{D}_z\phi_T)^{\top}\circ(K,\mathrm{id})\Omega\circ K\circ\mathcal{R}_{\omega\alpha}\hat{P}\circ\mathcal{R}_{\omega\alpha}E_{\mathrm{red}\hat{P}} - E_{\mathrm{red}\hat{P}}^{\top}\Omega_0\Lambda_{\hat{P}} - \Omega_0 \\ =& \hat{P}^{\top}(\mathrm{D}_z\phi_T)^{\top}\circ(K,\mathrm{id})\Big(\Omega\circ\phi_T\circ(K,\mathrm{id}) - \Delta\Omega\Big)\Big(I + E_{\mathrm{inv}\hat{P}}\circ\mathcal{R}_{\omega\alpha}\Big)\mathrm{D}_z\phi_T\circ(K,\mathrm{id})\hat{P} \\ &- \hat{P}^{\top}(\mathrm{D}_z\phi_T)^{\top}\circ(K,\mathrm{id})\Omega\circ K\circ\mathcal{R}_{\omega\alpha}\hat{P}\circ\mathcal{R}_{\omega\alpha}E_{\mathrm{red}\hat{P}} - E_{\mathrm{red}\hat{P}}^{\top}\Omega_0\Lambda_{\hat{P}} - \Omega_0, \end{aligned}$$

and using the symplecticity of ϕ_T and (3.46), we have

$$\begin{aligned} E_{\mathrm{symp}\Lambda_{\hat{P}}} =& E_{\mathrm{symp}\hat{P}} - E_{\mathrm{red}\hat{P}}^{\top}\Omega_0\Lambda_{\hat{P}} - \hat{P}^{\top}(\mathrm{D}_z\phi_T)^{\top}\circ(K,\mathrm{id})\Delta\Omega\,\mathrm{D}_z\phi_T\circ(K,\mathrm{id})\hat{P} \\ &+ \hat{P}^{\top}(\mathrm{D}_z\phi_T)^{\top}\circ(K,\mathrm{id})\Omega\circ K\circ\mathcal{R}_{\omega\alpha}E_{\mathrm{inv}\hat{P}}\circ\mathcal{R}_{\omega\alpha}\mathrm{D}_z\phi_T\circ(K,\mathrm{id})\hat{P} \\ &- \hat{P}^{\top}(\mathrm{D}_z\phi_T)^{\top}\circ(K,\mathrm{id})\Omega\circ K\circ\mathcal{R}_{\omega\alpha}\hat{P}\circ\mathcal{R}_{\omega\alpha}E_{\mathrm{red}\hat{P}}. \end{aligned}$$

Lastly, from (3.61) and lemmas 3.5, 3.6 and 3.7, we obtain

$$\begin{aligned} E_{\mathrm{sym}\Lambda^{-1}\hat{S}} =& B^{\top}\Omega_L B - E_{\mathrm{red}\hat{P}}^{22^{\top}}\hat{S} - \hat{N}^{\top}(\mathrm{D}_z\phi_T)^{\top}\circ(K,\mathrm{id})\Delta\Omega\,\mathrm{D}_z\phi_T\circ(K,\mathrm{id})\hat{N} \\ &+ \hat{N}^{\top}(\mathrm{D}_z\phi_T)^{\top}\circ(K,\mathrm{id})\Omega\circ K\circ\mathcal{R}_{\omega\alpha}E_{\mathrm{inv}\hat{P}}\circ\mathcal{R}_{\omega\alpha}\mathrm{D}_z\phi_T\circ(K,\mathrm{id})\hat{N} \\ &- \hat{N}^{\top}(\mathrm{D}_z\phi_T)^{\top}\circ(K,\mathrm{id})\Omega\circ K\circ\mathcal{R}_{\omega\alpha}\hat{P}E^{22}_{\mathrm{red}P} \end{aligned} \tag{3.62}$$

from where estimate (3.60) follows. □

Lemma 3.9. *Let* $A : \mathbb{T}^{d+\ell} \to \mathbb{R}^{n\times n}$ *satisfy* (2.34)-(2.36)*, where*

$$A = \begin{pmatrix} \mathcal{O}_{n-1} & A_2 \\ A_3 & A_4 \end{pmatrix}.$$

Then, for any $\delta \in (0, \rho/2)$*,* A *satisfies*

$$\|E_{\mathrm{sym}A}\|_{\rho-2\delta} \leq \frac{C^K_{E_{\mathrm{sym}A}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho + C^W_{E_{\mathrm{sym}A}}\|E_W\|_\rho. \tag{3.63}$$

Proof. First, notice that

$$A - A^\top = \left(\begin{array}{c|c} & A_2 - A_3^\top \\ \hline A_3 - A_2^\top & \end{array}\right).$$

If we multiply (2.34) by λ and subtract the transpose of (2.35), we obtain

$$\mathfrak{L}^{-\lambda 1}\left(A_2 - A_3^\top\right) = \hat{S}_2\lambda - \hat{S}_3^\top.$$

Similarly, we can multiply the transpose of (2.34) by λ and subtract it to (2.35) to obtain

$$\mathfrak{L}^{-\lambda 1}\left(A_3 - A_2^\top\right) = \hat{S}_3 - \lambda\hat{S}_2^\top.$$

Therefore

$$\begin{aligned} \|A_2 - A_3^\top\|_{\rho-2\delta} &\leq \frac{1}{1-|\lambda|}\|\hat{S}_2\lambda - \hat{S}_3^\top\|_{\rho-2\delta} \leq \frac{1}{1-|\lambda|}\|\hat{S}\Lambda^\top - \Lambda\hat{S}^\top\|_{\rho-2\delta}, \\ \|A_3 - A_2^\top\|_{\rho-2\delta} &\leq \frac{1}{1-|\lambda|}\|\hat{S}_3 - \lambda\hat{S}_2^\top\|_{\rho-2\delta} \leq \frac{1}{1-|\lambda|}\|\hat{S}\Lambda^\top - \Lambda\hat{S}^\top\|_{\rho-2\delta}, \end{aligned}$$

from where we obtain the estimate

$$\|E_{\mathrm{sym}A}\|_{\rho-2\delta} \leq \frac{1}{1-|\lambda|}\|\hat{S}\Lambda^\top - \Lambda\hat{S}^\top\|_{\rho-2\delta} \leq \frac{1}{1-\sigma_\lambda}\|E_{\mathrm{sym}\Lambda^{-1}\hat{S}}\|_{\rho-2\delta},$$

where we used that $\|\Lambda\|$=1. We can now use Lemma 3.8 and obtain (3.63).

□

3.2.9 Approximate symplecticity of the frame P

Let $E_{\mathrm{symp}P} : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times 2n}$ denote the error in the symplecticity of P. That is,

$$E_{\mathrm{symp}P} := P^\top\Omega\circ KP - \Omega_0. \tag{3.64}$$

Lemma 3.10. *For any* $\delta \in (0, \rho/2)$ *the symplecticity error* $E_{\mathrm{symp}P}$ *satisfies*

$$\|E_{\mathrm{symp}P}\|_{\rho-2\delta} \leq \frac{C^K_{E_{\mathrm{symp}P}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho + C^W_{E_{\mathrm{symp}P}}\|E_W\|_\rho. \tag{3.65}$$

Proof. From definition (3.64) and transformation (3.59) we obtain

$$E_{\mathrm{symp}P} = \left(\begin{array}{c|c} I_n & \\ \hline A^\top & I_n \end{array}\right) E_{\mathrm{symp}\hat{P}} \left(\begin{array}{c|c} I_n & A \\ \hline & I_n \end{array}\right) + \left(\begin{array}{c|c} & \\ \hline & A - A^\top \end{array}\right).$$

Consequently,

$$\begin{aligned} \|E_{\mathrm{symp}P}\|_{\rho-2\delta} &\leq \left(1 + \|A^\top\|_\rho\right)(1 + \|A\|_\rho)\, \|E_{\mathrm{symp}\hat{P}}\|_{\rho-2\delta} + \|E_{\mathrm{sym}A}\|_{\rho-2\delta} \\ &\leq (1 + C_{A^\top})(1 + C_A)\, \|E_{\mathrm{symp}\hat{P}}\|_{\rho-2\delta} + \|E_{\mathrm{sym}A}\|_{\rho-2\delta}. \end{aligned}$$

Then, using Lemmas 3.5 and 3.9, result (3.65) is a direct computation. □

3.2.10 Approximate reducibility of the frame P

Let us now define the error in the reducibility of the linearized dynamics in the coordinates given by the frame P as $E_{\mathrm{red}P} : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times 2n}$ given by

$$E_{\mathrm{red}P} := -\Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\, \mathrm{D}_z\phi_T \circ (K, \mathrm{id}) P - \left(\begin{array}{c|c} \Lambda & S \\ \hline & \Lambda^{-\top} \end{array}\right). \tag{3.66}$$

Lemma 3.11. *For any $\delta \in (0, \rho/2)$ the reducibility error $E_{\mathrm{red}P}$ satisfies*

$$\|E_{\mathrm{red}P}\|_{\rho-2\delta} \leq \frac{C^K_{E_{\mathrm{red}P}}}{\gamma\delta^{\tau+1}} \|E_K\|_\rho + C^W_{E_{\mathrm{red}P}} \|E_W\|_\rho. \tag{3.67}$$

Proof. As in the proof of Lemma 3.10, we use definition (3.66) and transformation (3.59) to obtain

$$E_{\mathrm{red}P} = \left(\begin{array}{c|c} I_n & -A^\top \circ \mathcal{R}_{\omega\alpha} \\ \hline & I_n \end{array}\right) E_{\mathrm{red}\hat{P}} \left(\begin{array}{c|c} I_n & A \\ \hline & I_n \end{array}\right),$$

where we used that A solves (2.29). Then,

$$\begin{aligned} \|E_{\mathrm{red}P}\|_{\rho-2\delta} &\leq \left(1 + \|A^\top\|_\rho\right)(1 + \|A\|_{\rho-2\delta})\, \|E_{\mathrm{red}\hat{P}}\|_{\rho-2\delta} \\ &\quad (1 + C_{A^\top})(1 + C_A)\, \|E_{\mathrm{red}\hat{P}}\|_{\rho-2\delta}, \end{aligned}$$

from where we obtain (3.67). □

3.3 Torus and bundle corrections

In this section, we provide two lemmas to control the error in the torus and bundle parameterizations after one step of the iterative scheme. The lemmas rely on the bounds for the approximate geometric properties obtained in Section 3.2.

3.3.1 The linearized equation for the torus

Let us now inspect the new error in the invariance equation after the iterative step of Section 2.5.1, i.e., after choosing ξ_K that solves (2.46)-(2.49). A straightforward calculation reveals

$$\begin{aligned} E_{\bar{K}} :=& \phi_T \circ (K + \Delta K, \mathrm{id}) - K \circ \mathcal{R}_{\omega\alpha} - \Delta K \circ \mathcal{R}_{\omega\alpha} \\ =& E_K + \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \Delta K - \Delta K \circ \mathcal{R}_{\omega\alpha} + \Delta^2 \phi_T, \end{aligned} \tag{3.68}$$

where

$$\Delta^2 \phi_T := \int_0^1 (1-s) \mathrm{D}_z^2 \phi_T \circ (K + s\Delta K, \mathrm{id}) \, [\Delta K, \Delta K] \, ds$$

is only well-defined in $\tilde{\mathcal{U}}$ for sufficiently small ΔK. Similarly as in Section 3.2.2, can guarantee so if

$$\frac{\|\Delta K\|_{\rho-2\delta}}{R - \|K - K_0\|_\rho} < 1. \tag{3.69}$$

Also, we will need $I_{2n} - \Omega_0 E_{\mathrm{symp}P}$ to be invertible in $\mathbb{T}^{d+\ell}_{\rho-2\delta}$. This condition is granted if

$$\|E_{\mathrm{symp}P}\|_{\rho-2\delta} \leq \nu < 1 \tag{3.70}$$

holds—recall the analogous assumption (3.55) in Section 3.2.7. For the remaining of this section, we incorporate (3.69) and (3.70) to our assumptions.

Lemma 3.12. *The new error in the torus invariance equation after correcting K by solving* (2.44) *is given by*

$$E_{\bar{K}} = P \circ \mathcal{R}_{\omega\alpha} \Big(I_{2n} - \Omega_0 E_{\mathrm{symp}P} \circ \mathcal{R}_{\omega\alpha} \Big)^{-1} E^K_{\mathrm{lin}} + \Delta^2 \phi_T, \tag{3.71}$$

where

$$E^K_{\mathrm{lin}} := E_{\mathrm{red}P} \xi_K + \Omega_0 E_{\mathrm{symp}P} \circ \mathcal{R}_{\omega\alpha} \xi_K \circ \mathcal{R}_{\omega\alpha} - \begin{pmatrix} 0_{n\times 1} \\ \langle \eta^3_K \rangle \\ 0_{1\times 1} \end{pmatrix}.$$

Additionally, for any $\delta \in (0, \rho/2)$

$$\|E_{\bar{K}}\|_{\rho-2\delta} \leq \frac{C^{KK}_{E_{\bar{K}}}}{\gamma^4 \delta^{4\tau}} \|E_K\|^2_\rho + \frac{C^{KW}_{E_{\bar{K}}}}{\gamma^2 \delta^{2\tau}} \|E_K\|_\rho \|E_W\|_\rho. \tag{3.72}$$

Proof. Let us begin by inspecting the sum

$$E_K + \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \Delta K - \Delta K \circ \mathcal{R}_{\omega\alpha}.$$

If we left-multiply it by $-\Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha} \, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}$ and by its inverse, we obtain

$$\begin{aligned} E_K + \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \Delta K - \Delta K \circ \mathcal{R}_{\omega\alpha} =& \Big(-\Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha} \, \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \Big)^{-1} \Bigg(- \eta_K \\ & + \Lambda_P \xi_K + E_{\mathrm{red}P} \xi_K - \Big(I_{2n} - \Omega_0 E_{\mathrm{symp}P} \circ \mathcal{R}_{\omega\alpha} \Big) \xi_K \circ \mathcal{R}_{\omega\alpha} \Bigg) \end{aligned}$$

$$= \left(-\Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\right)^{-1} \left(E_{\mathrm{red}P} \xi_K + \Omega_0 E_{\mathrm{symp}P} \circ \mathcal{R}_{\omega\alpha}\, \xi_K \circ \mathcal{R}_{\omega\alpha} - \begin{pmatrix} 0_{n\times 1} \\ \langle \eta_K^3 \rangle \\ 0_{1\times 1} \end{pmatrix} \right)$$
$$= \left(-\Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\right)^{-1} E^K_{\mathrm{lin}},$$

where we used definitions (3.64), (3.66), and that ξ_K satisfies (2.46)-(2.49). Recall that by assumption $\|E_{\mathrm{symp}P}\|_{\rho-2\delta} \le \nu < 1$ so $I_{2n} - \Omega_0 E_{\mathrm{symp}P} \circ \mathcal{R}_{\omega\alpha}$ is invertible. Then, from the previous expression, that

$$\left(- \Omega_0 P^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha}\right)^{-1} = P \circ \mathcal{R}_{\omega\alpha} \left(I_{2n} - \Omega_0 E_{\mathrm{symp}P} \circ \mathcal{R}_{\omega\alpha}\right)^{-1}, \tag{3.73}$$

and (3.68), we obtain (3.71). For estimate (3.72), we first obtain estimates for each η_K^i as

$$\|\eta_K^1\|_\rho \le C_{N^\top} c_\Omega \|E_K\|_\rho \le C_{\eta_K^1} \|E_K\|_\rho, \tag{3.74}$$
$$\|\eta_K^2\|_\rho \le C_{N^\top} c_\Omega \|E_K\|_\rho \le C_{\eta_K^2} \|E_K\|_\rho, \tag{3.75}$$
$$\|\eta_K^3\|_\rho \le \max\left\{\sigma_{(\mathrm{D}_\theta K)^\top},\ C_{\mathcal{X}^\top}\right\} c_\Omega \|E_K\|_\rho \le C_{\eta_K^3} \|E_K\|_\rho, \tag{3.76}$$
$$\|\eta_K^4\|_\rho \le \sigma_{W^\top} c_\Omega \|E_K\|_\rho \le C_{\eta_K^4} \|E_K\|_\rho, \tag{3.77}$$

and for each ξ_K^i we obtain the following estimates from (2.50)-(2.54)

$$|\left\langle \xi_K^3 \right\rangle| \le \sigma_{\langle \mathcal{S}\rangle^{-1}} \left(C_{\eta_K^1} \gamma \delta^\tau + C_{\hat{S}} c_{\mathfrak{R}} C_{\eta_K^3} \right) \frac{\|E_K\|_\rho}{\gamma \delta^\tau} \le \frac{C_{\langle \xi_K^3 \rangle}}{\gamma\delta^\tau} \|E_K\|_\rho, \tag{3.78}$$
$$\|\xi_K^1\|_{\rho-2\delta} \le c_{\mathfrak{R}} \left(C_{\eta_K^1} \gamma\delta^\tau + C_{\hat{S}} \left(c_{\mathfrak{R}} C_{\eta_K^3} + C_{\langle \xi_K^3\rangle} \right)\right) \frac{\|E_K\|_\rho}{\gamma^2 \delta^{2\tau}} \le \frac{C_{\xi_K^1}}{\gamma^2\delta^{2\tau}} \|E_K\|_\rho, \tag{3.79}$$
$$\|\xi_K^2\|_\rho \le \frac{1}{1-\sigma_\lambda} C_{\eta_K^2} \|E_K\|_\rho \le C_{\xi_K^2} \|E_K\|_\rho, \tag{3.80}$$
$$\|\xi_K^3\|_{\rho-\delta} \le \left(c_{\mathfrak{R}} C_{\eta_K^3} + C_{\langle \xi_K^3\rangle}\right) \frac{\|E_K\|_\rho}{\gamma\delta^\tau} \le \frac{C_{\xi_K^3}}{\gamma\delta^\tau} \|E_K\|_\rho, \tag{3.81}$$
$$\|\xi_K^4\|_\rho \le \frac{\sigma_\lambda}{1-\sigma_\lambda} C_{\eta_K^4} \|E_K\|_\rho \le C_{\xi_K^4} \|E_K\|_\rho. \tag{3.82}$$

Therefore,

$$\|\xi_K\|_{\rho-2\delta} \le \max\left\{ C_{\xi_K^1},\ C_{\xi_K^2} \gamma^2\delta^{2\tau},\ C_{\xi_K^3}\gamma\delta^\tau,\ C_{\xi_K^4}\gamma^2\delta^{2\tau}\right\} \frac{\|E_K\|_\rho}{\gamma^2\delta^{2\tau}} \le \frac{C_{\xi_K}}{\gamma^2\delta^{2\tau}} \|E_K\|_\rho, \tag{3.83}$$

and, for $\Delta K = P\xi_K$, we have the estimate

$$\begin{aligned} \|\Delta K\|_{\rho-2\delta} &\le \left(C_L \max\left\{ C_{\xi_K^1},\ C_{\xi_K^2}\gamma^2\delta^{2\tau}\right\} + C_N \max\left\{ C_{\xi_K^3}\gamma\delta^\tau,\ C_{\xi_K^4}\gamma^2\delta^{2\tau}\right\}\right) \frac{\|E_K\|_\rho}{\gamma^2\delta^{2\tau}} \\ &\le \frac{C_{\Delta K}}{\gamma^2\delta^{2\tau}} \|E_K\|_\rho. \end{aligned} \tag{3.84}$$

In order to obtain estimates for $|\langle \eta_K^3\rangle|$, let

$$\eta_K^3 = \begin{pmatrix} (\mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha})^\top \Omega \circ K \circ \mathcal{R}_{\omega\alpha} E_K \\ \mathcal{X}^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha} E_K \end{pmatrix} =: \begin{pmatrix} \eta_K^{31} \\ \eta_K^{32} \end{pmatrix}.$$

We now make use of the following lemma from [FMHM24b],

Lemma 3.13. *The averages of* η_K^{31} *and* η_K^{32} *are*

$$\begin{aligned}\langle\eta_K^{31}\rangle &= \langle(\mathrm{D}_\theta E_K)^\top\Delta^1 a + (\mathrm{D}_\theta K\circ\mathcal{R}_{\omega\alpha})^\top\Delta^2 a\rangle,\\ \langle\eta_K^{32}\rangle &= \langle\Delta^2 H\rangle - \langle(\Delta^1 a)^\top\mathrm{D}_\varphi E_K\hat{\alpha}\rangle - \langle(\Delta^2 a)^\top\mathrm{D}_\varphi K\circ\mathcal{R}_{\omega\alpha}\hat{\alpha}\rangle,\end{aligned}$$

where the Taylor remainders Δ^i *of order* i *in* E^K *are given by*

$$\begin{aligned}\Delta^1 a :=& a\circ\phi_T\circ(K,\mathrm{id}) - a\circ K\circ\mathcal{R}_{\omega\alpha}\\ =& \int_0^1 \mathrm{D}a\circ\Big(K\circ\mathcal{R}_{\omega\alpha} + sE_K\Big)E_K\ ds,\\ \Delta^2 a :=& a\circ\phi_T\circ(K,\mathrm{id}) - a\circ K\circ\mathcal{R}_{\omega\alpha} - \mathrm{D}a\circ K\circ\mathcal{R}_{\omega\alpha}E_K\\ =& \int_0^1 (1-s)\mathrm{D}^2 a\circ\Big(K\circ\mathcal{R}_{\omega\alpha} + sE_K\Big)[E_K,E_K]\ ds,\\ \Delta^2 H :=& H\circ\Big(\phi_T\circ(K,\mathrm{id}),R_\alpha\Big) - H\circ\Big(K\circ\mathcal{R}_{\omega\alpha},R_\alpha\Big) - \mathrm{D}_z H\circ(K\circ\mathcal{R}_{\omega\alpha},R_\alpha)E_K\\ =& \int_0^1 (1-s)\mathrm{D}_z^2 H\circ\Big(K\circ\mathcal{R}_{\omega\alpha} + sE_K, R_\alpha\Big)[E_K,E_K]\ ds.\end{aligned}$$

Since—by assumption—the integral Taylor remainders are well defined, we can obtain the estimates

$$|\Big\langle\eta_K^{31}\Big\rangle| \le \left(2nc_{\mathrm{D}a} + \sigma_{(\mathrm{D}_\theta K)^\top}c_{\mathrm{D}^2 a}\frac{\delta}{2}\right)\frac{\|E_K\|_\rho^2}{\delta} \le \frac{C_{\langle\eta_K^{31}\rangle}}{\delta}\|E_K\|_\rho^2, \tag{3.85}$$

$$|\Big\langle\eta_K^{32}\Big\rangle| \le \left(2n\ell c_{(\mathrm{D}a)^\top}|\hat{\alpha}| + (c_{\mathrm{D}_z^2 H} + 2n\ell c_{\mathrm{D}^2 a}\sigma_{\mathrm{D}_\varphi K}|\hat{\alpha}|)\frac{\delta}{2}\right)\frac{\|E_K\|_\rho^2}{\delta} \le \frac{C_{\langle\eta_K^{32}\rangle}}{\delta}\|E_K\|_\rho^2, \tag{3.86}$$

and

$$|\Big\langle\eta_K^{3}\Big\rangle| \le \max\Big\{C_{\langle\eta_K^{31}\rangle},\ C_{\langle\eta_K^{32}\rangle}\Big\}\frac{\|E_K\|_\rho^2}{\delta} \le \frac{C_{\langle\eta_K^{3}\rangle}}{\delta}\|E_K\|_\rho^2. \tag{3.87}$$

To control $E_{\mathrm{symp}P}$, we will use Lemma 3.10 which applies if $\|E_{\mathrm{symp}\hat{P}}\|_{\rho-2\delta} \le \hat{\nu} < 1$; which recall is one of our assumptions. Therefore, using Lemmas 3.10 and 3.11, we obtain

$$\begin{aligned}\|E_{\mathrm{lin}}^K\|_{\rho-2\delta} \le& \left(\Big(C_{E_{\mathrm{red}P}}^K + C_{E_{\mathrm{symp}P}}^K\Big)C_{\xi_K} + C_{\langle\eta_K^3\rangle}\gamma^3\delta^{3\tau}\right)\frac{\|E_K\|_\rho^2}{\gamma^3\delta^{3\tau+1}}\\ &+ \Big(C_{E_{\mathrm{red}P}}^W + C_{E_{\mathrm{symp}P}}^W\Big)C_{\xi_K}\frac{\|E_K\|_\rho\|E_W\|_\rho}{\gamma^2\delta^{2\tau}}\\ \le& \frac{C_{E_{\mathrm{lin}}^K}^K}{\gamma^3\delta^{3\tau+1}}\|E_K\|_\rho^2 + \frac{C_{E_{\mathrm{lin}}^K}^{KW}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho\|E_W\|_\rho.\end{aligned} \tag{3.88}$$

Observe that using Neumann series,

$$\|(I_{2n} - \Omega_0 E_{\mathrm{symp}P})^{-1}\|_{\rho-2\delta} \le \frac{1}{1-\nu}, \tag{3.89}$$

Also, note that $\Delta^2\phi_T$ is well-defined by assumption (3.69). Hence,

$$\|\Delta^2\phi_T\|_{\rho-2\delta} \le \frac{1}{2}c_{\mathrm{D}_z^2\phi_T}\|\Delta K\|_{\rho-2\delta}^2 \le \frac{1}{2}c_{\mathrm{D}_z^2\phi_T}(C_{\Delta K})^2\frac{\|E_K\|_\rho^2}{\gamma^4\delta^{4\tau}}.$$

We now have all the estimates to obtain

$$\|E_{\bar{K}}\|_{\rho-2\delta} \le \frac{C_P}{1-\nu}\left(\frac{C_{E_{\mathrm{lin}}^K}^K}{\gamma^3\delta^{3\tau+1}}\|E_K\|_\rho^2 + \frac{C_{E_{\mathrm{lin}}^K}^{KW}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho\|E_W\|_\rho\right) + \frac{c_{\mathrm{D}_z^2\phi_T}(C_{\Delta K})^2}{2\gamma^4\delta^{4\tau}}\|E_K\|_\rho^2 \tag{3.90}$$

$$\leq \left(\frac{C_P}{1-\nu}C^K_{E^K_{\rm lin}}\gamma\delta^{\tau-1} + \frac{1}{2}c_{{\rm D}^2_z\phi_T}C^2_{\Delta K}\right)\frac{\|E_K\|^2_\rho}{\gamma^4\delta^{4\tau}} + \frac{C_P}{1-\nu}C^{KW}_{E^K_{\rm lin}}\frac{\|E_K\|_\rho\|E_W\|_\rho}{\gamma^2\delta^{2\tau}}$$
$$\leq \frac{C^{KK}_{E_{\bar K}}}{\gamma^4\delta^{4\tau}}\|E_K\|^2_\rho + \frac{C^{KW}_{E_{\bar K}}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho\|E_W\|_\rho.$$

□

3.3.2 The linearized equation for the bundle

Let us now inspect the new error in the bundle invariance equation. First recall definition (2.56). The new error in the bundle invariance equation is the following

$$\begin{aligned} E_{\bar W} :=& {\rm D}_z\phi_T\circ(K+\Delta K, {\rm id})(W+\Delta W) - \Big(W\circ\mathcal{R}_{\omega\alpha} + \Delta W\circ\mathcal{R}_{\omega\alpha}\Big)(\lambda+\Delta\lambda) \\ =& \widetilde{E_W} + {\rm D}_z\phi_T\circ(K,{\rm id})\Delta W - \Delta W\circ\mathcal{R}_{\omega\alpha}\lambda - W\circ\mathcal{R}_{\omega\alpha}\Delta\lambda \\ &+ \Delta{\rm D}_z\phi_T[\Delta W] - \Delta W\circ\mathcal{R}_{\omega\alpha}\Delta\lambda, \end{aligned} \tag{3.91}$$

where, for any $\zeta : \mathbb{T}^{d+\ell}\to\mathbb{R}^{2n}$,

$$\Delta{\rm D}_z\phi_T[\zeta] = \int_0^1 {\rm D}^2_z\phi_T\circ(K+s\Delta K,{\rm id})[\Delta K,\zeta]ds.$$

Lemma 3.14. *The new error in the bundle invariance equation after correcting W by solving* (2.59) *is given by*

$$E_{\bar W} = P\circ\mathcal{R}_{\omega\alpha}\Big(I_{2n} - \Omega_0 E_{{\rm symp}P}\circ\mathcal{R}_{\omega\alpha}\Big)^{-1}E^W_{\rm lin} + \Delta{\rm D}_z\phi_T[\Delta W] - \Delta W\circ\mathcal{R}_{\omega\alpha}\Delta\lambda, \tag{3.92}$$

where

$$E^W_{\rm lin} := E_{{\rm red}P}\xi_W + \Omega_0 E_{{\rm symp}P}\circ\mathcal{R}_{\omega\alpha}\Big(\xi_W\circ\mathcal{R}_{\omega\alpha}\lambda + e_n\Delta\lambda\Big)$$

Additionally, for any $\delta\in(0,\rho/3)$

$$\|E_{\bar W}\|_{\rho-3\delta} \leq \frac{C^{KK}_{E_{\bar W}}}{\gamma^5\delta^{5\tau}}\|E_K\|^2_\rho + \frac{C^{WW}_{E_{\bar W}}}{\gamma\delta^\tau}\|E_W\|^2_\rho + \frac{C^{KW}_{E_{\bar W}}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho\|E_W\|_\rho. \tag{3.93}$$

Proof. Let us first inspect the sum

$$\widetilde{E_W} + {\rm D}_z\phi_T\circ(K,{\rm id})\Delta W - \Delta W\circ\mathcal{R}_{\omega\alpha}\lambda - W\circ\mathcal{R}_{\omega\alpha}\Delta\lambda.$$

If left-multiply it by $-\Omega_0 P^\top\circ\mathcal{R}_{\omega\alpha}\Omega\circ K\circ\mathcal{R}_{\omega\alpha}$ and by its inverse, we obtain

$$\begin{aligned} &\Big(-\Omega_0 P^\top\circ\mathcal{R}_{\omega\alpha}\Omega\circ K\circ\mathcal{R}_{\omega\alpha}\Big)^{-1}\Big(-\eta_W + \Lambda_P\xi_W + E_{{\rm red}P}\xi_W \\ &- (I_{2n} - \Omega_0 E_{{\rm symp}P}\circ\mathcal{R}_{\omega\alpha})\,(\xi_W\circ\mathcal{R}_{\omega\alpha}\lambda + e_n\Delta\lambda) \\ =& \Big(-\Omega_0 P^\top\circ\mathcal{R}_{\omega\alpha}\Omega\circ K\circ\mathcal{R}_{\omega\alpha}\Big)^{-1}\Big(E_{{\rm red}P}\xi_W + \Omega_0 E_{{\rm symp}P}\circ\mathcal{R}_{\omega\alpha}\Big(\xi_W\circ\mathcal{R}_{\omega\alpha}\lambda + e_n\Delta\lambda\Big)\Big) \\ =& \Big(-\Omega_0 P^\top\circ\mathcal{R}_{\omega\alpha}\Omega\circ K\circ\mathcal{R}_{\omega\alpha}\Big)^{-1}E^W_{\rm lin}, \end{aligned}$$

where we used definitions (3.64), (3.66), and that ξ_W satisfies (2.59). Result (3.92) then follows from using this last expression in (3.91), the assumption $\|E_{{\rm symp}P}\|_{\rho-2\delta}\leq\nu<1$, and

(3.73). In order to obtain estimate (3.93), note that $\Delta \mathrm{D}_z\phi_T$ is well-defined by assumption (3.69) and that

$$\|\widetilde{E_W}\|_{\rho-2\delta} \leq c_{\mathrm{D}_z^2\phi_T} C_{\Delta K}\sigma_W \frac{\|E_K\|_\rho}{\gamma^2\delta^{2\tau}} + \|E_W\|_\rho \leq \frac{C^K_{\widetilde{E_W}}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\widetilde{E_W}}\|E_W\|_\rho. \tag{3.94}$$

Then, we proceed to obtain estimates for each η^i_W as

$$\begin{aligned}
\|\eta^1_W\|_{\rho-2\delta} &\leq C_{N^\top} c_\Omega \left(\frac{C^K_{\widetilde{E_W}}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + \|E_W\|_\rho\right) && (3.95)\\
&\leq \frac{C^K_{\eta^1_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\eta^1_W}\|E_W\|_\rho, \\
\|\eta^2_W\|_{\rho-2\delta} &\leq C_{N^\top} c_\Omega \left(\frac{C^K_{\widetilde{E_W}}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + \|E_W\|_\rho\right) && (3.96)\\
&\leq \frac{C^K_{\eta^2_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\eta^2_W}\|E_W\|_\rho, \\
\|\eta^3_W\|_{\rho-2\delta} &\leq \max\left\{\sigma_{(\mathrm{D}_\theta K)^\top},\ C_{\mathcal{X}^\top}\right\} c_\Omega \left(\frac{C^K_{\widetilde{E_W}}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + \|E_W\|_\rho\right) && (3.97)\\
&\leq \frac{C^K_{\eta^3_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\eta^3_W}\|E_W\|_\rho, \\
\|\eta^4_W\|_{\rho-2\delta} &\leq \sigma_{W^\top} c_\Omega \left(\frac{C^K_{\widetilde{E_W}}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + \|E_W\|_\rho\right) && (3.98)\\
&\leq \frac{C^K_{\eta^4_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\eta^4_W}\|E_W\|_\rho,
\end{aligned}$$

and for each ξ^i_W we obtain from (2.65)-(2.68) the following

$$\begin{aligned}
\|\xi^1_W\|_{\rho-2\delta} &\leq \frac{1}{1-\sigma_\lambda}\left(\left(C^K_{\eta^1_W} + \frac{C_{\hat S} C^K_{\eta^3_W}}{1-\sigma_\lambda}\right)\frac{\|E_K\|_\rho}{\gamma^2\delta^{2\tau}} + \left(C^W_{\eta^1_W} + \frac{C_{\hat S} C^W_{\eta^3_W}}{1-\sigma_\lambda}\right)\|E_W\|_\rho\right) && (3.99)\\
&\leq \frac{C^K_{\xi^1_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\xi^1_W}\|E_W\|_\rho, \\
\|\xi^2_W\|_{\rho-3\delta} &\leq \frac{c_{\mathfrak{R}}\sigma_{\lambda^{-1}}}{\gamma\delta^\tau}\left(\frac{C^K_{\eta^2_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\eta^2_W}\|E_W\|_\rho\right) && (3.100)\\
&\leq \frac{C^K_{\xi^2_W}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{C^W_{\xi^2_W}}{\gamma\delta^\tau}\|E_W\|_\rho, \\
\|\xi^3_W\|_{\rho-2\delta} &\leq \frac{1}{1-\sigma_\lambda}\left(\frac{C^K_{\eta^3_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\eta^3_W}\|E_W\|_\rho\right) && (3.101)\\
&\leq \frac{C^K_{\xi^3_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\xi^3_W}\|E_W\|_\rho, \\
\|\xi^4_W\|_{\rho-2\delta} &\leq \frac{\sigma_\lambda}{1-(\sigma_\lambda)^2}\left(\frac{C^K_{\eta^4_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\eta^4_W}\|E_W\|_\rho\right) && (3.102)
\end{aligned}$$

$$\leq \frac{C^K_{\xi^4_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\xi^4_W}\|E_W\|_\rho.$$

Then,

$$\begin{aligned}
\|\xi_W\|_{\rho-3\delta} \leq \max\Bigg\{ & \frac{C^K_{\xi^1_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\xi^1_W}\|E_W\|_\rho,\ \frac{C^K_{\xi^2_W}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{C^W_{\xi^2_W}}{\gamma\delta^\tau}\|E_W\|_\rho, \\
& \frac{C^K_{\xi^3_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\xi^3_W}\|E_W\|_\rho,\ \frac{C^K_{\xi^4_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\xi^4_W}\|E_W\|_\rho\Bigg\} \\
\leq & \max\left\{C^K_{\xi^1_W}\gamma\delta^\tau,\ C^K_{\xi^2_W},\ C^K_{\xi^3_W}\gamma\delta^\tau,\ C^K_{\xi^4_W}\gamma\delta^\tau\right\}\frac{\|E_K\|_\rho}{\gamma^3\delta^{3\tau}} \\
& + \max\left\{C^W_{\xi^1_W}\gamma\delta^\tau,\ C^W_{\xi^2_W},\ C^W_{\xi^3_W}\gamma\delta^\tau,\ C^W_{\xi^4_W}\gamma\delta^\tau\right\}\frac{\|E_W\|_\rho}{\gamma\delta^\tau} \\
\leq & \frac{C^K_{\xi_W}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{C^W_{\xi_W}}{\gamma\delta^\tau}\|E_W\|_\rho.
\end{aligned} \tag{3.103}$$

We can then control the correction of the bundle parameterization and the rate of contraction as

$$\begin{aligned}
\|\Delta W\|_{\rho-3\delta} \leq\ & C_L\max\left\{\frac{C^K_{\xi^1_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\xi^1_W}\|E_W\|_\rho,\ \frac{C^K_{\xi^2_W}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{C^W_{\xi^2_W}}{\gamma\delta^\tau}\|E_W\|_\rho\right\} \\
& + C_N\max\left\{\frac{C^K_{\xi^3_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\xi^3_W}\|E_W\|_\rho,\ \frac{C^K_{\xi^4_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\xi^4_W}\|E_W\|_\rho\right\} \\
\leq\ & \left(C_L\max\left\{C^K_{\xi^1_W}\gamma\delta^\tau,\ C^K_{\xi^2_W}\right\} + C_N\max\left\{C^K_{\xi^3_W},\ C^K_{\xi^4_W}\right\}\gamma\delta^\tau\right)\frac{\|E_K\|_\rho}{\gamma^3\delta^{3\tau}} \\
& + \left(C_L\max\left\{C^W_{\xi^1_W}\gamma\delta^\tau,\ C^W_{\xi^2_W}\right\} + C_N\max\left\{C^W_{\xi^3_W},\ C^W_{\xi^4_W}\right\}\gamma\delta^\tau\right)\frac{\|E_W\|_\rho}{\gamma\delta^\tau} \\
\leq\ & \frac{C^K_{\Delta W}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{C^W_{\Delta W}}{\gamma\delta^\tau}\|E_W\|_\rho,
\end{aligned} \tag{3.104}$$

$$|\Delta\lambda| \leq \frac{C^K_{\eta^2_W}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\eta^2_W}\|E_W\|_\rho \leq \frac{C^K_{\Delta\lambda}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\Delta\lambda}\|E_W\|_\rho. \tag{3.105}$$

In order to control E^W_{lin}, we will use Lemmas 3.10 and 3.11. Recall that Lemma 3.10 applies from the assumption $\|E_{\text{symp}\hat{P}}\| \leq \hat{\nu} < 1$. Therefore, we can proceed as follows

$$\begin{aligned}
\|E^W_{\text{lin}}\|_{\rho-3\delta} \leq\ & \|E_{\text{red}P}\|_{\rho-2\delta}\|\xi_W\|_{\rho-3\delta} + \|E_{\text{symp}P}\|_{\rho-2\delta}\|\xi_W\|_{\rho-3\delta}|\lambda| + \|E_{\text{symp}P}\|_{\rho-2\delta}|\Delta\lambda| \\
\leq\ & \left(\frac{C^K_{E_{\text{red}P}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho + C^W_{E_{\text{red}P}}\|E_W\|_\rho\right)\left(\frac{C^K_{\xi_W}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{C^W_{\xi_W}}{\gamma\delta^\tau}\|E_W\|_\rho\right) \\
& + \left(\frac{C^K_{E_{\text{symp}P}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho + C^W_{E_{\text{symp}P}}\|E_W\|_\rho\right)\left(\frac{C^K_{\xi_W}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{C^W_{\xi_W}}{\gamma\delta^\tau}\|E_W\|_\rho\right)\sigma_\lambda \\
& + \left(\frac{C^K_{E_{\text{symp}P}}}{\gamma\delta^{\tau+1}}\|E_K\|_\rho + C^W_{E_{\text{symp}P}}\|E_W\|_\rho\right)\left(\frac{C^K_{\Delta\lambda}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\Delta\lambda}\|E_W\|_\rho\right)
\end{aligned} \tag{3.106}$$

$$
\begin{aligned}
&\leq \left(C^K_{E_{\mathrm{red}P}} C^K_{\xi_W} + C^K_{E_{\mathrm{symp}P}} C^K_{\xi_W} \sigma_\lambda + C^K_{E_{\mathrm{symp}P}} C^K_{\Delta\lambda} \gamma\delta^\tau\right) \frac{\|E_K\|_\rho^2}{\gamma^4 \delta^{4\tau+1}} \\
&\quad + \left(C^W_{E_{\mathrm{red}P}} C^W_{\xi_W} + C^W_{E_{\mathrm{symp}P}} C^W_{\xi_W} \sigma_\lambda + C^W_{E_{\mathrm{symp}P}} C^W_{\Delta\lambda} \gamma\delta^\tau\right) \frac{\|E_W\|_\rho^2}{\gamma\delta^\tau} \\
&\quad + \Bigg(C^W_{E_{\mathrm{red}P}} C^K_{\xi_W} + C^W_{E_{\mathrm{symp}P}} C^K_{\xi_W} + \left(C^K_{E_{\mathrm{red}P}} C^W_{\xi_W} + C^K_{E_{\mathrm{symp}P}} C^W_{\xi_W}\right) \gamma\delta^{\tau-1} \\
&\quad + C^W_{E_{\mathrm{symp}P}} C^K_{\Delta\lambda} \gamma\delta^\tau + C^K_{E_{\mathrm{symp}P}} C^W_{\Delta\lambda} \gamma^2 \delta^{2\tau-1}\Bigg) \frac{\|E_K\|_\rho \|E_W\|_\rho}{\gamma^3 \delta^{3\tau}} \\
&\leq \frac{C^{KK}_{E^W_{\mathrm{lin}}}}{\gamma^4 \delta^{4\tau+1}} \|E_K\|_\rho^2 + \frac{C^{KW}_{E^W_{\mathrm{lin}}}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho \|E_W\|_\rho + \frac{C^{WW}_{E^W_{\mathrm{lin}}}}{\gamma\delta^\tau} \|E_W\|_\rho^2 .
\end{aligned}
$$

Lastly, we can control the remaining terms in (3.92) as

$$
\begin{aligned}
\|\Delta \mathrm{D}_z \phi_T [\Delta W]\|_{\rho-3\delta} &\leq c_{\mathrm{D}_z^2 \phi_T} C_{\Delta K} \left(\frac{C^K_{\Delta W}}{\gamma^5 \delta^{5\tau}} \|E_K\|_\rho^2 + \frac{C^W_{\Delta W}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho \|E_W\|_\rho\right) \\
&\leq \frac{C^{KK}_{\Delta \mathrm{D}_z \phi_T [\Delta W]}}{\gamma^5 \delta^{5\tau}} \|E_K\|_\rho^2 + \frac{C^{KW}_{\Delta \mathrm{D}_z \phi_T [\Delta W]}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho \|E_W\|_\rho ,
\end{aligned} \tag{3.107}
$$

$$
\begin{aligned}
\|\Delta W \Delta\lambda\|_{\rho-3\delta} &\leq \frac{C^K_{\Delta W} C^K_{\Delta\lambda}}{\gamma^5 \delta^{5\tau}} \|E_K\|_\rho^2 + \frac{C^W_{\Delta W} C^W_{\Delta\lambda}}{\gamma\delta^\tau} \|E_W\|_\rho^2 \\
&\quad + \frac{C^K_{\Delta W} C^W_{\Delta\lambda} + C^W_{\Delta W} C^K_{\Delta\lambda}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho \|E_W\|_\rho \\
&\leq \frac{C^{KK}_{\Delta W \Delta\lambda}}{\gamma^5 \delta^{5\tau}} \|E_K\|_\rho^2 + \frac{C^{KW}_{\Delta W \Delta\lambda}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho \|E_W\|_\rho + \frac{C^{WW}_{\Delta W \Delta\lambda}}{\gamma\delta^\tau} \|E_W\|_\rho^2 .
\end{aligned} \tag{3.108}
$$

We can now use the assumption $\|E_{\mathrm{symp}P}\|_{\rho-2\delta} \leq \nu < 1$ and (3.89) to finally obtain the estimate (3.93) as

$$
\begin{aligned}
\|E_{\bar W}\|_{\rho-3\delta} &\leq \frac{C_P}{1-\nu} \left(\frac{C^{KK}_{E^W_{\mathrm{lin}}}}{\gamma^4 \delta^{4\tau+1}} \|E_K\|_\rho^2 + \frac{C^{KW}_{E^W_{\mathrm{lin}}}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho \|E_W\|_\rho + \frac{C^{WW}_{E^W_{\mathrm{lin}}}}{\gamma\delta^\tau} \|E_W\|_\rho^2\right) \\
&\quad + \frac{C^{KK}_{\Delta \mathrm{D}_z \phi_T [\Delta W]}}{\gamma^5 \delta^{5\tau}} \|E_K\|_\rho^2 + \frac{C^{KW}_{\Delta \mathrm{D}_z \phi_T [\Delta W]}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho \|E_W\|_\rho \\
&\quad + \frac{C^{KK}_{\Delta W \Delta\lambda}}{\gamma^5 \delta^{5\tau}} \|E_K\|_\rho^2 + \frac{C^{KW}_{\Delta W \Delta\lambda}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho \|E_W\|_\rho + \frac{C^{WW}_{\Delta W \Delta\lambda}}{\gamma\delta^\tau} \|E_W\|_\rho^2 \\
&\leq \left(\frac{C_P}{1-\nu} C^{KK}_{E^W_{\mathrm{lin}}} \gamma\delta^{\tau-1} + C^{KK}_{\Delta \mathrm{D}_z \phi_T [\Delta W]} + C^{KK}_{\Delta W \Delta\lambda}\right) \frac{\|E_K\|_\rho^2}{\gamma^5 \delta^{5\tau}} \\
&\quad + \left(\frac{C_P}{1-\nu} C^{KW}_{E^W_{\mathrm{lin}}} + C^{KW}_{\Delta \mathrm{D}_z \phi_T [\Delta W]} + C^{KW}_{\Delta W \Delta\lambda}\right) \frac{\|E_K\|_\rho \|E_W\|_\rho}{\gamma^3 \delta^{3\tau}} \\
&\quad + \left(\frac{C_P}{1-\nu} C^{WW}_{E^W_{\mathrm{lin}}} + C^{WW}_{\Delta W \Delta\lambda}\right) \frac{\|E_W\|_\rho^2}{\gamma\delta^\tau} \\
&\leq \frac{C^{KK}_{E_{\bar W}}}{\gamma^5 \delta^{5\tau}} \|E_K\|_\rho^2 + \frac{C^{KW}_{E_{\bar W}}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho \|E_W\|_\rho + \frac{C^{WW}_{E_{\bar W}}}{\gamma\delta^\tau} \|E_W\|_\rho^2 .
\end{aligned} \tag{3.109}
$$

□

3.4 Proof of the KAM theorem

As mentioned previously, the iterative scheme described in Section 2.5 allows the construction of a sequence of objects. The key idea for the proof of Theorem 3.1 consists in showing that the sequence for K, W, and λ converge—under certain conditions—to solutions of (2.3) and (2.6). First, we obtain an iterative lemma and derive conditions such that, if satisfied, the iterative lemma can be applied again to the objects obtained after one step of the Quasi-Newton method. Convergence then follows from ensuring that the conditions necessary for the iterative lemma are satisfied at every step of the iterative process.

3.4.1 The iterative lemma

After one step of the iterative scheme, we obtain the new parameterizations $\bar{K} \leftarrow K + \Delta K$ and $\bar{W} \leftarrow W + \Delta W$ and the new rate of contraction $\bar{\lambda} \leftarrow \lambda + \Delta\lambda$. From $\bar{K}, \bar{W}$, and $\bar{\lambda}$, we obtain the new objects $\bar{\hat{N}}, \bar{B}$, and $\bar{S}$. The following lemma allows the control of the new objects—and consequently the control of the iterative procedure.

Lemma 3.15 (The iterative lemma)**.** *Assume the hypothesis of Theorem 3.1. For every $\delta \in (0, \rho/3)$, there exists a constant C_Δ such that if*

$$\frac{C_\Delta}{\gamma\delta^\tau}\mathcal{E} < 1 \tag{3.110}$$

holds, where

$$\mathcal{E} = \max\left\{\frac{\|E_K\|_\rho}{\gamma^2\delta^{2\tau}},\ \|E_W\|_\rho\right\},$$

$$\begin{aligned} C_\Delta := \max\Bigg\{ & C_{E_{\mathrm{symp}P}}\gamma\delta^\tau,\ \frac{\max\{\gamma^2\delta^{2\tau}, C_{\Delta K}\}}{R - \|K - K_0\|_\rho}\gamma\delta^\tau,\ \frac{dC_{\Delta K}}{\sigma_{\mathrm{D}_\theta K} - \|\mathrm{D}_\theta K\|_\rho}\gamma\delta^{\tau-1}, \\ & \frac{\ell C_{\Delta K}}{\sigma_{\mathrm{D}_\varphi K} - \|\mathrm{D}_\varphi K\|_\rho}\gamma\delta^{\tau-1},\ \frac{2nC_{\Delta K}}{\sigma_{(\mathrm{D}_\theta K)^\top} - \|(\mathrm{D}_\theta K)^\top\|_\rho}\gamma\delta^{\tau-1}, \\ & \frac{2nC_{\Delta K}}{\sigma_{(\mathrm{D}_\varphi K)^\top} - \|(\mathrm{D}_\varphi K)^\top\|_\rho}\gamma\delta^{\tau-1},\ \frac{C_{\Delta W}}{\sigma_W - \|W\|_\rho},\ \frac{2nC_{\Delta W}}{\sigma_{W^\top} - \|W^\top\|_\rho}, \\ & \frac{C_{\Delta B}}{\sigma_B - \|B\|_\rho},\ \frac{C_{\Delta \hat{N}}}{\sigma_{\hat{N}} - \|\hat{N}\|_\rho},\ \frac{C_{\Delta \hat{N}^\top}}{\sigma_{\hat{N}^\top} - \|\hat{N}^\top\|_\rho},\ \frac{C_{\Delta\langle\mathcal{S}\rangle^{-1}}}{\sigma_{\langle\mathcal{S}\rangle^{-1}} - \left|\langle\mathcal{S}\rangle^{-1}\right|}, \\ & \frac{C_{\Delta\lambda}}{\sigma_\lambda - |\lambda|}\gamma\delta^\tau,\ \frac{C_{\Delta\lambda^{-1}}}{\sigma_{\lambda^{-1}} - |\lambda^{-1}|}\gamma\delta^\tau \Bigg\}, \end{aligned} \tag{3.111}$$

then we have new parameterizations $\bar{K} \in \left(\mathcal{A}\left(\mathbb{T}^{d+\ell}_{\rho-2\delta}\right)\right)^{2n}$, $\bar{W} \in \left(\mathcal{A}\left(\mathbb{T}^{d+\ell}_{\rho-3\delta}\right)\right)^{2n}$ and their associated objects satisfy

$$\|\mathrm{D}_\theta \bar{K}\|_{\rho-3\delta} < \sigma_{\mathrm{D}_\theta K}, \tag{3.112}$$

$$\|\mathrm{D}_\varphi \bar{K}\|_{\rho-3\delta} < \sigma_{\mathrm{D}_\varphi K}, \tag{3.113}$$

$$\|(\mathrm{D}_\theta \bar{K})^\top\|_{\rho-3\delta} < \sigma_{(\mathrm{D}_\theta K)^\top}, \tag{3.114}$$

$$\|(\mathrm{D}_\varphi \bar{K})^\top\|_{\rho-3\delta} < \sigma_{(\mathrm{D}_\varphi K)^\top}, \tag{3.115}$$

$$\|\bar{W}\|_{\rho-3\delta} < \sigma_W, \tag{3.116}$$
$$\|\bar{W}^\top\|_{\rho-3\delta} < \sigma_{W^\top}, \tag{3.117}$$
$$\|\hat{\bar{N}}\|_{\rho-3\delta} < \sigma_{\hat{N}}, \tag{3.118}$$
$$\|\hat{\bar{N}}^\top\|_{\rho-3\delta} < \sigma_{\hat{N}^\top}, \tag{3.119}$$
$$\|\bar{B}\|_{\rho-3\delta} < \sigma_B, \tag{3.120}$$
$$\left|\left\langle \bar{\mathcal{S}} \right\rangle^{-1}\right| < \sigma_{\langle \mathcal{S}\rangle^{-1}}. \tag{3.121}$$

Furthermore, the hyperbolicity is controlled by

$$|\bar{\lambda}| < \sigma_\lambda, \quad |\bar{\lambda}^{-1}| < \sigma_{\lambda^{-1}}, \tag{3.122}$$

and the new objects are close to the original ones in the following sense

$$\|\bar{K} - K\|_{\rho-2\delta} \leq C_{\Delta K}\mathcal{E}, \tag{3.123}$$
$$\|\mathrm{D}_\theta \bar{K} - \mathrm{D}_\theta K\|_{\rho-3\delta} \leq \frac{dC_{\Delta K}}{\delta}\mathcal{E}, \tag{3.124}$$
$$\|\mathrm{D}_\varphi \bar{K} - \mathrm{D}_\varphi K\|_{\rho-3\delta} \leq \frac{\ell C_{\Delta K}}{\delta}\mathcal{E}, \tag{3.125}$$
$$\|(\mathrm{D}_\theta \bar{K})^\top - (\mathrm{D}_\theta K)^\top\|_{\rho-3\delta} \leq \frac{2nC_{\Delta K}}{\delta}\mathcal{E}, \tag{3.126}$$
$$\|(\mathrm{D}_\varphi \bar{K})^\top - (\mathrm{D}_\varphi K)^\top\|_{\rho-3\delta} \leq \frac{2nC_{\Delta K}}{\delta}\mathcal{E}, \tag{3.127}$$
$$\|\bar{W} - W\|_{\rho-3\delta} \leq \frac{C_{\Delta W}}{\gamma\delta^\tau}\mathcal{E}, \tag{3.128}$$
$$\|\bar{W}^\top - W^\top\|_{\rho-3\delta} \leq \frac{2nC_{\Delta W}}{\gamma\delta^\tau}\mathcal{E}, \tag{3.129}$$
$$\|\bar{B} - B\|_{\rho-3\delta} \leq \frac{C_{\Delta B}}{\gamma\delta^\tau}\mathcal{E}, \tag{3.130}$$
$$\|\hat{\bar{N}} - \hat{N}\|_{\rho-3\delta} \leq \frac{C_{\Delta\hat{N}}}{\gamma\delta^\tau}\mathcal{E}, \tag{3.131}$$
$$\|\hat{\bar{N}}^\top - \hat{N}^\top\|_{\rho-3\delta} \leq \frac{C_{\Delta\hat{N}^\top}}{\gamma\delta^\tau}\mathcal{E}, \tag{3.132}$$
$$\left|\left\langle \bar{\mathcal{S}} \right\rangle^{-1} - \langle \mathcal{S}\rangle^{-1}\right| \leq \frac{C_{\Delta\langle \mathcal{S}\rangle^{-1}}}{\gamma\delta^\tau}\mathcal{E}, \tag{3.133}$$
$$|\bar{\lambda} - \lambda| \leq C_{\Delta\lambda}\mathcal{E}, \tag{3.134}$$
$$|\bar{\lambda}^{-1} - \lambda^{-1}| \leq C_{\Delta\lambda^{-1}}\mathcal{E}. \tag{3.135}$$

Lastly, the new errors in the invariance equations

$$E_{\bar{K}} = \phi_T \circ (\bar{K}, \mathrm{id}) - \bar{K}\circ \mathcal{R}_{\omega\alpha},$$
$$E_{\bar{W}} = \mathrm{D}_z\phi_T \circ (\bar{K}, \mathrm{id})\bar{W} - \bar{W}\circ\mathcal{R}_{\omega\alpha}\,\bar{\lambda},$$

satisfy

$$\|E_{\bar{K}}\|_{\rho-2\delta} \leq C_{E_{\bar{K}}}\mathcal{E}^2, \quad \|E_{\bar{W}}\|_{\rho-3\delta} \leq \frac{C_{E_{\bar{W}}}}{\gamma\delta^\tau}\mathcal{E}^2. \tag{3.136}$$

Proof. Recall that in order to use Lemmas 3.12 and 3.14, we need hypothesis (3.55), (3.70), (3.19), and (3.69) to hold. Let us take for the values of $\hat{\nu}$ and ν the following

$$\hat{\nu} = \left(C^K_{E_{\mathrm{symp}\hat{P}}} \gamma\delta^{\tau-1} + C^W_{E_{\mathrm{symp}\hat{P}}} \right) \mathcal{E} = C_{E_{\mathrm{symp}\hat{P}}} \mathcal{E}, \tag{3.137}$$

$$\nu = \left(C^K_{E_{\mathrm{symp}P}} \gamma\delta^{\tau-1} + C^W_{E_{\mathrm{symp}P}} \right) \mathcal{E} = C_{E_{\mathrm{symp}P}} \mathcal{E}, \tag{3.138}$$

which clearly satisfy $\|E_{\mathrm{symp}\hat{P}}\|_{\rho-2\delta} \leq \hat{\nu}$ and $\|E_{\mathrm{symp}P}\|_{\rho-2\delta} \leq \nu$. It is not hard to show that $\hat{\nu} \leq \nu$; see Table E.3. Then, from hypothesis (3.110) for the first term in (3.111) we conclude that (3.55) and (3.70) are satisfied. For (3.19) and (3.69), observe that by hypothesis (3.110) for the second term in (3.111)

$$\frac{\max\{\gamma^2\delta^{2\tau}, C_{\Delta K}\}\gamma\delta^\tau}{R - \|K - K_0\|_\rho} \frac{\mathcal{E}}{\gamma\delta^\tau} < 1.$$

Then, from the definition of $\mathcal{E}$ and estimate (3.84), we conclude that (3.19) and (3.69) are also satisfied. Furthermore, estimates (3.123), (3.128), and (3.129) are immediate from (3.84) and (3.104), respectively. For estimates (3.112)-(3.115), we proceed as follows

$$\begin{aligned}
\|\mathrm{D}_\theta \bar{K}\|_{\rho-3\delta} &\leq \|\mathrm{D}_\theta K\|_\rho + \|\mathrm{D}_\theta \Delta K\|_{\rho-3\delta} \leq \|\mathrm{D}_\theta K\|_\rho + \frac{d C_{\Delta K}}{\gamma^2\delta^{2\tau+1}} \|E_K\|_\rho, \\
\|\mathrm{D}_\varphi \bar{K}\|_{\rho-3\delta} &\leq \|\mathrm{D}_\varphi K\|_\rho + \|\mathrm{D}_\varphi \Delta K\|_{\rho-3\delta} \leq \|\mathrm{D}_\varphi K\|_\rho + \frac{\ell C_{\Delta K}}{\gamma^2\delta^{2\tau+1}} \|E_K\|_\rho, \\
\|(\mathrm{D}_\theta \bar{K})^\top\|_{\rho-3\delta} &\leq \|(\mathrm{D}_\theta K)^\top\|_\rho + \|(\mathrm{D}_\theta \Delta K)^\top\|_{\rho-3\delta} \leq \|(\mathrm{D}_\theta K)^\top\|_\rho + \frac{2n C_{\Delta K}}{\gamma^2\delta^{2\tau+1}} \|E_K\|_\rho, \\
\|(\mathrm{D}_\varphi \bar{K})^\top\|_{\rho-3\delta} &\leq \|(\mathrm{D}_\varphi K)^\top\|_\rho + \|(\mathrm{D}_\varphi \Delta K)^\top\|_{\rho-3\delta} \leq \|(\mathrm{D}_\varphi K)^\top\|_\rho + \frac{2n C_{\Delta K}}{\gamma^2\delta^{2\tau+1}} \|E_K\|_\rho,
\end{aligned}$$

where we used Cauchy estimates on ΔK. Note that these Cauchy estimates result directly in (3.124)-(3.127). Then,

$$\begin{aligned}
\|\mathrm{D}_\theta \bar{K}\|_{\rho-3\delta} &\leq \|\mathrm{D}_\theta K\|_\rho + \frac{d C_{\Delta K}}{\delta} \mathcal{E}, \\
\|\mathrm{D}_\varphi \bar{K}\|_{\rho-3\delta} &\leq \|\mathrm{D}_\varphi K\|_\rho + \frac{\ell C_{\Delta K}}{\delta} \mathcal{E}, \\
\|(\mathrm{D}_\theta \bar{K})^\top\|_{\rho-3\delta} &\leq \|(\mathrm{D}_\theta K)^\top\|_\rho + \frac{2n C_{\Delta K}}{\delta} \mathcal{E}, \\
\|(\mathrm{D}_\varphi \bar{K})^\top\|_{\rho-3\delta} &\leq \|(\mathrm{D}_\varphi K)^\top\|_\rho + \frac{2n C_{\Delta K}}{\delta} \mathcal{E},
\end{aligned}$$

and by hypothesis (3.110)—in particular for the third to sixth terms in (3.111)—we conclude

$$\begin{aligned}
\|\mathrm{D}_\theta \bar{K}\|_{\rho-3\delta} &< \sigma_{\mathrm{D}_\theta K}, & \|(\mathrm{D}_\theta \bar{K})^\top\|_{\rho-3\delta} &< \sigma_{(\mathrm{D}_\theta K)^\top}, \\
\|\mathrm{D}_\varphi \bar{K}\|_{\rho-3\delta} &< \sigma_{\mathrm{D}_\varphi K}, & \|(\mathrm{D}_\varphi \bar{K})^\top\|_{\rho-3\delta} &< \sigma_{(\mathrm{D}_\varphi K)^\top}.
\end{aligned}$$

Similarly, for the corrected parameterization of the bundle, from estimate (3.104) we have

$$\begin{aligned}
\|\bar{W}\|_{\rho-3\delta} &\leq \|W\|_\rho + \frac{C^K_{\Delta W}}{\gamma^3\delta^{3\tau}} \|E_K\|_\rho + \frac{C^W_{\Delta W}}{\gamma\delta^\tau} \|E_W\|_\rho \\
&\leq \|W\|_\rho + \frac{C_{\Delta W}}{\gamma\delta^\tau} \mathcal{E},
\end{aligned} \tag{3.139}$$

$$\begin{aligned}\|\bar{W}^\top\|_{\rho-3\delta} &\leq \|W^\top\|_\rho + \frac{2nC^K_{\Delta W}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{2nC^W_{\Delta W}}{\gamma\delta^\tau}\|E_W\|_\rho \\ &\leq \|W^\top\|_\rho + \frac{2nC_{\Delta W}}{\gamma\delta^\tau}\mathcal{E}.\end{aligned}$$

By hypothesis (3.110) for the seventh and eighth terms in (3.111), we obtain estimates (3.116) and (3.117). In order to control $\bar{L}$, note that

$$\Delta L := \bar{L} - L = \left(\mathrm{D}_\theta \Delta K \;\middle|\; \widetilde{\Delta X} - \mathrm{D}_\varphi \Delta K \hat{\alpha} \;\middle|\; \Delta W\right),$$

where

$$\widetilde{\Delta X} := \int_0^1 \mathrm{D}_z X \circ (K + s\Delta K, \mathrm{id})\, \Delta K ds,$$

which is well-defined in $\mathbb{T}^d_{\rho-2\delta} \times \mathbb{T}^\ell_{\rho-2\delta}$ by hypothesis (3.110) for the second term in (3.111). Then, using (3.84), (3.104), and Cauchy estimates, we obtain

$$\begin{aligned}\|\bar{L} - L\|_{\rho-3\delta} &\leq \left((d + \ell|\hat{\alpha}| + c_{\mathrm{D}_z X}\delta)\, C_{\Delta K}\gamma\delta^{\tau-1} + C^K_{\Delta W}\right)\frac{\|E_K\|_\rho}{\gamma^3\delta^{3\tau}} + \frac{C^W_{\Delta W}}{\gamma\delta^\tau}\|E_W\|_\rho \qquad (3.140)\\ &\leq \frac{C^K_{\Delta L}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{C^W_{\Delta L}}{\gamma\delta^\tau}\|E_W\|_\rho \\ &\leq \frac{C_{\Delta L}}{\gamma\delta^\tau}\mathcal{E},\\ \|\bar{L}^\top - L^\top\|_{\rho-3\delta} &\leq \max\left\{\frac{2n}{\delta}\|\Delta K\|_{\rho-2\delta},\ 2n\left(c_{(\mathrm{D}_z X)^\top} + \frac{|\hat{\alpha}^\top|}{\delta}\right)\|\Delta K\|_{\rho-2\delta},\ 2n\|\Delta W\|_{\rho-3\delta}\right\} \qquad (3.141)\\ &\leq \frac{2n}{\delta}\max\left\{1, c_{(\mathrm{D}_z X)^\top}\delta + |\hat{\alpha}^\top|\right\}\|\Delta K\|_{\rho-2\delta} + 2n\|\Delta W\|_{\rho-3\delta}\\ &\leq 2n\left(\max\left\{1, c_{(\mathrm{D}_z X)^\top}\delta + |\hat{\alpha}^\top|\right\}C_{\Delta K}\gamma\delta^{\tau-1} + C^K_{\Delta W}\right)\frac{\|E_K\|_\rho}{\gamma^3\delta^{3\tau}} + \frac{2nC^W_{\Delta W}}{\gamma\delta^\tau}\|E_W\|_\rho\\ &\leq \frac{C^K_{\Delta L^\top}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{C^W_{\Delta L^\top}}{\gamma\delta^\tau}\|E_W\|_\rho\\ &\leq \frac{C_{\Delta L^\top}}{\gamma\delta^\tau}\mathcal{E}.\end{aligned}$$

Let us now control $\Delta B := \bar{B} - B$. First, for the restricted metric, we have that

$$G_{\bar{L}} - G_L = L^\top \Delta G \bar{L} + L^\top G \circ K \Delta L + \Delta L^\top G \circ \bar{K} \bar{L},$$

where

$$\Delta G := G \circ \bar{K} - G \circ K = \int_0^1 \mathrm{D}G\,(K + s\Delta K)\,\Delta K ds,$$

which is well-defined in $\mathbb{T}^d_{\rho-2\delta} \times \mathbb{T}^\ell_{\rho-2\delta}$ by hypothesis (3.110) for the second term in (3.111). Then,

$$\begin{aligned}\|G_{\bar{L}} - G_L\|_{\rho-3\delta} &\leq C_{L^\top} c_{\mathrm{D}G} C_L \|\Delta K\|_{\rho-2\delta} + C_{L^\top} c_G \|\Delta L\|_{\rho-3\delta} + c_G C_L \|\Delta L^\top\|_{\rho-3\delta} \qquad (3.142)\\ &\leq \left(C_{L^\top} c_{\mathrm{D}G} C_L C_{\Delta K}\gamma\delta^\tau + C_{L^\top} c_G C^K_{\Delta L} + c_G C_L C^K_{\Delta L^\top}\right)\frac{\|E_K\|_\rho}{\gamma^3\delta^{3\tau}}\end{aligned}$$

$$
\begin{aligned}
&+ \left(C_{L^\top} c_G C^W_{\Delta L} + c_G C_L C^W_{\Delta L^\top}\right) \frac{\|E_W\|_\rho}{\gamma\delta^\tau} \\
&\leq \frac{C^K_{\Delta G_L}}{\gamma^3\delta^{3\tau}} \|E_K\|_\rho + \frac{C^W_{\Delta G_L}}{\gamma\delta^\tau} \|E_W\|_\rho \\
&\leq \frac{C_{\Delta G_L}}{\gamma\delta^\tau} \mathcal{E}.
\end{aligned}
$$

We will now use the following result that is presented for matrices but extends to matrix valued maps [FH24].

Lemma 3.16. *Let $M \in \mathbb{C}^{n\times n}$ be an invertible matrix satisfying $|M^{-1}| < \sigma$ and assume that $\bar{M} \in \mathbb{C}^{n\times n}$ satisfies*

$$
\frac{\sigma^2 |\bar{M} - M|}{\sigma - |M^{-1}|} < 1.
$$

Then, $\bar{M}$ is invertible and

$$
|\bar{M}^{-1}| < \sigma, \quad |\bar{M}^{-1} - M^{-1}| \leq \sigma^2 |\bar{M} - M|.
$$

Note that by hypothesis (3.110) for the ninth term in (3.111),

$$
\frac{(\sigma_B)^2 \|G_{\bar{L}} - G_L\|_{\rho-3\delta}}{\sigma_B - \|B\|_\rho} < 1.
$$

Therefore, applying Lemma 3.16 to $G_{\bar{L}}(\theta,\varphi)$ and $G_L(\theta,\varphi)$, we obtain (3.130) and (3.120) as

$$
\begin{aligned}
\|\bar{B} - B\|_{\rho-3\delta} &\leq (\sigma_B)^2 \|G_{\bar{L}} - G_L\|_{\rho-3\delta} \\
&\leq \frac{(\sigma_B)^2 C^K_{\Delta G_L}}{\gamma^3\delta^{3\tau}} \|E_K\|_\rho + \frac{(\sigma_B)^2 C^W_{\Delta G_L}}{\gamma\delta^\tau} \|E_W\|_\rho \\
&\leq \frac{C^K_{\Delta B}}{\gamma^3\delta^{3\tau}} \|E_K\|_\rho + \frac{C^W_{\Delta B}}{\gamma\delta^\tau} \|E_W\|_\rho \\
&\leq \frac{C_{\Delta B}}{\gamma\delta^\tau} \mathcal{E}, \\
\|\bar{B}\|_{\rho-3\delta} &< \sigma_B.
\end{aligned} \tag{3.143}
$$

Let us proceed and control $\Delta\hat{N} := \hat{\bar{N}} - \hat{N}$ and $\Delta\hat{N}^\top := \hat{\bar{N}}^\top - \hat{N}^\top$. Note that

$$
\hat{\bar{N}} - \hat{N} = J \circ K L \Delta B + J \circ K \Delta L \bar{B} + \Delta J \bar{L} \bar{B},
$$

where

$$
\Delta J := J \circ \bar{K} - J \circ K = \int_0^1 \mathrm{D}J\Big(K + s\Delta K\Big) \Delta K ds
$$

is well-defined in $\mathbb{T}^d_{\rho-2\delta} \times \mathbb{T}^\ell_{\rho-2\delta}$ by hypothesis (3.110) for the second term in (3.111). Then, we obtain (3.131) and (3.132) as

$$
\begin{aligned}
\|\hat{\bar{N}} - \hat{N}\|_{\rho-3\delta} &\leq c_J C_L \|\Delta B\|_{\rho-3\delta} + c_J \|\Delta L\|_{\rho-3\delta} \sigma_B + c_{\mathrm{D}J} \|\Delta K\|_{\rho-2\delta} C_L \sigma_B \\
&\leq \left(c_J \left(C_L C^K_{\Delta B} + C^K_{\Delta L} \sigma_B\right) + c_{\mathrm{D}J} C_{\Delta K} C_L \sigma_B \gamma\delta^\tau\right) \frac{\|E_K\|_\rho}{\gamma^3\delta^{3\tau}}
\end{aligned} \tag{3.144}
$$

$$
\begin{aligned}
&+ c_J \left(C_L C^W_{\Delta B} + C^W_{\Delta L} \sigma_B \right) \frac{\|E_W\|_\rho}{\gamma \delta^\tau} \\
&\leq \frac{C^K_{\Delta \hat{N}}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho + \frac{C^W_{\Delta \hat{N}}}{\gamma \delta^\tau} \|E_W\|_\rho \\
&\leq \frac{C_{\Delta \hat{N}}}{\gamma \delta^\tau} \mathcal{E},
\end{aligned}
$$

$$
\begin{aligned}
\|\hat{\bar{N}}^\top - \hat{N}^\top\|_{\rho-3\delta} &\leq C_{L^\top} c_{J^\top} \|\Delta B\|_{\rho-3\delta} + \sigma_B \|\Delta L^\top\|_{\rho-3\delta} c_{J^\top} + \sigma_B C_{L^\top} c_{(\mathrm{D}J)^\top} 2n \|\Delta K\|_{\rho-2\delta} \qquad (3.145) \\
&\leq \left(c_{J^\top} \left(C_{L^\top} C^K_{\Delta B} + \sigma_B C^K_{\Delta L^\top} \right) + \sigma_B C_{L^\top} c_{(\mathrm{D}J)^\top} 2n C_{\Delta K} \gamma \delta^\tau \right) \frac{\|E_K\|_\rho}{\gamma^3 \delta^{3\tau}} \\
&\quad + c_{J^\top} \left(C_{L^\top} C^W_{\Delta B} + \sigma_B C^W_{\Delta L^\top} \right) \frac{\|E_W\|_\rho}{\gamma \delta^\tau} \\
&\leq \frac{C^K_{\Delta \hat{N}^\top}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho + \frac{C^W_{\Delta \hat{N}^\top}}{\gamma \delta^\tau} \|E_W\|_\rho \\
&\leq \frac{C_{\Delta \hat{N}^\top}}{\gamma \delta^\tau} \mathcal{E}.
\end{aligned}
$$

Also, by hypothesis (3.110) for the tenth and eleventh terms in (3.111),

$$
\begin{aligned}
\|\hat{\bar{N}}\|_{\rho-3\delta} &\leq \|\hat{N}\|_\rho + \frac{C_{\Delta \hat{N}}}{\gamma \delta^\tau} \mathcal{E} < \sigma_{\hat{N}}, \\
\|\hat{\bar{N}}^\top\|_{\rho-3\delta} &\leq \|\hat{N}^\top\|_\rho + \frac{C_{\Delta \hat{N}^\top}}{\gamma \delta^\tau} \mathcal{E} < \sigma_{\hat{N}^\top},
\end{aligned}
$$

so we have estimates (3.118) and (3.119). Note that, according to (3.11), the new torsion satisfies

$$
\|\hat{\bar{S}}\|_{\rho-3\delta} < C_{\hat{S}}.
$$

Similarly as for $\Delta \hat{N}$, we obtain

$$
\begin{aligned}
\|\hat{\bar{S}} - \hat{S}\|_{\rho-3\delta} &\leq \sigma_{\hat{N}^\top} c_\Omega c_{\mathrm{D}_z \phi_T} \|\Delta \hat{N}\|_{\rho-3\delta} + \sigma_{\hat{N}^\top} \sigma_{\hat{N}} \left(c_{\mathrm{D}\Omega} c_{\mathrm{D}_z \phi_T} + c_\Omega c_{\mathrm{D}^2_z \phi_T} \right) \|\Delta K\|_{\rho-2\delta} \qquad (3.146) \\
&\quad + \|\Delta \hat{N}^\top\|_{\rho-3\delta} c_\Omega c_{\mathrm{D}_z \phi_T} \sigma_{\hat{N}} \\
&\leq \Big(\sigma_{\hat{N}^\top} c_\Omega c_{\mathrm{D}_z \phi_T} C^K_{\Delta \hat{N}} + \sigma_{\hat{N}^\top} \sigma_{\hat{N}} \left(c_{\mathrm{D}\Omega} c_{\mathrm{D}_z \phi_T} + c_\Omega c_{\mathrm{D}^2_z \phi_T} \right) C_{\Delta K} \gamma \delta^\tau \\
&\quad + C^K_{\Delta \hat{N}^\top} c_\Omega c_{\mathrm{D}_z \phi_T} \sigma_{\hat{N}} \Big) \frac{\|E_K\|_\rho}{\gamma^3 \delta^{3\tau}} + \left(\sigma_{\hat{N}^\top} c_\Omega c_{\mathrm{D}_z \phi_T} C^W_{\Delta \hat{N}} + C^W_{\Delta \hat{N}^\top} c_\Omega c_{\mathrm{D}_z \phi_T} \sigma_{\hat{N}} \right) \frac{\|E_W\|_\rho}{\gamma \delta^\tau} \\
&\leq \frac{C^K_{\Delta \hat{S}}}{\gamma^3 \delta^{3\tau}} \|E_K\|_\rho + \frac{C^W_{\Delta \hat{S}}}{\gamma \delta^\tau} \|E_W\|_\rho \\
&\leq \frac{C_{\Delta \hat{S}}}{\gamma \delta^\tau} \mathcal{E}.
\end{aligned}
$$

In order to obtain estimates (3.121) and (3.133), we need to control the reduced torsion $\mathcal{S}$. We will use Lemma 3.16 which holds if

$$
\frac{\left(\sigma_{\langle \mathcal{S} \rangle^{-1}} \right)^2 \left| \left\langle \bar{\mathcal{S}} \right\rangle - \langle \mathcal{S} \rangle \right|}{\sigma_{\langle \mathcal{S} \rangle^{-1}} - \left| \langle \mathcal{S} \rangle^{-1} \right|} < 1. \qquad (3.147)
$$

Observe that

$$
\left| \left\langle \bar{\mathcal{S}} \right\rangle - \langle \mathcal{S} \rangle \right| \leq \left| \left\langle \hat{\bar{S}} \right\rangle - \left\langle \hat{S} \right\rangle \right| \leq \|\hat{\bar{S}} - \hat{S}\|_{\rho-3\delta},
$$

and by hypothesis (3.110) for the twelfth term in (3.111),

$$\frac{\left(\sigma_{\langle\mathcal{S}\rangle^{-1}}\right)^2 \|\hat{\bar{S}} - \hat{S}\|_{\rho-3\delta}}{\sigma_{\langle\mathcal{S}\rangle^{-1}} - \left|\langle\mathcal{S}\rangle^{-1}\right|} < 1.$$

Therefore, (3.147) holds, $\left\langle\bar{\mathcal{S}}\right\rangle$ is invertible, and

$$\begin{aligned}
\left|\left\langle\bar{\mathcal{S}}\right\rangle^{-1} - \langle\mathcal{S}\rangle^{-1}\right| &\le \left(\sigma_{\langle\mathcal{S}\rangle^{-1}}\right)^2 \|\hat{\bar{S}} - \hat{S}\|_{\rho-3\delta} \\
&\le \frac{\left(\sigma_{\langle\mathcal{S}\rangle^{-1}}\right)^2 C^K_{\Delta\hat{S}}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{\left(\sigma_{\langle\mathcal{S}\rangle^{-1}}\right)^2 C^W_{\Delta\hat{S}}}{\gamma\delta^\tau}\|E_W\|_\rho \\
&\le \frac{C^K_{\Delta\langle\mathcal{S}\rangle^{-1}}}{\gamma^3\delta^{3\tau}}\|E_K\|_\rho + \frac{C^W_{\Delta\langle\mathcal{S}\rangle^{-1}}}{\gamma\delta^\tau}\|E_W\|_\rho \\
&\le \frac{C_{\Delta\langle\mathcal{S}\rangle^{-1}}}{\gamma\delta^\tau}\mathcal{E}, \\
\left|\left\langle\bar{\mathcal{S}}\right\rangle^{-1}\right| &< \sigma_{\langle\mathcal{S}\rangle^{-1}}.
\end{aligned} \tag{3.148}$$

For the new rate of contraction, observe that

$$|\bar{\lambda}| \le |\lambda| + |\Delta\lambda|, \quad |\bar{\lambda}^{-1}| \le |\lambda^{-1}| + |\Delta\lambda^{-1}|.$$

Therefore, the hyperbolicity controls (3.122) hold if

$$\frac{|\Delta\lambda|}{\sigma_\lambda - |\lambda|} < 1, \quad \frac{|\Delta\lambda^{-1}|}{\sigma_{\lambda^{-1}} - |\lambda^{-1}|} < 1.$$

Note that from bound (3.105) and hypothesis (3.110) for the thirteenth term in (3.111)

$$\frac{|\Delta\lambda|}{\sigma_\lambda - |\lambda|} \le \frac{1}{\sigma_\lambda - |\lambda|}\left(\frac{C^K_{\Delta\lambda}}{\gamma^2\delta^{2\tau}}\|E_K\|_\rho + C^W_{\Delta\lambda}\|E_W\|_\rho\right) \le \frac{C_{\Delta\lambda}}{\sigma_\lambda - |\lambda|}\mathcal{E} < 1. \tag{3.149}$$

Also, hypothesis (3.110) for the last term in (3.111) allow us to use Lemma 3.16 and obtain

$$\begin{aligned}
|\Delta\lambda^{-1}| &\le (\sigma_{\lambda^{-1}})^2 C_{\Delta\lambda}\mathcal{E}, \\
|\bar{\lambda}^{-1}| &\le \sigma_{\lambda^{-1}}.
\end{aligned} \tag{3.150}$$

For estimates (3.136) we will use Lemmas 3.12 and 3.14. From (3.72) we have that

$$\|E_{\bar{K}}\|_{\rho-2\delta} \le \left(C^{KK}_{E_{\bar{K}}} + C^{KW}_{E_{\bar{K}}}\right)\mathcal{E}^2 \le C_{E_{\bar{K}}}\mathcal{E}^2. \tag{3.151}$$

Also, from (3.93), we obtain

$$\begin{aligned}
\|E_{\bar{W}}\|_{\rho-3\delta} &\le \left(C^{KK}_{E_{\bar{W}}} + C^{WW}_{E_{\bar{W}}} + C^{KW}_{E_{\bar{W}}}\right)\frac{\mathcal{E}^2}{\gamma\delta^\tau} \\
&\le \frac{C_{E_{\bar{W}}}}{\gamma\delta^\tau}\mathcal{E}^2.
\end{aligned} \tag{3.152}$$

□

3.4.2 Convergence of the iterative procedure and proof of the KAM theorem

We show that under the assumptions of Theorem 3.1, condition (3.110) in Lemma 3.15 holds at every step of the iterative scheme. Therefore, Lemma 3.15 can be applied infinitely many times. In the limit, the sequences of objects converge and define invariant tori and its invariant bundles.

Consider the objects at the iterative step j denoted by K_j, W_j, λ_j, $\hat{N}_j$, B_j, and $\mathcal{S}_j$. Consider also the errors E_{K_j} and E_{W_j} in the invariance equations (2.41) (2.42), respectively, the analyticity domain ρ_j, and analyticity bite δ_j. For the step $j = 0$, consider the objects declared in theorem 3.1. That is, take $K_0 = K$, $W_0 = W$, $\lambda_0 = \lambda$, $\hat{N}_0 = \hat{N}$, $B_0 = B$, $\mathcal{S}_0 = \mathcal{S}$, $\rho_0 = \rho$, $\delta_0 = \delta$, and $\mathcal{E}_0 = \mathcal{E}$. Note that the domain of analyticity of the objects decreases at every step according to

$$\rho_j = \rho_{j-1} - 3\delta_{j-1},$$

where δ_j is the domain bite chosen for step j. Let us then choose the following geometric sequence of bites

$$\delta_j = \frac{\delta}{a^j},$$

for some $a > 1$—which results in the limiting width

$$\rho_\infty = \rho - 3\delta \frac{a}{a-1}$$

and

$$a = \frac{\rho - \rho_\infty}{\rho_1 - \rho_\infty}.$$

Let us now define

$$\mathcal{E}_j := \max\left\{\frac{\|E_{K_j}\|_{\rho_j}}{\gamma^2 \delta_j^{2\tau}}, \|E_{W_j}\|_{\rho_j}\right\}, \qquad C_{\mathcal{E}_j} := \max\left\{E_{\bar{K}_j} a^{2\tau}, E_{\bar{W}_j} \gamma \delta_j\right\}, \tag{3.153}$$

$$\varkappa_j := \frac{\mathcal{E}_j}{\gamma^2 \delta_j^{2\tau}}, \qquad \kappa_j := a^{2\tau} C_{\mathcal{E}_j} \varkappa_j$$

and assume (3.110) in Lemma 3.15 holds for the objects up to step $j-1$. Then,

$$\varkappa_j \le a^{2\tau} C_{\mathcal{E}_{j-1}} \varkappa_{j-1}^2 \le \kappa_{j-1} \varkappa_{j-1}$$

and

$$\kappa_j \le a^{2\tau} C_{\mathcal{E}_j} \kappa_{j-1} \varkappa_{j-1} \le \kappa_{j-1}^2,$$

where the last inequality holds as long as $C_{\mathcal{E}_j} \le C_{\mathcal{E}_{j-1}}$. In order to apply Lemma 3.15 again we need (3.110) to hold for the objects in the step j. Notice that C_Δ in (3.111), as well as the constants therein, depend on δ_j—polynomially by construction—and on $\hat{\nu}_j$ and ν_j. See Appendix E and definitions (3.137) and (3.138). Let us now assume $\kappa := \kappa_0 < 1$ and $\nu_0 < 1$, which are included in assumption 3.2. Recall that $\hat{\nu} \le \nu$. Additionally, the assumption $\kappa < 1$ implies decreasing sequences for $\hat{\nu}_j$, ν_j, $C_{\mathcal{E}_j}$, κ_j, $\varkappa_j$, and $\mathcal{E}_j$. In particular,

$$\kappa_j \le \kappa_{j-1}^2 \le \kappa^{2^j}$$

$$\mathcal{E}_j \leq \gamma^2 \delta_j^{2\tau} \varkappa_j \leq \gamma^2 \delta_j^{2\tau} \kappa_0^{2^{j-1}} \varkappa_0 \leq \left(\frac{\kappa}{a^{2\tau}}\right)^j \mathcal{E}. \tag{3.154}$$

In order to prove convergence, it will be necessary to control the constants within C_Δ uniformly with respect to j. Hence, the constants within C_Δ are taken for the largest values of δ, $\hat{\nu}$, and ν, i.e., $\delta = \delta_0$, $\hat{\nu} = \hat{\nu}_0$, and $\nu = \nu_0$. Nonetheless, the explicit δ in C_Δ can be taken to be the corresponding ones for the step j, i.e., they can be taken to be $\delta = \delta_j$.

We will now use estimate (3.154) to check condition (3.110) for step j. For the first term in (3.111), condition (3.110) reads

$$C_{E_{\mathrm{symp}P}} \mathcal{E}_j < 1$$

and, using (3.154), we have

$$C_{E_{\mathrm{symp}P}} \mathcal{E}_j \leq C_{E_{\mathrm{symp}P}} \left(\frac{\kappa}{a^{2\tau}}\right)^j \mathcal{E} \leq C_{E_{\mathrm{symp}P}} \mathcal{E} < 1$$

where the last inequality holds from the assumption $\nu_0 < 1$. For the second term in (3.111), let us first compute

$$\|K_j - K\|_{\rho_j} \leq \sum_{i=0}^{j-1} \|K_{i+1} - K_i\|_{\rho_{i+1}} \leq C_{\Delta K} \mathcal{E} \sum_{i=0}^{j-1} \left(\frac{\kappa}{a^{2\tau}}\right)^i \leq \frac{a^{2\tau}}{a^{2\tau} - \kappa} C_{\Delta K} \mathcal{E} := \mathfrak{C}_{\Delta K} \mathcal{E}, \tag{3.155}$$

where we used (3.123) and (3.154). Notice that the previous bound holds for all j—hence, we can obtain estimate (3.4). Similarly, using Lemma 3.15, we obtain the estimates (3.5)-(3.7). Let us now consider the inequality

$$\frac{\gamma^2 \delta_j^{2\tau} \mathcal{E}_j}{R - \|K_j - K\|_{\rho_j}} < 1$$

which is satisfied if

$$\gamma^2 \delta_j^{2\tau} \mathcal{E}_j + \mathfrak{C}_{\Delta K} \mathcal{E} < R$$

or, alternatively,

$$\frac{\gamma^2 \delta^{2\tau} + \mathfrak{C}_{\Delta K}}{R} \mathcal{E} < 1,$$

where we used that $\delta_j < \delta$ and $\mathcal{E}_j < \mathcal{E}$. It is not hard to show that the previous condition implies (3.110) for the second term in (3.111) at every step; which is included in assumption (3.2). In order to express condition (3.110) for the third term in (3.111), we need to control $\|\mathrm{D}_\theta K_j\|_{\rho_j}$. Notice that we can recurrently obtain the bound

$$\begin{aligned}
\|\mathrm{D}_\theta K_j\|_{\rho_j} &\leq \|\mathrm{D}_\theta K\|_\rho + \sum_{i=0}^{j-1} \frac{dC_{\Delta K}}{\delta_i} \mathcal{E}_i \\
&\leq \|\mathrm{D}_\theta K\|_\rho + \frac{dC_{\Delta K}}{\delta} \sum_{i=0}^{\infty} \left(\frac{\kappa}{a^{2\tau-1}}\right)^i \mathcal{E} \\
&\leq \|\mathrm{D}_\theta K\|_\rho + \frac{dC_{\Delta K} a^{2\tau-1}}{\delta(a^{2\tau-1} - \kappa)} \mathcal{E} < \sigma_{\mathrm{D}_\theta K},
\end{aligned}$$

where the last inequality holds if

$$\frac{dC_{\Delta K}}{\delta(\sigma_{\mathrm{D}_\theta K} - \|\mathrm{D}_\theta K\|_\rho)} \frac{a^{2\tau-1}}{a^{2\tau-1}-\kappa}\mathcal{E} < 1,$$

which is included in assumption (3.2). Notice that this condition guarantees (3.110) for the third term in (3.111). We can proceed similarly and obtain conditions such that (3.110) holds at every step for the remaining terms in (3.111). Such conditions are then included in (3.2) for which we have the explicit constants

$$\mathfrak{C} := \max\left\{a^{2\tau} C_{\mathcal{E}},\ \mathfrak{C}_\Delta \gamma\delta^\tau\right\}, \tag{3.156}$$

$$\begin{aligned}\mathfrak{C}_\Delta := \max\Bigg\{ & C_{E_{\mathrm{symp}P}}\gamma\delta^\tau,\ \frac{\mathfrak{C}_{\tilde{\mathcal{U}}}}{R}\gamma\delta^\tau,\ \frac{\mathfrak{C}_{\Delta \mathrm{D}_\theta K}}{\sigma_{\mathrm{D}_\theta K} - \|\mathrm{D}_\theta K\|_\rho}\gamma\delta^{\tau-1},\ \frac{\mathfrak{C}_{\Delta \mathrm{D}_\varphi K}}{\sigma_{\mathrm{D}_\varphi K} - \|\mathrm{D}_\varphi K\|_\rho}\gamma\delta^{\tau-1}, \\ & \frac{\mathfrak{C}_{\Delta (\mathrm{D}_\theta K)^\top}}{\sigma_{(\mathrm{D}_\theta K)^\top} - \|(\mathrm{D}_\theta K)^\top\|_\rho}\gamma\delta^{\tau-1},\ \frac{\mathfrak{C}_{\Delta (\mathrm{D}_\varphi K)^\top}}{\sigma_{(\mathrm{D}_\varphi K)^\top} - \|(\mathrm{D}_\varphi K)^\top\|_\rho}\gamma\delta^{\tau-1},\ \frac{\mathfrak{C}_{\Delta W}}{\sigma_W - \|W\|_\rho}, \\ & \frac{\mathfrak{C}_{\Delta W^\top}}{\sigma_{W^\top} - \|W^\top\|_\rho},\ \frac{\mathfrak{C}_{\Delta B}}{\sigma_B - \|B\|_\rho},\ \frac{\mathfrak{C}_{\Delta \hat{N}}}{\sigma_{\hat{N}} - \|\hat{N}\|_\rho},\ \frac{\mathfrak{C}_{\Delta \hat{N}^\top}}{\sigma_{\hat{N}^\top} - \|\hat{N}^\top\|_\rho},\ \frac{\mathfrak{C}_{\Delta \langle\mathcal{S}\rangle^{-1}}}{\sigma_{\langle\mathcal{S}\rangle^{-1}} - |\langle\mathcal{S}\rangle^{-1}|}, \\ & \frac{\mathfrak{C}_{\Delta\lambda}}{\sigma_\lambda - |\lambda|}\gamma\delta^\tau,\ \frac{\mathfrak{C}_{\Delta\lambda^{-1}}}{\sigma_{\lambda^{-1}} - |\lambda^{-1}|}\gamma\delta^\tau \Bigg\},\end{aligned}$$

where

$$\begin{aligned}
&\mathfrak{C}_{\tilde{\mathcal{U}}} := \gamma^2\delta^{2\tau} + \mathfrak{C}_{\Delta K}, && \mathfrak{C}_{\Delta \mathrm{D}_\theta K} := \frac{a^{2\tau-1}}{a^{2\tau-1}-\kappa} dC_{\Delta K}, && \mathfrak{C}_{\Delta \mathrm{D}_\varphi K} := \frac{a^{2\tau-1}}{a^{2\tau-1}-\kappa}\ell C_{\Delta K}, \\
&\mathfrak{C}_{\Delta (\mathrm{D}_\theta K)^\top} := \frac{a^{2\tau-1}}{a^{2\tau-1}-\kappa} 2nC_{\Delta K}, && \mathfrak{C}_{\Delta (\mathrm{D}_\varphi K)^\top} := \mathfrak{C}_{\Delta (\mathrm{D}_\theta K)^\top}, && \mathfrak{C}_{\Delta W} := \frac{a^\tau}{a^\tau-\kappa} C_{\Delta W}, \\
&\mathfrak{C}_{\Delta W^\top} := \frac{a^\tau}{a^\tau-\kappa} 2nC_{\Delta W}, && \mathfrak{C}_{\Delta B} := \frac{a^\tau}{a^\tau-\kappa} C_{\Delta B}, && \mathfrak{C}_{\Delta \hat{N}} := \frac{a^\tau}{a^\tau-\kappa} C_{\Delta \hat{N}}, \\
&\mathfrak{C}_{\Delta \hat{N}^\top} := \frac{a^\tau}{a^\tau-\kappa} C_{\Delta \hat{N}^\top}, && \mathfrak{C}_{\Delta \langle\mathcal{S}\rangle^{-1}} := \frac{a^\tau}{a^\tau-\kappa} C_{\Delta \langle\mathcal{S}\rangle^{-1}}, && \mathfrak{C}_{\Delta\lambda} := \frac{a^{2\tau}}{a^{2\tau}-\kappa} C_{\Delta\lambda} \\
& && \mathfrak{C}_{\Delta\lambda^{-1}} := \frac{a^{2\tau}}{a^{2\tau}-\kappa} C_{\Delta\lambda^{-1}}.
\end{aligned}$$

Observe that the first term in (3.156) corresponds to the condition of having a decreasing sequence of errors whereas the second term corresponds to the conditions necessary for applying Lemma 3.15 infinitely many times.

Chapter 4

Whiskered tori: a unified approach

Both in theory and applications, it is of great interest to identify the asymptotic behavior of the orbits of a dynamical system—specially when the orbits approach invariant sets. Such kind of behavior is described by the stable and unstable manifolds of hyperbolic objects, e.g., fixed points, periodic orbits, or whiskered tori. A torus is said to be whiskered if there exist invariant directions in its tangent space that contract and expand exponentially under a flow or a map. These hyperbolic directions correspond to the invariant bundles which are the linear analogues to the stable and unstable manifolds of the torus—also known as whiskers. Hyperbolic tori and their dynamical connections have deep implications in the kind of dynamics to be expected and provide a framework to study dynamical systems both locally and globally.

The existence of homoclinic and heteroclinic orbits was already known to Poincaré, in what he called doubly asymptotic solutions [Poi87], which began to unravel the complexity of dynamical systems exhibiting homoclinic and heteroclinic phenomena. In [Arn64], Arnold suggested that transition chains of whiskered tori constitute a general mechanism for diffusion in dynamical systems which has been studied extensively. See for instance the perturbative geometrical methods of [DdlLS06] and references therein. Rigorous proofs of chaotic behaviour came from the work of Smale who showed that if a dynamical system has transverse homoclinic solutions, there exist an invariant Cantor set containing such solutions where the dynamics is topologically conjugate to the horseshoe map and, consequently, to the shift automorphism on the space of bi-infinite sequences [Sma65, Sma67].

In applications, dynamical systems techniques have been used in space mission design over the last decades [GLMS01, GSLM01, GJSM01a, GJSM01b]. Dynamical connections between hyperbolic objects can serve as zero-fuel trajectories; for instance, the Genesis [HBW97] and Artemis [SBA$^+$14] space missions used trajectories close to heteroclinic connections for the design of their itineraries. Given the fundamental role of stable and unstable manifolds in shaping system dynamics, considerable effort has been devoted to developing rigorous and computational methods for identifying and analyzing whiskered tori across various dynamical settings.

Under the parameterization method, invariant tori and their whiskers were computed in [HdlL06b, HdlL06a] for quasi-periodically forced maps. Following the same paradigm, [FdlLS09] obtains a-posteriori theorems for the existence of whiskered tori in exact symplectic maps and exact symplectic vector fields. Similarly, fast iterative schemes were described in [HdlLS12] for both Lagrangian and whiskered tori in Hamiltonian systems. For resonant tori see [KAdlL21a] and for a different approach based on quasi-periodic Floquet transformations for response tori see [GJNO22]. The invariant bundles and whiskers can then subsequently be used to compute dynamical connections. For instance, numerical techniques were developed in [BMO09, BMO13] to compute families of heteroclinic and homoclinic connections between periodic orbits in Hamiltonian systems whereas computer-assisted methods for individual connections were used in [WZ03, WZ05]. Naturally, dynamical connections between higher dimensional objects also exist. See for instance [GKL+04, CM08] for the CRTBP and [KAdlL21a, KAdlL21b] for connections between resonant tori in the planar circular and elliptic problems, respectively. For entire families between center manifolds see [BHM24].

The goal of this chapter is to contribute in the computation of invariant tori and their whiskers by means of KAM schemes. We provide a unified approach to compute tori and whiskers simultaneously by solving a functional equation extending the framework of [HdlLS12]. We focus in the case of rank-1 whiskers but the methodology can be adapted to the case of whiskers of arbitrary rank. Following the flow map methods of Chapters 2 and 3, we obtain tori and whiskers in autonomous and quasi-periodic Hamiltonian systems. As opposed to the order-by-order methods, which compute the torus and its bundles before computing the whiskers term by term, we correct every term of the parameterizations and double the number of correct terms with each iteration.

The obtained parameterizations provide a local picture that can later be globalized to study dynamical connections. Due to computational limitations, the local description is affected by the order of the expansions and the quality of the computed coefficients. In order to maximize the region of the phase space where we have an accurate description of the dynamics, we take careful consideration into the order of the parameterizations and the quality of the output of the iterative scheme, i.e., we aim to maximize the so called *fundamental domain* of the parameterizations provided by our methods. We have applied our techniques in the CRTBP where we compute a large set of non-resonant $2-$dimensional invariant tori and their stable/unstable manifolds. Furthermore, we outline the schemes for the quasi-periodic case and compute a few 3-dimensional tori and their invariant manifolds in the ERTBP.

The structure of the chapter is the following. In Sections 4.1 and 4.2 we discuss whiskered tori and their geometric properties for autonomous and quasi-periodic Hamiltonian systems, respectively. These sections also describe the construction of adapted frames and iterative schemes for the computation and continuation of whiskers. Section 4.3 discusses how to study the fundamental domain of the parameterizations and Section 4.4 addresses computational aspects of the methods of Sections 4.1 and 4.2. Lastly, in Sections 4.5 and 4.6, we apply our techniques to the Circular and Elliptic Restricted Three-Body Problems, respectively.

4.1 Whiskers in autonomous Hamiltonian systems

We will first cover the computation of whiskered tori for autonomous Hamiltonian systems and leave the quasi-periodic case for Section 4.2. The setting is analogous to that of

Section 1.2. For any smooth function $H : U \to \mathbb{R}$, we can construct a Hamiltonian vector field $X : U \to \mathbb{R}^{2n}$ as

$$X(z) := \Omega(z)^{-1} \mathrm{D}H(z)^{\top}, \tag{4.1}$$

where $\Omega : U \to \mathbb{R}^{2n\times 2n}$ is the matrix representation of the symplectic form $\boldsymbol{\omega}$. Let $\phi : D \subset \mathbb{R} \times U \to U$ denote the flow for the vector field X where D is the domain of definition of the flow, open. We will use the standard notation $\phi(t,z) = \phi_t(z)$.

For fixed t, the diffeomorphism ϕ_t is exact symplectic—that is, ϕ_t preserves the symplectic form as

$$\phi_t^* \boldsymbol{\omega} = \boldsymbol{\omega},$$

and the 1-form $\phi_t^* \boldsymbol{\alpha} - \boldsymbol{\alpha}$ is exact for some smooth primitive function, see [HM21] for the primitive function in an autonomous setting and Appendix A for the primitive function in quasi-periodic Hamiltonian systems.

We follow the approach of Chapter 2 and consider $(d+1)$-dimensional invariant tori $\hat{\mathcal{K}} \subset U$ with frequency vector $\hat{\omega} \in \mathbb{R}^{d+1}$ and $d = n-2$. For a suitable T, we can look for a codimension 1 generating torus $\mathcal{K} \subset \hat{\mathcal{K}}$, invariant under ϕ_T, see Section 2.1. More specifically, let us define $\hat{\omega} := \frac{1}{T}(\omega, 1)$, where we assume $\omega \in \mathbb{R}^d$ to be Diophantine. We can look for parameterizations $K : \mathbb{T}^d \to U$ that conjugate the dynamics in $\mathbb{T}^d$ under ϕ_T to a rigid rotation with rotation vector ω. That is, we require the parameterization to satisfy the following invariance equation

$$\phi_T \circ K - K \circ \mathcal{R}_\omega = 0, \tag{4.2}$$

where the map $\mathcal{R}_\omega$ defines the dynamics in $\mathcal{K}$ and is given by

$$\begin{aligned} \mathcal{R}_\omega : \quad \mathbb{T}^d &\longrightarrow \mathbb{T}^d \\ \theta &\longmapsto \theta + \omega. \end{aligned}$$

An invariant torus $\mathcal{K}$ is whiskered if it has directions that contract exponentially fast in the future and in the past under iteration of the differential of ϕ_T. These directions are the invariant bundles of the torus and together with the tangent to the torus and its symplectic conjugate, span the tangent bundle of U restricted to $\mathcal{K}$. Tangent to the exponentially contracting directions, the torus has attached invariant manifolds that converge exponentially fast to the torus—the whiskers.

In order to make the definition precise, let us first revisit the invariance equation (4.2) and observe that $\mathcal{K}$ is invariant under the map

$$\phi_T^{\circ n} := \underbrace{\phi_T \circ \cdots \circ \phi_T}_{n \text{ times}}.$$

Note that since $\phi_T^{-1} = \phi_{-T}$, we can define $\phi_T^{\circ n}$ for all $n \in \mathbb{Z}$.

Definition 4.1. An invariant torus $\mathcal{K} = K(\mathbb{T}^d)$ is said to be whiskered if for every $\theta \in \mathbb{T}^d$, $T_{K(\theta)}U$ admits an invariant splitting

$$T_{K(\theta)}U = E^s_{K(\theta)} \oplus E^u_{K(\theta)} \oplus E^c_{K(\theta)}$$

such that there exist constants $C > 0, \lambda < 1$, and $\mu > 1$, satisfying $\lambda\mu < 1$ and

$$\begin{aligned} v \in E^s_{K(\theta)} &\Longleftrightarrow |\Big(\mathrm{D}\phi_T^{\circ n} \circ K\Big)(v)| \leq C\lambda^n |v| \quad n > 0 \\ v \in E^u_{K(\theta)} &\Longleftrightarrow |\Big(\mathrm{D}\phi_T^{\circ n} \circ K\Big)(v)| \leq C\lambda^{-n} |v| \quad n < 0 \\ v \in E^c_{K(\theta)} &\Longleftrightarrow |\Big(\mathrm{D}\phi_T^{\circ n} \circ K\Big)(v)| \leq C\mu^{|n|} |v| \quad n \in \mathbb{Z}. \end{aligned}$$

It is possible to make Definition 4.3 in terms of *cocycles* instead of using the map $\mathrm{D}\phi_T^{\circ n}$. We leave such formulation for the quasi-periodic case of Section 4.2

Remark 4.1. We consider rank-1 bundles—hence reducible to constants. Also, since ϕ_T is symplectic, the rates of contraction/expansion of the stable and unstable subspaces are reciprocals.

4.1.1 Invariance equation for whiskers

We are interested in an invariant manifold $\mathcal{W}$, containing a rotational torus $\mathcal{K}$ and its associated whisker. We look for a parameterization $W : \mathbb{T}^d \times \mathbb{R} \to U$ such that under the map ϕ_T, the parameterization conjugates the dynamics on the torus to a rigid rotation and the dynamics on the whisker to a constant contraction. Note that we can do this without loss of generality since the case where the dynamics on the whisker are conjugate to a constant expansion is completely analogous but for the symplectomorphism ϕ_{-T}. Therefore, in what follows, we assume $|\lambda| < 1$. The conjugacy condition translates into W satisfying the invariance equation

$$\phi_T \circ W - W \circ \mathcal{R}^\lambda_\omega = 0, \tag{4.3}$$

where the map

$$\begin{aligned} \mathcal{R}^\lambda_\omega : \quad \mathbb{T}^d \times \mathbb{R} &\longrightarrow \mathbb{T}^d \times \mathbb{R} \\ (\theta, s) &\longmapsto (\theta + \omega, \lambda s) \end{aligned}$$

represents the internal dynamics of $\mathcal{W}$ under ϕ_T. Note that at $s = 0$, we recover the invariance equation for $\mathcal{K}$ which is parameterized by $K(\theta) := W(\theta, 0)$.

If we differentiate the invariance equation (4.3), we obtain

$$\mathrm{D}\phi_T \circ W \ \mathrm{D}W = \mathrm{D}W \circ \mathcal{R}^\lambda_\omega \ \mathrm{D}\mathcal{R}^\lambda_\omega, \tag{4.4}$$

where

$$\mathrm{D}\mathcal{R}^\lambda_\omega = \left(\begin{array}{c|c} I_d & \\ \hline & \lambda \end{array}\right).$$

This expression reveals that the invariance equation for $\mathcal{W}$ carries a bundle invariant under $\mathrm{D}\phi_T$.

Remark 4.2. If the vector field X is reversible, like the CRTBP, we can obtain parameterizations of the unstable whisker from that of the stable whisker. More specifically, if there exists an involution $\varrho : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ satisfying

$$X \circ \varrho = -\mathrm{D}\varrho X,$$

then we can obtain a parameterization of the unstable whisker, say $\tilde{W}$, from the stable whisker, and vice versa, as $\tilde{W} = \varrho \circ W$. Note that the torus $\mathcal{K}$ parameterized by $W(\theta, 0)$ is different from the torus parameterized by $\tilde{W}(\theta, 0)$. Nonetheless, both $W(\theta, 0)$ and $\tilde{W}(\theta, 0)$ generate the same torus $\hat{\mathcal{K}}$, see Chapter 2.

4.1.2 Geometric properties and adapted frames for $\mathcal{W}$

It is a classical result that tori invariant under exact symplectic maps with dynamics conjugate to ergodic rotations are isotropic manifolds—for the proof in the quasi-periodic case see the proof of Lemma 2.1. Such result also holds for the invariant manifolds of whiskered tori due to hyperbolicity. In fact, we can construct invariant Lagrangian frames for $\mathcal{W}$ that allow the existence of coordinates that reduce the linearized dynamics to upper triangular form. The construction is analogous to that of Section 2.4 and is as follows.

Recall that, according to (4.4), $\mathrm{D}W$ is invariant under $\mathrm{D}\phi_T$. Additionally, a quick observation reveals that

$$\mathrm{D}\phi_T \circ W \, X \circ W = X \circ W \circ \mathcal{R}^{\lambda}_{\omega}, \tag{4.5}$$

where we used the commutativity of flows and the invariance of $\mathcal{W}$. Consequently $X \circ W$ is also invariant under the differential of ϕ_T on $\mathcal{W}$. Let us construct a subframe $L : \mathbb{T}^d \times \mathbb{R} \to \mathbb{R}^{2n \times n}$ as

$$L := \begin{pmatrix} \mathrm{D}_\theta W & X \circ W & \mathrm{D}_s W \end{pmatrix}. \tag{4.6}$$

Proposition 4.2. *The subframe L constructed as in* (4.6) *is a Lagrangian subframe.*

Proof. If L is a Lagrangian subframe, it needs to satisfy

$$\Omega_L := L^\top \Omega \circ W L = 0.$$

First note that from (4.4) and (4.5) we have that

$$\mathrm{D}\phi_T \circ W L - L \circ \mathcal{R}^{\lambda}_{\omega} \Lambda = 0, \tag{4.7}$$

where Λ is as in (2.12). Therefore, we can see that L is invariant under $\mathrm{D}\phi_T \circ W$. Additionally, using the symplecticity of ϕ_T, (4.3), and (4.7) we obtain

$$\Lambda^{-\top} \Omega_L - \Omega_L \circ \mathcal{R}^{\lambda}_{\omega} \, \Lambda = 0. \tag{4.8}$$

Let us split Ω_L into blocks of sizes $(n-1) \times (n-1), (n-1) \times 1, 1 \times (n-1)$, and 1×1 as

$$\Omega_L = \begin{pmatrix} \Omega_L^{11} & \Omega_L^{12} \\ \Omega_L^{21} & \Omega_L^{22} \end{pmatrix}.$$

Then, (4.8) translates into

$$\begin{aligned} \Omega_L^{11} - {}& \Omega_L^{11} \circ \mathcal{R}^{\lambda}_{\omega} = 0, \\ \Omega_L^{12} - {}& \lambda \Omega_L^{12} \circ \mathcal{R}^{\lambda}_{\omega} = 0, \\ \lambda^{-1} \Omega_L^{21} - {}& \Omega_L^{21} \circ \mathcal{R}^{\lambda}_{\omega} = 0, \end{aligned}$$

$$\lambda^{-1}\Omega_L^{22} - \lambda\Omega_L^{22}\circ\mathcal{R}_\omega^\lambda = 0,$$

and if we use Fourier-Taylor series, we obtain

$$\sum_{j,k}\left(\Omega_L^{11}\right)_{jk}\left(1 - e^{i2\pi k\omega}\lambda^j\right)e^{i2\pi k\theta}s^j = 0,$$
$$\sum_{j,k}\left(\Omega_L^{12}\right)_{jk}\left(1 - e^{i2\pi k\omega}\lambda^{j+1}\right)e^{i2\pi k\theta}s^j = 0,$$
$$\sum_{j,k}\left(\Omega_L^{21}\right)_{jk}\left(\lambda^{-1} - e^{i2\pi k\omega}\lambda^j\right)e^{i2\pi k\theta}s^j = 0,$$
$$\sum_{j,k}\left(\Omega_L^{22}\right)_{jk}\left(\lambda^{-1} - e^{i2\pi k\omega}\lambda^{j+1}\right)e^{i2\pi k\theta}s^j = 0$$

for all θ and s. Hence, necessarily, $\Omega_L^{11} = (\Omega_L^{11})_{00}$, $\Omega_L^{12} = \Omega_L^{21} = \Omega_L^{22} = 0$—which shows that Ω_L is constant. Therefore, we can inspect the Lagrangian character of L on the torus $\mathcal{K}$. That is,

$$\Omega_L(\theta, s) = \Omega_L(\theta, 0).$$

Furthermore, it is known, see e.g. [HM21], that L generates Lagrangian subspaces on $T_{\mathcal{K}}U$, i.e., $\Omega_L(\theta, 0) = 0$. Therefore, $\Omega_L = 0$. For the proof in the quasi-periodic case, see the proof of Proposition 2.2.

It remains to be shown that $\mathrm{D}W$ and $X\circ W$ generate an $n-$dimensional subspace. This can be done without much work by looking at the vector field invariance equation for the whisker generated by $\mathcal{W}$. □

Remark 4.3. Note that the fact that L is a Lagrangian frame implies that $\mathcal{W}$ is an isotropic manifold, i.e,

$$(\mathrm{D}W)^\top\Omega\circ W\mathrm{D}W = 0.$$

Let us now complement the subframe L. In particular, we can complement L with a normal bundle $N : \mathbb{T}^d \times \mathbb{R} \to \mathbb{R}^{2n\times n}$ to a full symplectic basis of $T_{\mathcal{W}}U$ given by a frame $P : \mathbb{T}^d \times \mathbb{R} \to \mathbb{R}^{2n\times 2n}$ as

$$P := (L\ N)\,, \tag{4.9}$$

where

$$N := J\circ WLG_L^{-1} \tag{4.10}$$

and

$$G_L := L^\top G\circ WL.$$

Note that every Lagrangian subspace can be complemented with another Lagrangian subspace to a full symplectic basis [CdS01]. We emphasize that the frame P has geometric properties that are fundamental in our methods. In particular, by construction, P is symplectic with respect to the standard symplectic form. That is,

$$P^\top\Omega\circ WP = \Omega_0 = \left(\begin{array}{c|c} & -I_n \\ \hline I_n & \end{array}\right) \tag{4.11}$$

This identity shows that, on $\mathcal{W}$, the symplectic form in the coordinates given by P reduces to the standard symplectic form. If we right-multiply (4.11) by P^{-1} and then right-multiply it by $-\Omega_0$, we obtain the following formula for the inverse of P

$$P^{-1} = -\Omega_0 P^\top\Omega\circ W. \tag{4.12}$$

Also, the differential of ϕ_T on $\mathcal{W}$ in the coordinates given by P reduce to an upper triangular matrix-valued map constant along the diagonal. That is

$$(P \circ \mathcal{R}_\omega^\lambda)^{-1} \mathrm{D}\phi_T \circ W P = \left(\begin{array}{c|c} \Lambda & S \\ \hline & \Lambda^{-\top} \end{array} \right), \tag{4.13}$$

where $S : \mathbb{T}^d \times \mathbb{R} \to \mathbb{R}^{n \times n}$ is known as the torsion of $\mathcal{W}$ and is defined as

$$S = N^\top \circ \mathcal{R}_\omega^\lambda \Omega \circ W \circ \mathcal{R}_\omega^\lambda \mathrm{D}\phi_T \circ W N. \tag{4.14}$$

The reducibility of $\mathrm{D}\phi_T$ on $\mathcal{W}$ follows from (4.7) and that P satisfies symplecticity, i.e., (4.11). Recall that in Chapter 2, we further reduced the torsion S on the torus $\mathcal{K}$ so we could obtain both the stable and the unstable bundles simultaneously as part of the frame P. This followed from constructing the subframe N as in (2.26) to obtain such reduction. Following the analogous approach on the whisker, it is only possible to reduce the torsion by means of a linear symplectic transformation to

$$\mathcal{S} = \begin{pmatrix} \langle S_0^1 \rangle & \langle S_1^2 \rangle s \\ \langle S_1^3 \rangle s & \langle S_2^4 \rangle s^2 \end{pmatrix}.$$

Therefore, the torsion S can be reduced to constant in $\mathbb{T}^d$ and only with terms up to order 2 in the coordinate s—consequently we cannot obtain both the stable and unstable whiskers by means of a linear symplectic transformation on P.

4.1.3 Computation of invariant tori and whiskers

For the computation of $\mathcal{W}$, assume we have a parameterization $W : \mathbb{T}^d \times \mathbb{R} \to U$ and λ that satisfy (4.3) approximately. Let us define the error in the invariance of $\mathcal{W}$ as $E : \mathbb{T}^d \times \mathbb{R} \to \mathbb{R}^{2n}$ given by

$$E := \phi_T \circ W - W \circ \mathcal{R}_\omega^\lambda. \tag{4.15}$$

Similarly as in Section 2.5, our objective is to find a map $\Delta W : \mathbb{T}^d \times \mathbb{R} \to \mathbb{R}^{2n}$ and $\Delta\lambda \in \mathbb{R}$, such that for $\bar{W} = W + \Delta W$ and $\bar{\lambda} = \lambda + \Delta\lambda$, the new error $\bar{E}$ satisfies $\|\bar{E}\| = \mathcal{O}(\|E\|^2)$ for some suitable norm. More specifically, we will construct sequences for W and λ by means of an iterative scheme such that in the limit we have an exact solution of (4.3).

In the approximately invariant case, where the invariance equation is satisfied approximately, the range of W will be approximately isotropic. Consequently, the frame P will be approximately symplectic and the reducibility of the dynamics will only hold approximately. The error in these properties is actually controlled by the error in the invariance equation similarly as in Chapter 3. In the iterative scheme we can neglect these errors, as their contribution is at least of second order and still have a quadratically convergent procedure. We focus here in the algorithms and leave the rigorous treatment of the iterative process as well as convergence for future work.

We require that in the linearized invariance equation, the corrections are such that they cancel E. That is, we require

$$\mathrm{D}\phi_T \circ W \Delta W - \Delta W \circ \mathcal{R}_\omega^\lambda = -E + \partial_s W \circ \mathcal{R}_\omega^\lambda \; \Delta\lambda s.$$

If we express ΔW in the coordinates given by P, i.e., $\Delta W = P\xi$, and use (4.2.1), we obtain

$$\left(P\circ\mathcal{R}_\omega^\lambda\right)^{-1} \mathrm{D}\phi_T\circ W P\xi - \xi\circ\mathcal{R}_\omega^\lambda = \left(P\circ\mathcal{R}_\omega^\lambda\right)^{-1}\Big(- E + \partial_s W\circ\mathcal{R}_\omega^\lambda \Delta\lambda s\Big).$$

We can now use (4.12) and, after neglecting higher order terms, we obtain

$$\left(\begin{array}{c|c} \Lambda & S \\ \hline & \Lambda^{-\top} \end{array}\right)\xi - \xi\circ\mathcal{R}_\omega^\lambda = \eta + e_n\Delta\lambda s, \tag{4.16}$$

where

$$\eta := \Omega_0\left(P\circ\mathcal{R}_\omega^\lambda\right)^\top \Omega\circ W\circ\mathcal{R}_\omega^\lambda\ E, \quad e_n := \begin{pmatrix} 0_{n-1} \\ 1 \\ 0_{n-1} \\ 0 \end{pmatrix}. \tag{4.17}$$

If we split ξ and η into $(n-1)\times 1\times(n-1)\times 1$ components as

$$\xi = \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \\ \xi^4 \end{pmatrix}, \quad \eta = \begin{pmatrix} \eta^1 \\ \eta^2 \\ \eta^3 \\ \eta^4 \end{pmatrix}$$

and S into blocks of sizes $(n-1)\times(n-1), (n-1)\times 1, 1\times(n-1)$, and 1×1 as

$$S = \begin{pmatrix} S^1 & S^2 \\ S^3 & S^4 \end{pmatrix},$$

we obtain the following system of cohomological equations

$$\xi^1 + S^1\xi^3 + S^2\xi^4 - \xi^1\circ\mathcal{R}_\omega^\lambda = \eta^1 \tag{4.18}$$
$$\lambda\xi^2 + S^3\xi^3 + S^4\xi^4 - \xi^2\circ\mathcal{R}_\omega^\lambda = \eta^2 + \Delta\lambda s \tag{4.19}$$
$$\xi^3 - \xi^3\circ\mathcal{R}_\omega^\lambda = \eta^3 \tag{4.20}$$
$$\lambda^{-1}\xi^4 - \xi^4\circ\mathcal{R}_\omega^\lambda = \eta^4. \tag{4.21}$$

Let us choose Fourier-Taylor representations for ξ, S, and η. See Section 1.4. We can first solve for ξ^4 as described in Section 1.5. Note that we can solve for ξ^3 up to $\langle\xi_0^3\rangle$ if $\langle\eta_0^3\rangle = 0$. This average is actually quadratically small in E, see [HM21] for the proof, so it can be neglected in the iterative scheme. Once we have solved for ξ^4 and for ξ^3 with $\langle\xi_0^3\rangle$ free, we obtain the following equations for ξ^1 and ξ^2

$$\xi^1 - \xi^1\circ\mathcal{R}_\omega^\lambda = \delta^1 \tag{4.22}$$
$$\lambda\xi^2 - \xi^2\circ\mathcal{R}_\omega^\lambda = \delta^2 \tag{4.23}$$

where

$$\delta^1 := \eta^1 - S^2\xi^4 - S^1\xi^3 \tag{4.24}$$
$$\delta^2 := \eta^2 - S^4\xi^4 - S^3\xi^3 + \Delta\lambda s. \tag{4.25}$$

Note that in (4.22) we have an obstruction at order 0 and in (4.23) we have an obstruction at order 1, see Section 1.5. Therefore, in order to solve for ξ^1 and for ξ^2, we need that $\langle\delta_0^1\rangle = 0$ and $\langle\delta_1^2\rangle = 0$. We can use the freedom in $\langle\xi_0^3\rangle$ and the parameter $\Delta\lambda$ to adjust these averages. Let

$$\xi^3 = \tilde{\xi}^3 + \langle\xi_0^3\rangle, \tag{4.26}$$

where $\langle\tilde{\xi}_0^3\rangle = 0$, and let us define

$$\zeta^1 := \eta^1 - S^2\xi^4 - S^1\tilde{\xi}^3 \tag{4.27}$$
$$\zeta^2 := \eta^2 - S^4\xi^4 - S^3\tilde{\xi}^3. \tag{4.28}$$

We then choose $\langle\xi_0^3\rangle$ and $\Delta\lambda$ such that

$$\left(\begin{array}{c|c} \langle S_0^1\rangle & \\ \hline \langle S_1^3\rangle & -1 \end{array}\right)\begin{pmatrix}\langle\xi_0^3\rangle \\ \Delta\lambda\end{pmatrix} = \begin{pmatrix}\langle\zeta_0^1\rangle \\ \langle\zeta_1^2\rangle\end{pmatrix}, \tag{4.29}$$

given the non-degeneracy condition

$$\det\langle \mathrm{S}_0^1\rangle \neq 0. \tag{4.30}$$

Recall that (4.22) at order 0 and (4.23) at order 1 are small divisors cohomological equations—therefore, $\langle\xi_0^1\rangle$ and $\langle\xi_1^2\rangle$ are free. This underdeterminacy reflects the freedom in the choice of K, and in the choice of the first order approximation of $\mathcal{W}$, i.e., the freedom in the scaling of the invariant bundle of $\mathcal{K}$, see remark 4.4. A simple choice is to take $\langle\xi_0^1\rangle = 0$ and $\langle\xi_1^2\rangle = 0$. Then, we can solve for ξ^1 and ξ^2 as described in Section 1.5. Lastly, we obtain the corrected parameterization and the corrected rate of contraction as $\bar{W} = W + P\xi$ and $\bar{\lambda} = \lambda + \Delta\lambda$, respectively.

Remark 4.4. The parameterization W is not unique—we can scale the coordinate s by some scaling factor ρ and obtain the reparameterization $W_\rho(\theta, s) = W(\theta, \rho s)$. Using the Fourier-Taylor representation of Section 1.4, the coefficients of W_ρ are $W_{\rho_k}(\theta) = \rho^k W_k(\theta)$. Albeit, formally, all parameterizations are mathematically equivalent, for normalization purposes, we can choose an scaling factor such that the norms of the coefficients of the parameterizations neither grow nor decrease too fast with k. For such a purpose, we can take $\rho = 1/r$, where r the radius of convergence of the series in s that we can obtain with a root test.

Remark 4.5. Suppose we represent W as a Fourier-Taylor series, see Section 1.4, and assume we have, for instance, the torus and the bundle. That is, we have the terms W_0 and W_1. Similarly as in Chapter 3, it can be shown that the iterative procedure is quadratically convergent. Consequently, after it iterations we have $2^{it+1} - 1$ terms for W.

We conclude this section with algorithm 4.1 for one step of the iterative scheme.

Algorithm 4.1 Iterative step

Given W and λ satisfying (4.3) approximately, compute the corrected parameterization $\bar{W}$ and rate of contraction $\bar{\lambda}$ by following these steps:

1) Construct L, N, and P from (4.6), (4.10), and (4.9), respectively.
2) Compute S from to (4.2.1).
3) Compute the error in the invariance equation E from (4.15).
4) Construct η from the definition (4.17).
5) Solve (4.21) to obtain ξ^4.
6) Solve (4.20) to obtain $\tilde{\xi}^3$ with $\langle \tilde{\xi}^3_0 \rangle = 0$.
7) Compute ζ^1 and ζ^2 from (4.27) and (4.28), respectively.
8) Solve the linear system (4.29) given the non-degeneracy condition (4.30) in order to obtain $\langle \xi^3_0 \rangle$ and $\Delta\lambda$.
9) Construct ξ^3 from (4.26) and compute δ^1 and δ^2 from (4.24) and (4.25), respectively.
10) Set $\langle \xi^1_0 \rangle = 0$ and $\langle \xi^2_1 \rangle = 0$. Solve (4.22) and (4.23) to obtain ξ^1 and ξ^2.
11) Set $\bar{W} \leftarrow W + P\xi$ and $\bar{\lambda} \leftarrow \lambda + \Delta\lambda$.

4.1.4 Continuation with respect to T

Assume we have a ϕ_T invariant torus and its whisker. That is, we have (W, λ) satisfying (4.3). We are now interested in another solution (W^*, λ^*) of (4.3), defining a ϕ_{T^*} invariant whisker $\mathcal{W}^*$ for some $T^* \neq T$. As is standard in continuation methods, we provide in this section a method to compute the derivatives of (W, λ) with respect to T from where we can compute a first order approximation of (W^*, λ^*).

Let Eq. (4.3) define implicitly W and λ as functions of T. Then, differentiating (4.3) with respect to T, we obtain

$$\mathrm{D}_z\phi_T \circ W \partial_T W + X \circ W \circ \mathcal{R}^\lambda_\omega - \partial_T W \circ \mathcal{R}^\lambda_\omega - \mathrm{D}_s W \circ \mathcal{R}^\lambda_\omega \partial_T \lambda s = 0.$$

If we express $\partial_T W$ in the coordinates given by the frame P, i.e., $\partial_T W = P\xi$, and left-multiply the previous expression by $\left(P \circ \mathcal{R}^\lambda_\omega\right)^{-1}$, we obtain the following system of equations

$$\left(\begin{array}{c|c} \Lambda & S \\ \hline & \Lambda^{-\top} \end{array}\right) \xi - \xi \circ \mathcal{R}^\lambda_\omega = \begin{pmatrix} 0_{n-2} \\ -1 \\ \partial_T \lambda s \\ 0_n \end{pmatrix}. \tag{4.31}$$

This system is completely analogous to (4.16), but with

$$\eta = \begin{pmatrix} 0_{n-2} \\ -1 \\ 0_{n+1} \end{pmatrix} \tag{4.32}$$

and $\partial_T \lambda$ instead of $\Delta\lambda$. In this case, we have $\xi^4 = 0$ and $\xi^3 = \langle \xi^3_0 \rangle$. We can choose $\langle \xi^3_0 \rangle$ and $\partial_T \lambda$ to adjust averages for the equations for ξ^1 and ξ^2 to be solvable. That is, we choose $\langle \xi_0^3 \rangle$ and $\partial_T \lambda$ such that

$$\left(\begin{array}{c|c} \langle S_0^1 \rangle & \\ \hline \langle S_1^3 \rangle & -1 \end{array}\right) \begin{pmatrix} \langle \xi_0^3 \rangle \\ \partial_T \lambda \end{pmatrix} = \begin{pmatrix} -e_{n-1} \\ 0 \end{pmatrix}, \tag{4.33}$$

where e_{n-1} is as in (4.17). Then, we have no obstructions in the cohomological equations for ξ^1 and ξ^2 and we can proceed as in the previous section.

The computation of derivatives of (W, λ) with respect to T is summarized with algorithm 4.2.

Algorithm 4.2 Computation of derivatives with respect to T

Let $(W, \lambda)_T$ be implicit functions of T by equation (4.3). Find $\partial_T W$ and $\partial_T \lambda$ by following these steps:

1) Construct L, N, and P from (4.6), (4.10), and (4.9), respectively.
2) Compute S from (4.2.1).
3) Construct η from the definition (4.32).
4) Set $\xi^4 = 0$.
5) Solve the linear system (4.33) given the non-degeneracy condition (4.30) in order to obtain $\langle \xi_0^3 \rangle$ and $\partial_T \lambda$.
6) Set $\xi^3 = \langle \xi_0^3 \rangle$.
7) Compute δ^1 and δ^2 from (4.24) and (4.25), respectively, with $\partial_T \lambda$ instead of $\Delta\lambda$.
8) Set $\langle \xi_0^1 \rangle = 0$, $\langle \xi_1^2 \rangle = 0$, and solve (4.22) and (4.23) to obtain ξ^1 and ξ^2, respectively.
9) Set $\partial_T W = P\xi$.

4.2 Whiskers in quasi-periodic Hamiltonian systems

Let us now follow the quasi-periodic setting of Section 1.2. We will introduce some constructs that simplify the definition of whiskered tori in quasi-periodic systems. These constructs, besides providing simplicity, lead to interesting connections between partly hyperbolic invariant tori and hyperbolic operators.

Suppose we have K satisfying (2.3). Let us then define the *transfer matrix* $M : \mathbb{T}^{d+\ell} \to \mathbb{R}^{2n\times 2n}$ defined on the invariant torus $\mathcal{K}$ as

$$M(\theta, \varphi) = \mathrm{D}_z \phi_T\Big(K(\theta, \varphi), \varphi\Big).$$

In order to study the dynamics under iteration of M, for $n > 0$ let us define

$$M^{(n)} = M \circ \mathcal{R}_{\omega\alpha}^{\circ(n-1)} \dots M \circ \mathcal{R}_{\omega\alpha} M,$$

where

$$\mathcal{R}_{\omega\alpha}^{\circ n} := \underbrace{\mathcal{R}_{\omega\alpha} \circ \dots \circ \mathcal{R}_{\omega\alpha}}_{n \text{ times}}.$$

Observe that M satisfies the cocycle relation

$$M^{(n+m)} = M^{(n)} \circ \mathcal{R}_{\omega\alpha}^{\circ m} M^{(m)},$$

which we can use to define $M^{(-n)}$ as

$$M^{(-n)} = \left(M^{(n)} \circ \mathcal{R}_{\omega\alpha}^{\circ(-n)}\right)^{-1},$$

where

$$\mathcal{R}_{\omega\alpha}^{\circ(-n)} := \underbrace{R_{-\omega-\alpha} \circ \cdots \circ R_{-\omega-\alpha}}_{n \text{ times}}.$$

Definition 4.3. An invariant torus $\mathcal{K} = K(\mathbb{T}^{d+\ell})$ is said to be fiberwise whiskered if for every $(\theta, \varphi) \in \mathbb{T}^{d+\ell}$, $T_{K(\theta,\varphi)}U$ admits an invariant splitting

$$T_{K(\theta,\varphi)}U = E^s_{K(\theta,\varphi)} \oplus E^u_{K(\theta,\varphi)} \oplus E^c_{K(\theta,\varphi)}$$

such that there exist constants $C > 0, \lambda < 1$, and $\mu > 1$, satisfying $\lambda\mu < 1$ and

$$\begin{aligned}
v \in E^s_{K(\theta,\varphi)} &\iff |M^{(n)}(v)| \le C\lambda^n |v| \quad n > 0\\
v \in E^u_{K(\theta,\varphi)} &\iff |M^{(n)}(v)| \le C\lambda^{-n} |v| \quad n < 0\\
v \in E^c_{K(\theta,\varphi)} &\iff |M^{(n)}(v)| \le C\mu^{|n|} |v| \quad n \in \mathbb{Z}.
\end{aligned}$$

The transfer matrix M induces a *transfer operator* whose spectral properties are directly linked with the dynamical properties of the invariant torus $\mathcal{K}$. We will not follow such approach here but it is useful to acknowledge this relationship. For some references see e.g., [Mat68, SS74, HPS70, Mañ78, HCF+16].

4.2.1 Geometric properties and adapted frames for $\mathcal{W}$

Assume we have $W : \mathbb{T}^{d+\ell} \times \mathbb{R} \to U$ satisfying the invariance equation

$$\phi_T \circ (W, \mathrm{id}) - W \circ \mathcal{R}^\lambda_{\omega\alpha} = 0, \tag{4.34}$$

where

$$\begin{aligned}
\mathcal{R}^\lambda_{\omega\alpha} : \quad \mathbb{T}^{d+\ell} \times \mathbb{R} &\longrightarrow \mathbb{T}^{d+\ell} \times \mathbb{R}\\
(\theta, \varphi, s) &\longmapsto (\mathcal{R}_{\omega\alpha}(\theta, \varphi), \lambda s).
\end{aligned}$$

That is, we have a parameterization of a whiskered torus and its stable/unstable whisker $\mathcal{W}$ invariant under ϕ_T and with dynamics conjugated to $\mathcal{R}^\lambda_{\omega\alpha}$.

Similarly as in Section 4.1.2, we can construct adapted frames for $\mathcal{W}$ that reduce the linearized dynamics around $\mathcal{W}$ to upper triangular form. This is possible due to geometric properties analogous to those of Section 4.1.2 which in this case hold fiberwise.

For each $\varphi \in \mathbb{T}^\ell$, let us construct a subframe $L : \mathbb{T}^{d+\ell} \times \mathbb{R} \to \mathbb{R}^{2n\times n}$ invariant under the differential of ϕ_T. First, differentiate (4.34) with respect to the θ, φ, and s coordinates to obtain

$$\begin{aligned}
\mathrm{D}_z\phi_T \circ (W, \mathrm{id})\mathrm{D}_\theta W &= \mathrm{D}_\theta W \circ \mathcal{R}^\lambda_{\omega\alpha},\\
\mathrm{D}_z\phi_T \circ (W, \mathrm{id})\mathrm{D}_\varphi W + \partial_\varphi\phi_T \circ (W, \mathrm{id}) &= \mathrm{D}_\varphi W \circ \mathcal{R}^\lambda_{\omega\alpha},\\
\mathrm{D}_z\phi_T \circ (W, \mathrm{id})\mathrm{D}_s W &= \mathrm{D}_s W \circ \mathcal{R}^\lambda_{\omega\alpha}\lambda.
\end{aligned} \tag{4.35}$$

Observe that the bundles parameterized by $\mathrm{D}_\theta W$ and $\mathrm{D}_s W$ are invariant under $\mathrm{D}_z\phi_T$ on $\mathcal{W}$ whereas the bundle parameterized by $\mathrm{D}_\varphi W$ is not.

Similarly as for the construction of the subframe L for the torus, see Section 2.4.1, we can use (4.35) and (2.19) at $z = W(\theta, \varphi)$ to show that the following holds

$$\mathrm{D}_z\phi_T \circ (W, \mathrm{id})\mathcal{X} - \mathcal{X} \circ \mathcal{R}^\lambda_{\omega\alpha} = 0,$$

where

$$\mathcal{X} := X \circ (W, \mathrm{id}) - \mathrm{D}_\varphi W \hat{\alpha}.$$

Consequently, the subframe L constructed as

$$L := \begin{pmatrix} \mathrm{D}_\theta W & \mathcal{X} & \mathrm{D}_s W \end{pmatrix} \tag{4.36}$$

satisfies

$$\mathrm{D}_z\phi_T \circ (W, \mathrm{id})L - L \circ \mathcal{R}^\lambda_{\omega\alpha} \Lambda = 0.$$

Naturally, at $s = 0$, we recover (2.16).

In the quasi-periodic case, the isotropy of $\mathcal{W}$ holds fiberwise. That is, for each $\varphi \in \mathbb{T}^\ell$, the manifold parameterized by $W(\cdot, \varphi)$ is isotropic. Due to the fiberwise isotropy of $\mathcal{W}$, the subframe L is fiberwise Lagrangian. The proof is analogous to that of Proposition 4.2 which generalizes fiberwise. Therefore, we can proceed as in Section 4.1.2 and complement the subframe L with another fiberwise Lagrangian subframe N such that we obtain a symplectic frame satisfying the analogous condition to (4.11) and reducing the linearized dynamics as

$$(P \circ \mathcal{R}^\lambda_{\omega\alpha})^{-1} \mathrm{D}\phi_T \circ (W, \mathrm{id}) P = \left(\begin{array}{c|c} \Lambda & S \\ \hline & \Lambda^{-\top} \end{array}\right),$$

where the torsion $S : \mathbb{T}^{d+\ell} \times \mathbb{R} \to \mathbb{R}^{n \times n}$ is defined as

$$S = N^\top \circ \mathcal{R}^\lambda_{\omega\alpha} \; \Omega \circ W \circ \mathcal{R}^\lambda_{\omega\alpha} \; \mathrm{D}\phi_T \circ (W, \mathrm{id}) \; N.$$

An iterative procedure for the computation of whiskers for quasi-periodic Hamiltonians follows—*mutatis mutandis*—from Section 4.1.3. We include some numerical results in Section 4.6.

4.3 Fundamental domain

In this section we study the domain of validity of the parameterizations. We focus in the autonomous case which is the main case study for the numerical explorations of Section 4.5. Nonetheless, the same ideas apply to study the fundamental domains of parameterizations in the quasi-periodic case.

Assume we have a parameterization W given by a truncated Fourier-Taylor series at order N. The series W will only be valid in $\mathbb{T}^d \times \mathcal{D}$ for some open interval $\mathcal{D} \subset \mathbb{R}$—we refer to $\mathcal{D}$ as the fundamental domain of W. Let us consider the following norm for the error (4.15)

$$\|E\|_{\mathcal{D}} := \sup_{\substack{\theta \in \mathbb{T}^d \\ s \in \mathcal{D}}} \|E(\theta, s)\|_\infty. \tag{4.37}$$

Then, we can define $\mathcal{D}$ as the largest interval where

$$\|E\|_{\mathcal{D}} < \varepsilon, \tag{4.38}$$

holds for some $\varepsilon > 0$ small. If we proceed formally, we can construct a Fourier-Taylor series for E of the form

$$E(\theta, s) = \sum_{j=0}^{\infty} E_j(\theta) s^j. \tag{4.39}$$

Then, using the definition (4.37), we have the inequality

$$\|E\|_{\mathcal{D}} = \sup_{\substack{\theta \in \mathbb{T}^d \\ s \in \mathcal{D}}} \|\sum_{j=0}^{\infty} E_j(\theta) s^j\|_\infty \leq \sum_{j=0}^{\infty} \left(\sup_{\substack{\theta \in \mathbb{T}^d \\ s \in \mathcal{D}}} \|E_j(\theta) s^j\|_\infty \right)$$

and if we define

$$\|E_j\| := \sup_{\theta \in \mathbb{T}^d} \|E_j(\theta)\|_\infty, \quad D := \sup_{s \in \mathcal{D}} |s|, \tag{4.40}$$

we can estimate the fundamental domain by requiring

$$\|E\|_{\mathcal{D}} \leq \sum_{j=0}^{N} \|E_j\| D^j + \sum_{j=N+1}^{\infty} \|E_j\| D^j < \varepsilon. \tag{4.41}$$

Observe that if we want (4.41) to hold, necessarily for any k

$$\|E_k\| D^k < \varepsilon. \tag{4.42}$$

Let us now inspect which term contributes the most in the sums of (4.41) so we can obtain a useful bound for D from (4.42).

Observe that E_j, for $j \leq N$, should be small since these are the terms that we control with our iterative procedure. The terms for $j > N$ are terms of order j of $\phi_T \circ W$ since we do not have terms of order higher than N for W. Consequently, let us consider that the term that contributes the most in (4.41) is in the infinite sum. Note that the true radius of convergence for the series (4.39), say $\tilde{D}$, satisfies $D \leq \tilde{D}$ since we have

$$\|E\|_{\tilde{\mathcal{D}}} < \infty, \quad \|E\|_{\mathcal{D}} < \varepsilon,$$

where $\tilde{\mathcal{D}} = [-\tilde{D}, \tilde{D}]$. Furthermore, E is analytic so we can bound each E_j using Cauchy estimates by

$$\|E_j\| \leq \frac{C}{\tilde{D}^j}$$

for some constant C. Consequently, in the infinite sum, we can bound each term by

$$C \left(\frac{D}{\tilde{D}} \right)^j,$$

from where we see that the largest upper bound corresponds to the term $j = N + 1$. We can then obtain the following necessary condition for D

$$D < \left(\frac{\varepsilon}{\|E_{N+1}\|} \right)^{\frac{1}{N+1}}. \tag{4.43}$$

Alternatively, for some grid of values $\{\theta_j\}_{j\in\mathrm{I}}$, where I is some index set, and some partition of the interval $[-D, D]$, we can estimate $\|E\|_{\mathcal{D}}$ as

$$\|E\|_{\mathcal{D}} \approx \max_{\substack{j\in\mathrm{I}\\ 0\le i\le p}} \Big(\|\phi_T\Big(W(\theta_j, s_i)\Big) - W(\theta_j+\omega, \lambda s_i)\|_\infty\Big) \tag{4.44}$$

where p is some positive integer, $s_0 = -D$, $s_p = D$, and $s_i < s_{i+1}$. To estimate the fundamental domain, construct a sequence $\{D_i\}_{i=0}$ where $D_{i+1} = D_i + \Delta D$, for some ΔD small. Then, we obtain the estimate $\mathcal{D} = [-D_k, D_k]$ where $D_k \in \{D_i\}_{i=0}$ is the largest element in the sequence satisfying (4.38) for $\|E\|_{\mathcal{D}}$ estimated as in as in (4.44).

4.4 Comments on implementations

Throughout this section, assume $d = 1$—which is the case for the numerical experiments of Section 4.5—and that we have the following Fourier-Taylor representation

$$W(\theta, s) \sim \sum_{i=0}^{N} W_i(\theta)s^i \sim \sum_{i=0}^{N}\sum_{k\in\mathcal{I}} \hat{W}_{ik}e^{i2\pi k\theta}s^i, \tag{4.45}$$

for some finite N and where each W_i is given by N_F Fourier coefficients.

We follow [HM21, FMHM24b] and implement a low-pass filter after each iterative step of Algorithm 4.1. This is, for some filtering factor $r_f \in [\frac{1}{4}, \frac{1}{2})$, we set to zero the coefficients $\hat{W}_{ik}$ for $|k| > r_f \cdot N_F$. In order to decide the number of coefficients for the Fourier series, i.e., N_F, we control the exponential decay since our objects are real analytic. In particular, for any W_i, we compute the tail as

$$t(W_i) = \sum_{|k|>r_t\cdot N_F} \|\hat{W}_{ik}\|_\infty,$$

where $r_t < r_f$ is some tail factor. Then, we can decide to increase N_F if we detect the tails do not decrease sufficiently fast with k.

Recall that, according to Remark 4.4, we can scale the series in the coordinate s by its estimated radius of convergence for the sake of normalization. A possibility is to run algorithm 4.1 twice: the first to estimate the radius of convergence and the second, after the scaling has been done. However, if we are interested in high order parameterizations, say up to a maximum order N_{max}, estimating the radius of convergence after performing algorithm 4.1 to order N_{max} might not give good estimates. For this reason, we run algorithm 4.1 first to order $N_i < N_{max}$. Then we estimate the radius of convergence of the parameterization and scale the series. Then, we iterate at order $N_{i+1} > N_i$ until the radius of convergence is close to 1 or until $N_i = N_{max}$. See algorithm 4.3.

Furthermore, in order to mitigate the hyperbolicity, we follow [HM21, FMHM24b] and implement a multiple-shooting strategy. For some positive integer m, we look for whiskers $\{\mathcal{W}^i\}_{i=0}^{m-1}$ and parameterizations $\{W^i\}_{i=0}^{m-1}$, where $W^i : \mathbb{T}^d\times\mathbb{R} \to U$, such that the following system of functional equations is satisfied

$$\phi_{T_m}\circ W^i - W^{i+1}\circ\mathcal{R}_{\omega_m}^{\lambda_m} = 0, \quad \text{for } i = 0, \dots, m-1 \tag{4.46}$$

where the superscript of W is defined modulo m, $T_m := T/m$, $\omega_m := \omega/m$, and $\lambda_m := \lambda^{\frac{1}{m}}$. Assuming $|\lambda| < 1$, we choose m such that the expansion rate $|\lambda|^{\frac{-1}{m}}$ is small enough. For the sake of simplicity, we described the methodology for $m = 1$—the generalization to $m > 1$, albeit cumbersome, presents no additional complications.

Lastly, we parallelize the integration of jets, both for the flow and the variational equations. Our Jet Transport implementation substitutes the floating point operations of a RKF78 integrator by a simpler yet similar version of the truncated formal power series manipulator described in chapter 2 of [HCF+16]; which is publicly available. It does so by means of a C++ wrapper that overloads the arithmetic operators of a RKF78 C routine. The step size control performed is Fehlbergh's but taking into account all the coefficients of the jet. This is justified by the fact that, assuming analyticity, a shot calculation shows that the error in all the coefficients of the jet is of the order of the numerical integrator in the step size. Reference [GJJC+23] shows actually more: the jet of coefficients obtained through, e.g., Taylor and Runge-Kutta methods is the same (up to rounding errors) as the one that would be obtained by integrating the variational equations up to the order of the jets. For publicly available implementations see [JZ05, PHB19]

We conclude this section with a proposal for a practical implementation of the iterative scheme of Section 4.1.3 in Algorithm 4.3.

Algorithm 4.3 Practical implementation

Let λ be some approximate rate of contraction and let W_0 and W_1 be zero and first order terms of W satisfying (2.6) approximately for some fixed T and rotation vector ω. Let ε_t, ε_W, $\tilde{\varepsilon}$ be tolerances, n_t, n_{it} maximum number of iterates, $\tilde{r} \in (0, 1)$ a radius of convergence, ΔN an integer increment, N_i an initial order for scaling, N a desired order, and $N_{F_{max}}$a maximum number of Fourier coefficients. Compute λ and W at order N by following these steps:

1) Compute the tails $t(W_0)$ and $t(W_1)$. If $\max\{t(W_0),\, t(W_1)\} > \varepsilon_t$ set $N_F \leftarrow \min\{2N_F,\, N_{F_{max}}\}$. Try up to n_t times.
2) Perform Algorithm 4.1 at order N_i until $\|E_{N_i}\| < \varepsilon_W$ or up to n_{it} times.
3) Estimate the radius of convergence as $r \approx \|E_{N_i}\|^{-1/N_i}$.
4) If $r < \tilde{r}$ or $r > 1$, scale the series for W by $1/r$. Else, go to step 6.
5) If $N_i = N$, go to step 6. Else, set $N_i \leftarrow \min\{N_i + \Delta N,\, N\}$ and go to step 2.
6) Perform Algorithm 4.1 at order N until $\|E_N\| < \tilde{\varepsilon}$ or up to n_{it} times. Obtain W at order N and λ.

4.5 Numerical experiments: whiskers in the CRTBP

We test our algorithms in the Hamiltonian vector field of the CRTBP. That is, we have a vector field generated by the gravitational attraction of two massive bodies, known as primaries, moving in circular orbits around their common barycenter. Let m_1 and m_2 denote the masses of the primaries and let us also define the mass parameter $\mu := \frac{m_2}{m_1+m_2}$. For a suitable rescaling in space and time, we can define a rotating frame, also known as a synodic frame, such that the position of the primaries is fixed in x_1 axis, see e.g. [Vic67]. The coordinates of the primaries are then $(\mu, 0, 0)$ and $(\mu - 1, 0, 0)$ and their period of revolution is 2π. The motion of the infinitesimal body is then described by autonomous

Hamiltonian

$$H(x,p) = \frac{1}{2}(p_1^2 + p_2^2 + p_3^2) - x_1 p_2 + x_2 p_1 - \frac{1-\mu}{r_1} - \frac{\mu}{r_2},$$

where $x, p \in \mathbb{R}^3$, $r_1^2 := (x_1 - \mu)^2 + x_2^2 + x_3^2$, and $r_2^2 := (x_1 - \mu + 1)^2 + x_2^2 + x_3^2$.

The CRTBP has five fixed points denoted by L_i with $i = 1, 2, \ldots, 5$. We focus on the L_1 point, which is the one located between the primaries, for the Earth-Moon systems. That is, for the value of the parameter $\mu = 1.215058560962404 \cdot 10^{-2}$. For some qualitative information on the CRTBP, see Section 2.8.2.

4.5.1 The Lissajous family

In order to identify each torus of the Lissajous family, we follow [GM01, HM21] and define the rotation number, denoted by ρ, as in Section 2.8.2.

For the computation of the Lissajous family of whiskered tori and their whiskers, we have two main options: to perform continuations in T for families with fixed ρ following Section 4.1.4 or to compute tori from approximate solutions of (4.3) following the iterative scheme of Section 4.1.3. Since we are interested in high order expansions, we choose the later approach because running continuations in T—for large step sizes—is only effective at low orders. In other words, keeping high order terms for the parameterizations does not help convergence since for large step sizes the error in the high order terms is large. It is more effective to perform continuations at low order and then increase the order with algorithm 4.1. We have run some tests which show that for orders say $N \leq 10$, the continuations in T are effective for moderate step sizes.

For the second approach, we need approximate parameterizations. To this end, we follow the methods of [HM21] and compute parameterizations of invariant tori and their invariant bundles for the Lissajous family—that is, we have the terms W_0 and W_1 of the Fourier-Taylor expansion of $\mathcal{W}$. The tori obtained are included in Figure 4.1, where the color bar labels the number of Fourier coefficients used. For the tori represented in the Figure, the rate of expansion is roughly between 500 and 3000.

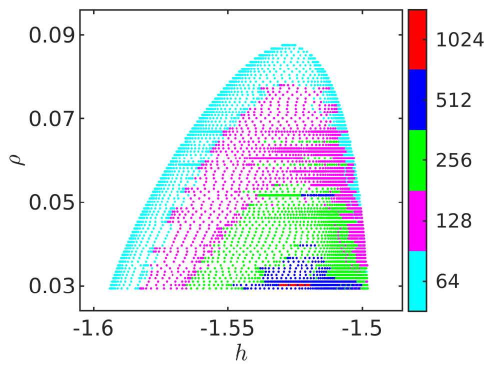

Figure 4.1: Energy-rotation number representation of the Lissajous tori around the L_1 point computed following the approach of [HM21]. The colorbar labels the number of Fourier coefficients of the parameterizations.

Remark 4.6. We emphasize that albeit in Figure 4.1 we represent the total number of Fourier coefficients used to compute tori and their bundles, the effective number of coefficients is lower than that represented in the Figure because of the low-pass filter implemented to prevent numerical instabilities, see Section 4.4.

4.5.2 Computation of Lissajous whiskers

In this section, we apply Algorithm 4.3 to each torus obtained for Figure 4.1. Recall that in the previous section, we obtained the terms W_0 and W_1 that we can use as seed for the iterative procedure. We first compute parameterizations up to order $N = 10, 12, 16, 20, 50$ with $\varepsilon_t = 10^{-9},\ \varepsilon_W = \tilde{\varepsilon} = 10^{-4},\ n_t = 1,\ n_{it} = 7,\ \tilde{r} = 0.9,\ \Delta N = 10$, and $N_{F_{max}} = 1024$. See Algorithm 4.3. For the multiple-shooting strategy, we take $m = 4$, see (4.46).

In order to asses the quality of the parameterizations, we construct an observable d_{max} to measure pointwise the maximum distance in phase space (up to the discretization) between tori and the limit of validity of their whiskers parameterizations in the corresponding fiber. That is, we consider

$$d_{max} = \max_{\substack{j \in \mathrm{I} \\ 0 \leq i \leq p}} \|W(\theta_j, s_i) - W(\theta_j, 0)\|_2, \tag{4.47}$$

where $\{\theta\}_{j\in \mathrm{I}}$ is an equally spaced grid of $\mathbb{T}^d$, I is some finite multi-index set with elements in $\mathbb{N}^d$, p is some positive integer, $s_0 = -D, s_p = D,\ s_i < s_{i+1}$, and $\mathcal{D} = [-D, D]$ the fundamental domain of W. For the estimates of the fundamental domain, we tested both of the methods described in Section 4.3. That is, using the upper bound (4.43) and by requiring (4.38) to hold using (4.44). Both for $\varepsilon = 10^{-7}$. We obtained very similar results being the estimate provided by (4.43) slightly larger. This shows that the main source of error is the truncation error which we cannot control. In what follows, we use criterion (4.38) with the approximation (4.44) to estimate the fundamental domain.

We include in Figure 4.2 the observable d_{max} in the energy-rotation number representation. We have also computed the observable d_{max} when we use the linear approximation of the whiskers, i.e., when $N = 1$. The results are not included in the Figure since in this case d_{max} is $\mathcal{O}(10^{-4})$. We emphasize that d_{max} reflects the validity of the parameterizations in phase space whereas D reflects the validity in the coordinates of the parameterizations—which depend on the scaling of W_1. We can observe that increasing the order of the parameterizations from $N = 10$ to $N = 30$ results in larger domains for almost all tori. However, if we increase the order to $N = 50$, there are tori for which we obtain larger domains but, in certain region of the energy-rotation number representation, we obtain worse parameterizations. In other words, increasing the order of the parameterizations does not always result in larger fundamental domains. It is then clear that if, we want to maximize the fundamental domains for each tori, we need to choose suitable orders for each parameterization.

Let us now explore individual truncation orders and set a maximum order for the parameterizations $N_{max} = 50$. For the tori obtained at order 50 in Figure 4.2, we truncate the series as follows. For each torus, we consider Fourier-Taylor representations for the error,

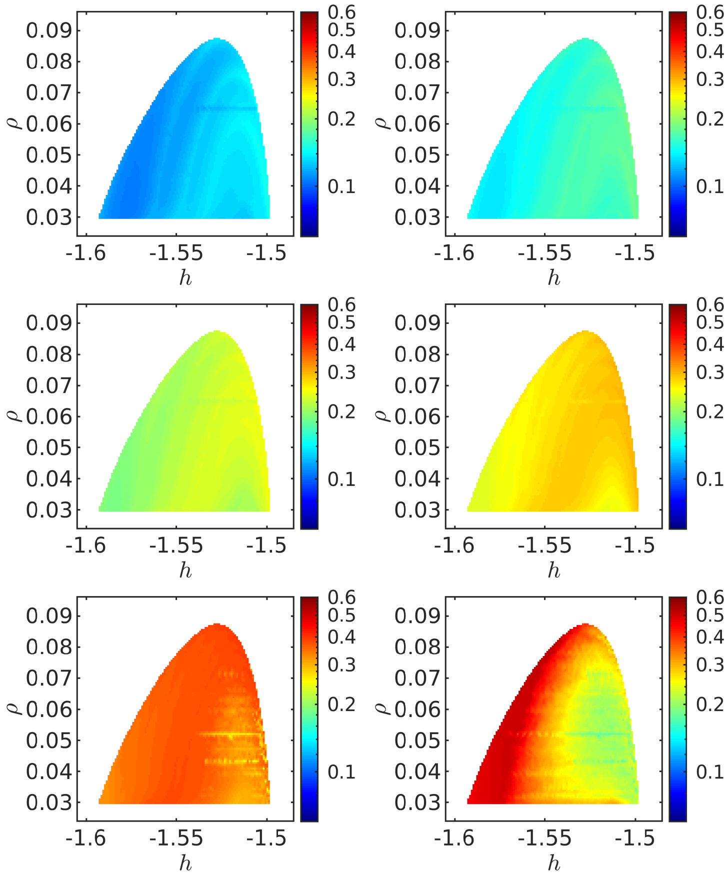

Figure 4.2: Observable d_{max} for fixed orders (from left to right) $N = 10, 12, 16, 20, 30, 50$. Colorbar in logarithmic scale.

defined in (4.15), at order N_{max}. That is, we consider

$$E(\theta, s) \sim \sum_{i=0}^{N_{max}} E_i(\theta) s^i.$$

Then, we truncate the parameterizations for each torus at the largest order $N \leq N_{max}$ for which

$$\|E_j\| < \text{tol} \tag{4.48}$$

holds for all $j \leq N$ and some tol > 0. We run tests for tol $= 10^{-4}$, 1 and include the orders obtained in Figure 4.3 and the observable d_{max} in Figure 4.4. Note that in the truncated series for E, the term E_j contributes tot the error as $\|E_j\| D^j$ where $D < 1$.

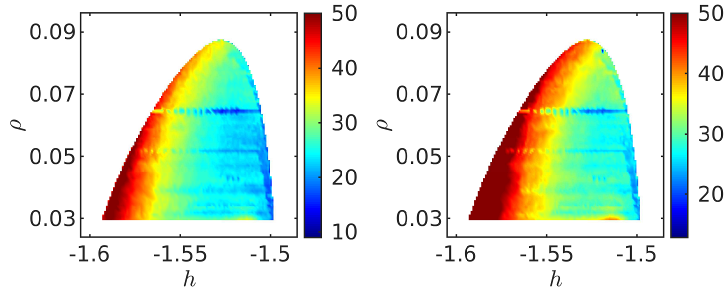

Figure 4.3: Truncation order for tol $= 10^{-4}$ (left) and tol $= 1$ (right).

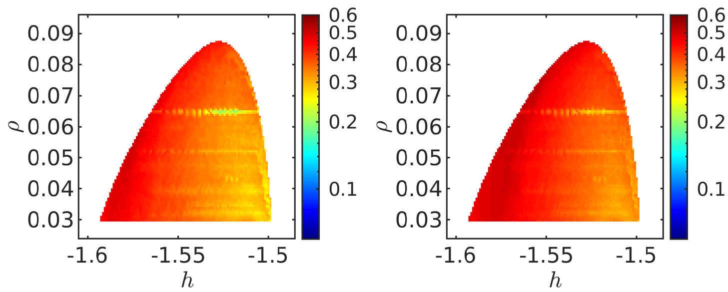

Figure 4.4: Observable d_{max} for *tol* $= 10^{-4}$ (left) and *tol* $= 1$ (right). Colorbar in logarithmic scale.

We can observe that, with this approach, we improve the results obtained in Figure 4.2: for most tori, we obtain values of the observable d_{max} larger that the obtained for the fixed order results of Figure 4.2. This shows that for each torus in the energy-rotation number representation requires different number of Fourier coefficients and different truncation orders in their Fourier-Taylor parameterizations. Recall that the number of Fourier coefficients was determined by the seed that we used for the computation of whiskers coming from [HM21] and from their tails, see Algorithm 4.3.

As a last experiment, we proceed to compute the optimal orders of the parameterizations via direct computation. More specifically, we use the parameterizations obtained in Figure 4.2 at fixed order $N_{max} = 50$. Then, for some $N \leq N_{max}$ and $j = 1, \ldots, N$ we construct increasing sequences for the fundamental domains in phase space of truncated parameterizations at order j according to Section 4.3. That is, we construct the sequences $\{d^j_{max}\}_{1\leq j\leq N}$ where N is the largest integer for which $\{d^j_{max}\}_{1\leq j\leq N}$ is strictly increasing. Then, we take such N as the order that optimizes the fundamental domain of the parameterizations. The results are included in Figure 4.5, where we can observe that the results are rather similar to those obtained in Figures 4.3-4.4 using criterion (4.48) for tol $= 1$. Lastly, we include a contour plot in Figure 4.6 for the observable d_{max} obtained with criterion (4.48) and via direct computation. It is then clear that criterion (4.48) with tol $= 1$ is a suitable choice for choosing truncation orders that maximize the fundamental domains in phase space. It is worth mentioning that the computation of the optimal orders via direct computation cannot deal with folds since we are only constructing the sequences $\{d^j_{max}\}_{1\leq j\leq N}$. We do not, in general, expect folds locally close to the invariant tori. If necessary, one could instead inspect the computationally more expensive sequences $\{d^j_{max}\}_{1\leq j\leq N_{max}}$ if folds are expected.

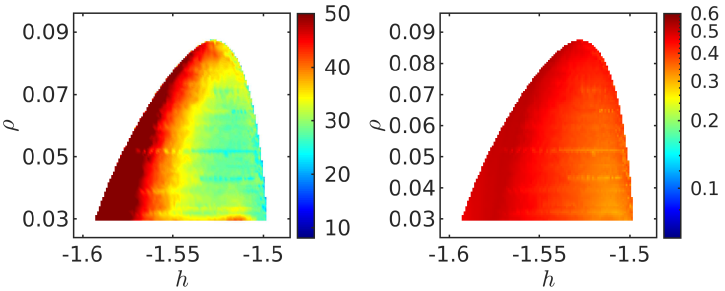

Figure 4.5: Optimal orders (left) and maximum distance (right) obtained via direct computation.

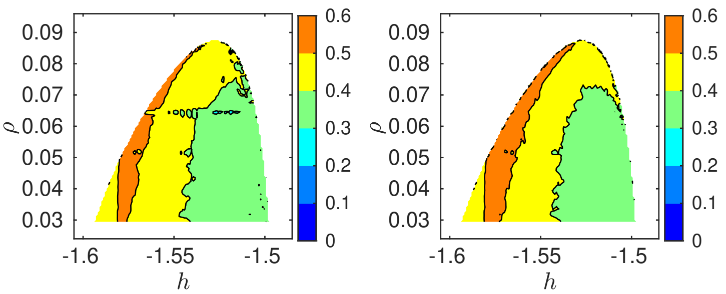

Figure 4.6: Observable d_{max} obtained from criterion (4.48) for $tol = 1$ (left) and optimal distance obtained via direct computation(right).

We end this section with an illustration of the computing times of the procedure. We run Algorithm 4.3 but without scaling the bundle, i.e., we run only the steps 1) and 2) at $N_i = N$. In particular, we take the torus labeled by $h = -1.5319$ and $\rho = 0.0723$ and compute its stable whisker for different numbers of Fourier coefficients and at different orders, see Table 4.1. We also include the error of invariance in the torus, bundle, the last term of the Fourier-Taylor representation, the limit of the fundamental domain in the coordinates of the parameterization (denoted by D), and the observable d_{max}. For the specifications of the machine used for the computations see the second paragraph of Section 2.8.

We first observe that when for $N_F = 64$, we do not have enough Fourier coefficients to obtain parameterizations satisfying the invariance equation with an error smaller than $\mathcal{O}(10^{-8})$. Also, note that while $\|E_N\|$ grows with N, its contribution to the error is $\|E_N\| D^N$. Therefore, large values of $\|E_N\|$ do not result in large errors. It simply implies that D needs to be sufficiently small to guarantee an accurate description of the parameterizations in the fundamental domain of W. Also, we have observed that the computational time grows linearly with N_F. This is not surprising since a profiling of the code reveals that 98.4% of the total time is numerical integration. Regarding the computing time behaviour with the order, from an operation count standpoint, a quadratic increase of computing time with order should be expected. Profiling tests have also shown that quadratic growth is not observed in Table 4.1 because of the computational effort required by the book-keeping and memory management related to the manipulation of power series in the Jet Transport implementation.

Table 4.1: Computational times

N_F	N	$\|E_0\|$	$\|E_1\|$	$\|E_N\|$	D	d_{max}	it	t
64	10	$2.3 \cdot 10^{-8}$	$6.5 \cdot 10^{-8}$	$2.6 \cdot 10^{-6}$	0.333	0.122	5	2.68s
	12	$2.3 \cdot 10^{-8}$	$6.5 \cdot 10^{-8}$	$3.5 \cdot 10^{-6}$	0.424	0.162	5	2.76s
	16	$2.3 \cdot 10^{-8}$	$6.5 \cdot 10^{-8}$	$2.8 \cdot 10^{-5}$	0.534	0.214	5	3.05s
	20	$2.3 \cdot 10^{-8}$	$6.5 \cdot 10^{-8}$	$4.7 \cdot 10^{-6}$	0.567	0.231	6	3.85s
	30	$2.3 \cdot 10^{-8}$	$6.5 \cdot 10^{-8}$	$3.4 \cdot 10^{-1}$	0.552	0.223	7	5.49s
	50	$2.3 \cdot 10^{-8}$	$6.5 \cdot 10^{-8}$	$2.0 \cdot 10^{9}$	0.457	0.177	7	7.15s
256	10	$1.4 \cdot 10^{-14}$	$1.1 \cdot 10^{-14}$	$1.1 \cdot 10^{-13}$	0.335	0.123	5	9.90s
	12	$1.4 \cdot 10^{-14}$	$1.1 \cdot 10^{-14}$	$1.3 \cdot 10^{-12}$	0.427	0.163	5	10.25s
	16	$1.4 \cdot 10^{-14}$	$1.1 \cdot 10^{-14}$	$1.2 \cdot 10^{-9}$	0.557	0.226	5	11.02s
	20	$1.6 \cdot 10^{-14}$	$1.3 \cdot 10^{-14}$	$1.0 \cdot 10^{-13}$	0.657	0.279	6	14.25s
	30	$1.6 \cdot 10^{-14}$	$1.3 \cdot 10^{-14}$	$1.4 \cdot 10^{-8}$	0.839	0.383	6	16.02s
	50	$1.6 \cdot 10^{-14}$	$1.3 \cdot 10^{-14}$	$3.7 \cdot 10^{-6}$	0.692	0.298	6	19.30s
1024	10	$1.5 \cdot 10^{-14}$	$1.3 \cdot 10^{-14}$	$1.6 \cdot 10^{-12}$	0.335	0.123	5	37.99s
	12	$1.5 \cdot 10^{-14}$	$1.3 \cdot 10^{-14}$	$2.0 \cdot 10^{-12}$	0.427	0.163	5	39.60s
	16	$1.5 \cdot 10^{-14}$	$1.3 \cdot 10^{-14}$	$1.3 \cdot 10^{-8}$	0.557	0.226	5	42.07s
	20	$1.4 \cdot 10^{-14}$	$1.3 \cdot 10^{-14}$	$1.4 \cdot 10^{-13}$	0.657	0.279	6	54.78s
	30	$1.4 \cdot 10^{-14}$	$1.3 \cdot 10^{-14}$	$1.5 \cdot 10^{-8}$	0.839	0.383	6	62.88s
	50	$1.4 \cdot 10^{-14}$	$1.3 \cdot 10^{-14}$	$3.8 \cdot 10^{-6}$	0.646	0.273	6	76.07s

4.6 Numerical experiments: whiskers in the ERTBP

In this Section, we explore the computation of whiskered tori in the periodic Hamiltonian of the Earth-Moon ERTBP, i.e., for parameters $\mu = 1.215058560962404 \cdot 10^{-2}$ and $e = 0.0549$. For some details on the dynamical system see Section 2.8.1.

Our approach is the following. First, we compute invariant tori and their invariant bundles as described in Chapter 2. Then, we use the methods of this chapter for quasi-periodic Hamiltonian systems outlined in Section 4.2 to compute invariant tori and their whiskers with Algorithm 4.3 adapted to the quasi-periodic case. In this case, we limit ourselves to the computation of a few whiskers in the ERTBP and leave a systematic exploration of the phase space for future work. In particular, we have taken two tori in the Earth-Moon CRTBP labeled with different values of the Hamiltonian h and the same value of $\rho = 0.0644$, see Section 2.8.2. We include in Table 4.2 the initial and final number of Fourier coefficients, the label h, the rate of expansion, the number of iterative steps, the norms of the first and last coefficients of the Fourier-Taylor series of the error of invariance, the observable d_{max} and the computing time. As in the previous section, for the computing time we do not scale the bundle, i.e., we run only steps 1) and 2) at order $N = 30$. Furthermore, simply for illustration purposes, we plot in Fig. 4.7 in red the generated 3−dimensional invariant tori and in blue the whiskers of two out of the four 2−dimensional generators. Recall that we use a multiple-shooting strategy, see Section 4.4.

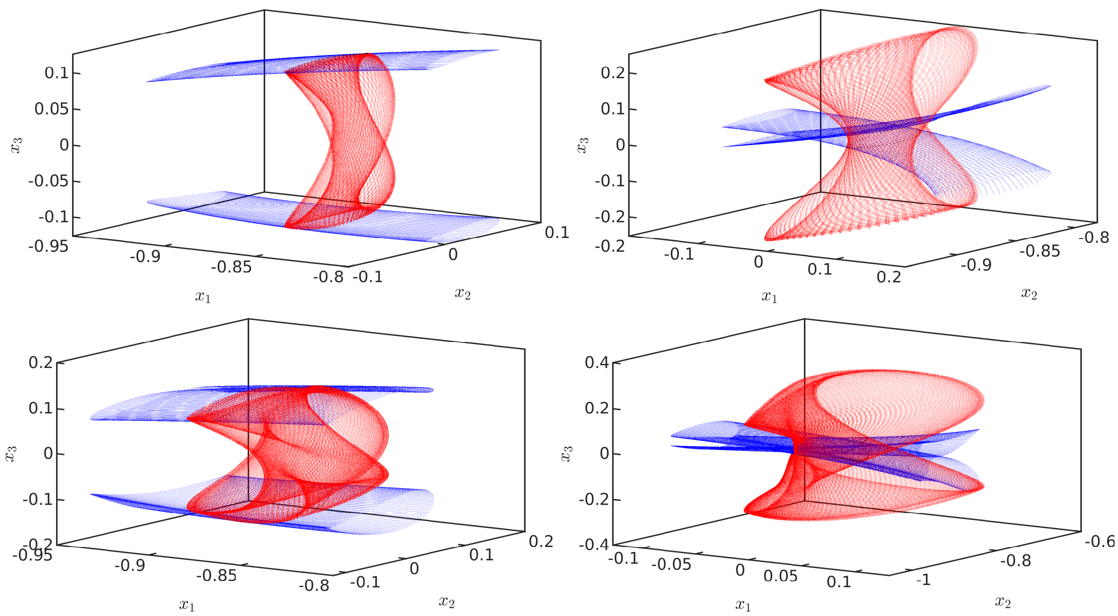

Figure 4.7: First column, projections into configuration space. Second column, projections into momenta space. First row, objects labeled by $h = -1.56378$. Second row, objects labeled by $h = -1.55032$. In red generated 3−dimensional tori and in blue whiskers of two 2−dimensinal generators.

Table 4.2: Computational times quasi-periodic whiskers

N_{F_0}	N_{F_f}	h	λ	steps	$\|E_K\|$	$\|E_{30}\|$	d_{max}	t
$(64, 16)$	$(64, 32)$	-1.56378	1909.58	5	$1.1 \cdot 10^{-13}$	$3.2 \cdot 10^{-9}$	0.379	2m 39s
$(128, 32)$	$(128, 32)$	-1.55032	1368.66	6	$5.3 \cdot 10^{-12}$	$1.2 \cdot 10^{-8}$	0.380	5m 32s

Conclusions

In Chapter 2, we designed KAM iterative schemes for quasi-periodic Hamiltonian systems with an arbitrary number of frequencies. These schemes are formulated using time-T maps, where T corresponds to the period associated with one of the internal frequencies of the torus.

The benefits of this approach are twofold: it enables a convenient and straightforward continuation from autonomous to quasi-periodic systems and it reduces the number of coefficients required for the parameterizations. The latter can be attributed to the following:

- The torus is more analytic in the temporal direction.
- The time dependency acts as a perturbation of an autonomous system. Then, if the perturbation is small, the effect of the time-dependency is small and fewer coefficients are necessary.

Overall, this approach is advantageous not only in terms of computational cost but also because it provides the freedom to choose which frequency to remove. This can be beneficial, for instance, for tori with fast and slow frequencies. We also designed continuation schemes using geometric properties such that the prediction step is done as efficiently as the KAM iterates. For the numerical experiments, we explored a large subset of the phase space computing $3-$dimensional (non-resonant) Lissajous tori and quasi-halo tori. The material of this chapter has been published in [FMHM24b].

In Chapter 3, we obtained an a-posteriori theorem for the methods of Chapter 2. This theorem is suitable for conducting computer-assisted proofs on the existence of invariant tori and their associated invariant bundles. By means of an extra symplectic transformation, we obtain both the stable and unstable bundles. Moreover, since we have taken periods associated with the internal frequencies of the torus, the theorem also applies to autonomous Hamiltonian systems. In this case, it is enough to set certain quantities to zero in the constants of Appendix E. For the proof of the theorem, we proved a lemma on quadratically small averages that allows us to prove convergence of the KAM scheme without using a translated curve theorem and a vanishing lemma. We also introduced fiberwise symplectic deformations and moment maps that allow us to prove a zero average lemma that justify the computation of derivatives of the parameterizations with respect to parameters. This allows better predictions for the continuation of the invariant objects with respect to parameters of the Hamiltonian. The material of this chapter have been collected in the preprint [FMHM24a].

In Chapter 4, we designed KAM schemes to compute high order expansions of the stable/unstable manifolds of invariant tori in autonomous and quasi-periodic Hamiltonian systems. The scheme is based on an iterative method that solves for every term of the expansion simultaneously as opposed to the order by order methods. This means that if we have the torus and the bundle exactly, after *it* iterations of the scheme we have parameterizations of the whiskers at order $2^{it+1} - 1$. Moreover, we include strategies for continuation with respect to parameters following the spirit of Chapter 2 and some strategies to estimate the fundamental domain of the parameterizations. For the numerical explorations, we performed massive computations in the CRTBP by computing Lissajous $2-$dimensional tori and their whiskers at different orders. We attempted to adapt the parameterizations of each torus such that the fundamental domain of the parameterizations are as large as possible. Lastly, we computed some $3-$dimensional tori and their whiskers in the ERTBP. The contents of this chapter have been collected in the preprint [FMHdlLM24].

The framework, algorithms, and theorem obtained provide an excellent starting point for future projects. In particular, the algorithms have been formulated for an arbitrary number of frequencies of the Hamiltonian. Therefore, we have the tools to compute invariant tori in models with more frequencies like the frequency models of [GMM02]. Future lines of research include the computation of homoclinic and heteroclinic connections between invariant tori in quasi-periodic Hamiltonian systems. These connections can serve in the design of space missions trajectories to conduct complex itineraries. Furthermore, in the methods of the thesis, we reduced the dynamics in the invariant bundles to constant contraction/expansion. Whilst effective, this comes at the cost of an additional analyticity loss at each iterative step—which can be avoided by generalizing the methods hereby described for invariant tori without reducing the dynamics in the bundles. Additionally, designing KAM schemes for tori with bundles of arbitrary rank, extending the a-posteriori theorem to this case, or obtaining sharp theorems in the spirit of [Vil17, FH24] are promising directions for future research.

Appendix A

On the primitive function of ϕ_t

We dedicate this section to the explicit expression of the primitive function of the uniparametric $t-$family of fiberwise exact symplectomorphisms $\phi_t : U \times \mathbb{T}^\ell \to U$, for $U \subset \mathbb{R}^{2n}$ open and $t \in \mathbb{R}$.

Lemma A.1. *The primitive function $p_t : U \times \mathbb{T}^\ell \to \mathbb{R}$ of the fiberwise exact symplectomorphism $\phi_t : U \times \mathbb{T}^\ell \to U$ is given by*

$$p_t = \int_0^t \Big(a^\top \circ \phi_s \; X \circ (\phi_s, \psi_s) - H \circ (\phi_s, \psi_s) \Big) \, ds, \tag{A.1}$$

where recall that $\psi_s(\varphi) = \varphi + \hat{\alpha} s$. That is, the primitive function p_t, as given above, satisfies

$$\mathrm{D}_z p_t = a^\top \circ \phi_t \; \mathrm{D}_z \phi_T - a^\top. \tag{A.2}$$

Moreover,

$$\mathrm{D}_\varphi p_t = a^\top \circ \phi_t \; \mathrm{D}_\varphi \phi_t - \int_0^t \mathrm{D}_\varphi H \circ (\phi_s, \psi_s) ds. \tag{A.3}$$

Proof. In order to prove p_t is as given in (A.1), it suffices to show that p_t satisfies (A.2). Let us begin by differentiating p_t with respect to z

$$\begin{aligned}\mathrm{D}_z p_t = \int_0^t \Big(& X^\top \circ (\phi_s, \psi_s) \mathrm{D} a \circ \phi_s \mathrm{D}_z \phi_s + a^\top \circ \phi_s \mathrm{D}_z X \circ (\phi_s, \psi_s) \mathrm{D}_z \phi_s \\ & - \mathrm{D}_z H \circ (\phi_s, \psi_s) \mathrm{D}_z \phi_s \Big) ds.\end{aligned}$$

We use that $\mathrm{D}_z H = -X^\top \Omega$ and that $\Omega = (\mathrm{D}a)^\top - \mathrm{D}a$ to obtain

$$\begin{aligned}\mathrm{D}_z p_t &= \int_0^t \Big(X^\top \circ (\phi_s, \psi_s) (\mathrm{D}a)^\top \circ \phi_s \mathrm{D}_z \phi_s + a^\top \circ \phi_s \mathrm{D}_z X \circ (\phi_s, \psi_s) \mathrm{D}_z \phi_s \Big) ds \\ &= \int_0^t \left(\frac{d}{ds} (a \circ \phi_s)^\top \mathrm{D}_z \phi_s + a^\top \circ \phi_s \frac{d}{ds} \mathrm{D}_z \phi_s \right) ds \\ &= \int_0^t \frac{d}{ds} \Big(a^\top \circ \phi_s \mathrm{D}_z \phi_s \Big) ds \\ &= a^\top \circ \phi_t \mathrm{D}_z \phi_t - a.\end{aligned}$$

Let us now differentiate p_t with respect to φ

$$\begin{aligned}\mathrm{D}_\varphi p_t = \int_0^t \Big(& X^\top \circ (\phi_s, \psi_s) \mathrm{D}a \circ \phi_s \mathrm{D}_\varphi \phi_s + a^\top \circ \phi_s \mathrm{D}_z X \circ (\phi_s, \psi_s) \mathrm{D}_\varphi \phi_s \\ & + a^\top \circ \phi_s \mathrm{D}_\varphi X \circ (\phi_s, \psi_s) - \mathrm{D}_z H \circ (\phi_s, \psi_s) \mathrm{D}_\varphi \phi_s - \mathrm{D}_\varphi H \circ (\phi_s, \psi_s) \Big) ds,\end{aligned}$$

which we can rewrite as

$$\begin{aligned}\mathrm{D}_\varphi p_t &= \int_0^t \Big(X^\top \circ (\phi_s, \psi_s) (\mathrm{D}a)^\top \circ \phi_s \mathrm{D}_\varphi \phi_s + a^\top \circ \phi_s \frac{d}{ds} \mathrm{D}_\varphi \phi_s - \mathrm{D}_\varphi H \circ (\phi_s, \psi_s) \Big) ds \\ &= \int_0^t \left(\frac{d}{ds} \Big(a^\top \circ \phi_s \mathrm{D}_\varphi \phi_s \Big) - \mathrm{D}_\varphi H \circ (\phi_s, \psi_s) \right) ds \\ &= a^\top \circ \phi_t \mathrm{D}_\varphi \phi_t - \int_0^t \mathrm{D}_\varphi H \circ (\phi_s, \psi_s) \; ds.\end{aligned}$$

$\square$

Appendix B

Quadratically small averages

For the iterative step described in Section 2.5.1 to be consistent, we need the averages of η^3 given by

$$\eta^3 = \begin{pmatrix} (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha} E_K \\ \mathcal{X}^\top \circ \mathcal{R}_{\omega\alpha}\, \Omega \circ K \circ \mathcal{R}_{\omega\alpha} E_K \end{pmatrix} =: \begin{pmatrix} \eta^{31} \\ \eta^{32} \end{pmatrix}$$

to be quadratically small with respect to the error in the torus invariance E_K. Note that for some suitable norm, the derivatives of E_K can be controlled by $\|E_K\|$ using Cauchy estimates. The following lemma provides the explicit formulas for $\langle \eta^3 \rangle$.

Lemma B.1. *The averages of η^{31} and η^{32} are given by*

$$\begin{aligned} \left\langle \eta^{31} \right\rangle &= \left\langle (\mathrm{D}_\theta E_K)^\top \Delta^1 a + (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \Delta^2 a \right\rangle \\ \left\langle \eta^{32} \right\rangle &= \left\langle \Delta^2 H \right\rangle - \left\langle \left(\Delta^1 a\right)^\top \mathrm{D}_\varphi E_K \hat{\alpha} \right\rangle - \left\langle \left(\Delta^2 a\right)^\top \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha}\, \hat{\alpha} \right\rangle, \end{aligned}$$

where the Taylor remainders Δ^i of order i in E_K are given by

$$\begin{aligned} \Delta^1 a :&= a \circ \phi_T \circ (K, \mathrm{id}) - a \circ K \circ \mathcal{R}_{\omega\alpha} \\ &= \int_0^1 \mathrm{D}a\Big(K \circ \mathcal{R}_{\omega\alpha} + s E_K\Big) E_K \; ds, \\ \Delta^2 a :&= a \circ \phi_T \circ (K, \mathrm{id}) - a \circ K \circ \mathcal{R}_{\omega\alpha} - \mathrm{D}a \circ K \circ \mathcal{R}_{\omega\alpha} E_K \\ &= \int_0^1 (1-s) \mathrm{D}^2 a\Big(K \circ \mathcal{R}_{\omega\alpha} + s E_K\Big)[E_K, E_K] \; ds, \\ \Delta^2 H :&= H \circ \Big(\phi_T \circ (K, \mathrm{id}), \mathcal{R}_\alpha\Big) - H \circ (K \circ \mathcal{R}_{\omega\alpha}, \mathcal{R}_\alpha) - \mathrm{D}_z H \circ (K \circ \mathcal{R}_{\omega\alpha}, \mathcal{R}_\alpha) E_K \\ &= \int_0^1 (1-s) \mathrm{D}_z^2 H\Big(K \circ \mathcal{R}_{\omega\alpha} + s E_K, \mathcal{R}_\alpha\Big)[E_K, E_K] \; ds. \end{aligned}$$

Proof. Let us start with $\langle \eta^{31} \rangle$ by using the exactness of the symplectic form as expressed in (1.4) to obtain

$$\left\langle \eta^{31} \right\rangle = \left\langle (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} (\mathrm{D}a)^\top \circ K \circ \mathcal{R}_{\omega\alpha} E_K \right\rangle - \left\langle (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \mathrm{D}a \circ K \circ \mathcal{R}_{\omega\alpha} E_K \right\rangle.$$

We now use that

$$\begin{aligned} 0 &= \left\langle \mathrm{D}_\theta \Big(a^\top \circ K \circ \mathcal{R}_{\omega\alpha} E_K \Big) \right\rangle \\ &= \left\langle {E_K}^\top \mathrm{D}_\theta \Big(a \circ K \circ \mathcal{R}_{\omega\alpha} \Big) + a^\top \circ K \circ \mathcal{R}_{\omega\alpha} \mathrm{D}_\theta E_K \right\rangle , \end{aligned}$$

so we have

$$\left\langle \eta^{31} \right\rangle = - \left\langle (\mathrm{D}_\theta E_K)^\top a \circ K \circ \mathcal{R}_{\omega\alpha} + (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \Big(\Delta^1 a - \Delta^2 a) \Big) \right\rangle .$$

We will now use

$$\begin{aligned} \mathrm{D}_\theta E_K &= \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\theta K - \mathrm{D}_\theta K \circ \mathcal{R}_{\omega\alpha} , \\ \left\langle (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} a \circ K \circ \mathcal{R}_{\omega\alpha} \right\rangle &= \left\langle (\mathrm{D}_\theta K)^\top a \circ K \right\rangle , \end{aligned}$$

Eq. (A.2), and

$$\mathrm{D}_\theta \Big(p_T \circ (K, \mathrm{id}) \Big) = \Big(a^\top \circ \phi_T \circ (K, \mathrm{id}) \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) - a^\top \circ K \Big) \mathrm{D}_\theta K$$

to rewrite $\langle \eta^{31} \rangle$ as

$$\begin{aligned} \left\langle \eta^{31} \right\rangle &= \left\langle (\mathrm{D}_\theta E_K)^\top \Delta^1 a + (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \Delta^2 a \right\rangle \\ &\quad - \left\langle (\mathrm{D}_\theta K)^\top \Big((\mathrm{D}_z \phi_T)^\top \circ (K, \mathrm{id}) a \circ \phi_T \circ (K, \mathrm{id}) + a \circ K \Big) \right\rangle \\ &= \left\langle (\mathrm{D}_\theta E_K)^\top \Delta^1 a + (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \Delta^2 a \right\rangle - \left\langle \mathrm{D}_\theta \Big(p_T \circ (K, \mathrm{id}) \Big)^\top \right\rangle \\ &= \left\langle (\mathrm{D}_\theta E_K)^\top \Delta^1 a + (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \Delta^2 a \right\rangle , \end{aligned}$$

For η^{32}, let us expand $\mathcal{X}$ as

$$\left\langle \eta^{32} \right\rangle = \left\langle X^\top \circ (K \circ \mathcal{R}_{\omega\alpha}, \mathcal{R}_\alpha) \Omega \circ K \circ \mathcal{R}_{\omega\alpha} E_K \right\rangle - \left\langle \hat{\alpha}^\top (\mathrm{D}_\varphi K)^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} E_K \right\rangle$$

and define

$$\begin{aligned} \eta_1^{32} &:= X^\top \circ (K \circ \mathcal{R}_{\omega\alpha}, \mathcal{R}_\alpha) \Omega \circ K \circ \mathcal{R}_{\omega\alpha} E_K , \\ \eta_2^{32} &:= - \hat{\alpha}^\top (\mathrm{D}_\varphi K)^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} E_K \end{aligned}$$

in order to inspect each term of η^{32} separately. We first use that $\mathrm{D}_z H = -X^\top \Omega$ to express $\langle \eta_1^{32} \rangle$ as

$$\left\langle \eta_1^{32} \right\rangle = - \left\langle \mathrm{D}_z H \circ (K \circ \mathcal{R}_{\omega\alpha}, \mathcal{R}_\alpha) E_K \right\rangle .$$

Hence,

$$\left\langle \eta_1^{32} \right\rangle = \left\langle \Delta^2 H \right\rangle - \left\langle H \circ \Big(\phi_T \circ (K, \mathrm{id}), \mathcal{R}_\alpha \Big) - H \circ (K, \mathrm{id}) \right\rangle$$

where we used that

$$\left\langle H \circ (K \circ \mathcal{R}_{\omega\alpha}, \mathcal{R}_\alpha) \right\rangle = \left\langle H \circ (K, \mathrm{id}) \right\rangle .$$

We now use the fundamental theorem of calculus and obtain

$$\begin{aligned} H \circ \Big(\phi_T \circ (K, \mathrm{id}), \mathcal{R}_\alpha \Big) - H \circ (K, \mathrm{id}) &= \int_0^T \frac{d}{ds} \Big(H \circ \Big(\phi_s \circ (K, \mathrm{id}), \psi_s \Big) \Big) \, ds \\ &= \int_0^T \Big(\mathrm{D}_z H \circ \Big(\phi_s \circ (K, \mathrm{id}), \psi_s \Big) X \circ \Big(\phi_s \circ (K, \mathrm{id}), \psi_s \Big) \Big) \, ds \\ &\quad + \int_0^T \Big(\mathrm{D}_\varphi H \circ \Big(\phi_s \circ (K, \mathrm{id}), \psi_s \Big) \hat{\alpha} \Big) \, ds , \end{aligned}$$

where the first integral varnishes because

$$\mathrm{D}_z H X = -X^\top \Omega X = 0.$$

Consequently,

$$\left\langle \eta_1^{32} \right\rangle = \left\langle \Delta^2 H \right\rangle - \left\langle \int_0^T \left(\mathrm{D}_\varphi H \circ \Big(\phi_s \circ (K, \mathrm{id}) \Big), \psi_s) \hat{\alpha} \right) \, ds \right\rangle.$$

We now use (A.3) to express $\langle \eta_1^{32} \rangle$ as

$$\begin{aligned}
\left\langle \eta_1^{32} \right\rangle &= \left\langle \Delta^2 H \right\rangle - \left\langle a^\top \circ \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi \phi_T \circ (K, \mathrm{id}) \hat{\alpha} - \mathrm{D}_\varphi p_T \circ (K, \mathrm{id}) \hat{\alpha} \right\rangle \\
&= \left\langle \Delta^2 H \right\rangle - \left\langle a^\top \circ \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi \phi_T \circ (K, \mathrm{id}) \hat{\alpha} \right\rangle - \left\langle \mathrm{D}_z p_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi K \hat{\alpha} \right\rangle,
\end{aligned}$$

where we used

$$0 = \left\langle \mathrm{D}_\varphi \Big(p_T \circ (K, \mathrm{id}) \Big) \right\rangle = \left\langle \mathrm{D}_z p_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi K + \mathrm{D}_\varphi p_T \circ (K, \mathrm{id}) \right\rangle.$$

We then use that

$$\mathrm{D}_\varphi E_K = \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi K + \mathrm{D}_\varphi \phi_T \circ (K, \mathrm{id}) - \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha},$$

in order to rewrite $\langle \eta_1^{32} \rangle$ as

$$\begin{aligned}
\left\langle \eta_1^{32} \right\rangle &= \left\langle \Delta^2 H \right\rangle - \left\langle a^\top \circ \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi E_K \hat{\alpha} \right\rangle + \left\langle a^\top \circ \phi_T \circ (K, \mathrm{id}) \mathrm{D}_z \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi K \hat{\alpha} \right\rangle \\
&\quad - \left\langle a^\top \circ \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha} \hat{\alpha} \right\rangle - \left\langle \mathrm{D}_z p_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi K \hat{\alpha} \right\rangle \\
&= \left\langle \Delta^2 H \right\rangle - \left\langle a^\top \circ \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha} \hat{\alpha} \right\rangle \\
&\quad - \left\langle a^\top \circ \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi E_K \hat{\alpha} - a^\top \circ K \circ \mathcal{R}_{\omega\alpha} \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha} \hat{\alpha} \right\rangle,
\end{aligned}$$

where we used Eq. (A.2) in the last equality and that

$$\left\langle a^\top \circ K \mathrm{D}_\varphi K \hat{\alpha} \right\rangle = \left\langle a^\top \circ K \circ \mathcal{R}_{\omega\alpha} \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha} \hat{\alpha} \right\rangle.$$

We leave the term η_1^{32} as it is. For the term η_2^{32}, we use (1.4) to obtain

$$\left\langle \eta_2^{32} \right\rangle = \left\langle \hat{\alpha}^\top (\mathrm{D}_\varphi K)^\top \circ \mathcal{R}_{\omega\alpha} \mathrm{D} a \circ K \circ \mathcal{R}_{\omega\alpha} E_K \right\rangle - \left\langle \hat{\alpha}^\top (\mathrm{D}_\varphi K)^\top \circ \mathcal{R}_{\omega\alpha} (\mathrm{D} a)^\top \circ K \circ \mathcal{R}_{\omega\alpha} E_K \right\rangle$$

Then, we use that

$$\begin{aligned}
0 &= \left\langle \mathrm{D}_\varphi \Big(a^\top \circ K \circ \mathcal{R}_{\omega\alpha} E_K \Big) \right\rangle \\
&= \left\langle {E_K}^\top \mathrm{D}_\varphi \Big(a \circ K \circ \mathcal{R}_{\omega\alpha} \Big) + a^\top \circ K \circ \mathcal{R}_{\omega\alpha} \mathrm{D}_\varphi E_K \right\rangle
\end{aligned}$$

and the definitions for $\Delta^1 a$ and $\Delta^2 a$ to obtain

$$\left\langle \eta_2^{32} \right\rangle = \left\langle a^\top \circ K \circ \mathcal{R}_{\omega\alpha} \mathrm{D}_\varphi E_K \hat{\alpha} \right\rangle + \left\langle \Big(\Delta^1 a - \Delta^2 a \Big)^\top \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha} \hat{\alpha} \right\rangle,$$

where we used that $\eta_2^{32} = \left(\eta_2^{32}\right)^\top$ since η_2^{32} is a scalar. Lastly, we obtain $\langle \eta^{32} \rangle$ as

$$\begin{aligned}
\left\langle \eta^{32} \right\rangle &= \left\langle \eta_1^{32} \right\rangle + \left\langle \eta_2^{32} \right\rangle \\
&= \left\langle \Delta^2 H \right\rangle - \left\langle a^\top \circ \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi E_K \hat{\alpha} - a^\top \circ K \circ \mathcal{R}_{\omega\alpha} \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha} \hat{\alpha} \right\rangle \\
&\quad - \left\langle a^\top \circ \phi_T \circ (K, \mathrm{id}) \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha} \hat{\alpha} \right\rangle + \left\langle a^\top \circ K \circ \mathcal{R}_{\omega\alpha} \mathrm{D}_\varphi E_K \hat{\alpha} \right\rangle \\
&\quad + \left\langle \left(\Delta^1 a - \Delta^2 a \right)^\top \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha} \hat{\alpha} \right\rangle \\
&= \left\langle \Delta^2 H \right\rangle - \left\langle (\Delta^1 a)^\top \mathrm{D}_\varphi E_K \hat{\alpha} \right\rangle - \left\langle (\Delta^2 a)^\top \mathrm{D}_\varphi K \circ \mathcal{R}_{\omega\alpha} \hat{\alpha} \right\rangle .
\end{aligned}$$

□

Appendix C

Fiberwise symplectic deformations and moment maps

In this section we introduce the notion of fiberwise symplectic deformations and establish their main properties. For a general exposition on symplectic deformations see [GHdlL14].

Definition C.1. A fiberwise symplectic deformation in $U \subset \mathbb{R}^{2n}$ with base $G \subset \mathbb{R}^m$ is a smooth diffeomorphism

$$\begin{aligned} \Phi: \quad U \times G &\longrightarrow U \times G \\ (z, \boldsymbol{t}) &\longmapsto \Big(\phi_{\boldsymbol{t}}(z), \tau(\boldsymbol{t})\Big), \end{aligned}$$

such that for all $\boldsymbol{t} \in G$, $\phi_{\boldsymbol{t}} : U \to U$ is symplectic. If for all $\boldsymbol{t} \in G$, $\phi_{\boldsymbol{t}}(z) = \phi(z, \boldsymbol{t})$ is exact symplectic, we will say that the fiberwise deformation is Hamiltonian.

Remark C.1. The inverse of Φ is given by

$$\begin{aligned} \Phi^{-1}: \quad U \times G &\longrightarrow U \times G \\ (z, \boldsymbol{t}) &\longmapsto \Big(\phi^{-1}_{\tau^{-1}(\boldsymbol{t})}(z), \tau^{-1}(\boldsymbol{t})\Big). \end{aligned}$$

Definition C.2. The primitive function of a fiberwise Hamiltonian deformation $\Phi : U \times G \to U \times G$ is a smooth function $p : U \times G \to \mathbb{R}$ such that for all $\boldsymbol{t} \in G$, the function $p_{\boldsymbol{t}} = p(\cdot, \boldsymbol{t})$ is the primitive function of $\phi_{\boldsymbol{t}}$.

Definition C.3. Let $\Phi : U \times G \to U \times G$ be a fiberwise Hamiltonian deformation and let $p : U \times G \to \mathbb{R}$ be the primitive function of Φ.

i) The generator of Φ is the function $\mathcal{F} : U \times G \to \mathbb{R}^{2n \times m}$ defined as

$$\mathcal{F} := \mathrm{D}_{\boldsymbol{t}}\phi \circ \Phi^{-1}, \tag{C.1}$$

where $\mathcal{F}_{\boldsymbol{t}}(z) = \mathcal{F}(z, \boldsymbol{t}) = \Big(\mathcal{F}^1(z, \boldsymbol{t}) \;\; \mathcal{F}^2(z, \boldsymbol{t}) \;\; \ldots \;\; \mathcal{F}^m(z, \boldsymbol{t})\Big)$.

ii) The moment map of Φ is the function $\mathcal{M} : U \times G \to \mathbb{R}^m$ defined as

$$\mathcal{M}^\top := a^\top \mathcal{F} - \mathrm{D}_{\boldsymbol{t}} p \circ \Phi^{-1},$$

where $\mathcal{M}_{\boldsymbol{t}}(z) = \mathcal{M}(z, \boldsymbol{t}) = \Big(\mathcal{M}^1(z, \boldsymbol{t}) \;\; \mathcal{M}^2(z, \boldsymbol{t}) \;\; \ldots \;\; \mathcal{M}^m(z, \boldsymbol{t})\Big)^\top$.

Lemma C.4. *For all $\boldsymbol{t} \in G$ and $z \in U$, the moment map $\mathcal{M}$ satisfies*

$$\mathcal{F} = \Omega^{-1}\Big(\mathrm{D}_z\mathcal{M}\Big)^{\top}.$$

Proof. For $i = 1, 2, \ldots, m$ and $\boldsymbol{t} = (t_1, \ldots, t_m)$ let us differentiate $\mathcal{M}^i \circ \Phi$ with respect to z as

$$\begin{aligned}
\mathrm{D}_z\Big(\mathcal{M}^i \circ \Phi\Big) &= \mathrm{D}_z\Big(a^{\top}\circ\phi\ \partial_{t_i}\phi - \partial_{t_i}p\Big) \\
&= a^{\top}\circ\phi\ \partial_{t_i}\mathrm{D}_z\phi + (\partial_{t_i}\phi)^{\top}\mathrm{D}a\circ\phi\ \mathrm{D}_z\phi - \partial_{t_i}\Big(a^{\top}\circ\phi\mathrm{D}_z\phi - a^{\top}\Big) \\
&= a^{\top}\circ\phi\ \partial_{t_i}\mathrm{D}_z\phi + (\partial_{t_i}\phi)^{\top}\mathrm{D}a\circ\phi\mathrm{D}_z\phi - a^{\top}\circ\phi\partial_{t_i}\mathrm{D}_z\phi - (\partial_{t_i}\phi)^{\top}(\mathrm{D}a\circ\phi)^{\top}\mathrm{D}_z\phi \\
&= -(\partial_{t_i}\phi)^{\top}\Omega\circ\phi\mathrm{D}_z\phi.
\end{aligned}$$

On the other hand, applying the chain rule we also have

$$\mathrm{D}_z\Big(\mathcal{M}^i \circ \Phi\Big) = \mathrm{D}_z\mathcal{M}^i \circ \Phi \mathrm{D}_z\phi.$$

Then, we obtain

$$\mathrm{D}_z\mathcal{M}^i \circ \Phi = -(\partial_{t_i}\phi)^{\top}\Omega\circ\phi,$$

and evaluating at $\Phi^{-1}(z, \boldsymbol{t})$ results in

$$\mathrm{D}_z\mathcal{M}^i = -\Big(\partial_{t_i}\phi\circ\Phi^{-1}\Big)^{\top}\Omega.$$

Equivalently, using the definition of the generator given by (C.1), we conclude

$$\mathcal{F}^i = \Omega^{-1}\Big(\mathrm{D}_z\mathcal{M}^i\Big)^{\top}.$$

□

Remark C.2. Note that, for $i = 1, 2, ..., m$, the moment map $\mathcal{M}$ gives the Hamiltonian $\mathcal{M}^i$ for the vector field $\mathcal{F}^i$ obtained by differentiating the symplectomorphism $\phi_{\boldsymbol{t}}$ with respect to t_i.

Let us now consider flows of a quasi-periodic Hamiltonian system, defined by the function H, with frequencies $\hat{\alpha} \in \mathbb{R}^{\ell}$. As described in Section 1.2, for each (z, φ) the flow $\tilde{\phi}$: $I_{z,\varphi} \times U \times \mathbb{T}^{\ell} \to U \times \mathbb{T}^{\ell}$, where $I_{z,\varphi} \subset \mathbb{R}$ is the maximal interval of existence for initial conditions (z, φ), adopts the form

$$\tilde{\phi}_t(z, \varphi) = \begin{pmatrix} \phi_t(z, \varphi) \\ \psi_t(\varphi) \end{pmatrix},$$

where $\psi_t(\varphi) = \psi(t, \varphi) = \varphi + \hat{\alpha}t$ is a linear flow on $\mathbb{T}^{\ell}$ with frequency $\hat{\alpha}$ and the evolution operator ϕ_t is fiberwise exact symplectic for all $t \in I_{z,\varphi}$. We observe that for fixed t, and each (z, φ), we can identify time-t maps with the following fiberwise Hamiltonian deformations

$$\Phi : \quad U \times G \longrightarrow U \times G$$

$$(z, \boldsymbol{t}) \longmapsto \Big(\phi_{\boldsymbol{t}}(z), \tau(\boldsymbol{t})\Big),$$

where $G \subset \mathbb{R}^{\ell+1}$, $\boldsymbol{t} = (t, \varphi)$, $\tau(\boldsymbol{t}) = \Big(t, \psi_{\boldsymbol{t}}\Big)$, $\psi_{\boldsymbol{t}} = \varphi + \hat{\alpha} t$, and $\phi_{\boldsymbol{t}}(z) = \phi_t(z, \varphi)$.

For a quasi-periodic Hamiltonian H_ε that depends on some parameter $\varepsilon \in \mathbb{R}$, we can again identify time-t maps with fiberwise Hamiltonian deformations as follows

$$\begin{aligned} \Phi : \quad U \times G &\longrightarrow U \times G \\ (z, \boldsymbol{t}) &\longmapsto \Big(\phi_{\boldsymbol{t}}(z), \tau(\boldsymbol{t})\Big), \end{aligned} \tag{C.2}$$

where $G \subset \mathbb{R}^{\ell+2}$, $\boldsymbol{t} = (t, \varepsilon, \varphi)$, $\tau(\boldsymbol{t}) = (t, \varepsilon, \psi_{\boldsymbol{t}}\Big)$, $\psi_{\boldsymbol{t}} = \varphi + \hat{\alpha}(\varepsilon) t$, and $\phi_{\boldsymbol{t}}(z) = \phi_{t,\varepsilon}(z, \varphi)$. Note than in our case, $\hat{\alpha}(\varepsilon) = \hat{\alpha}$ is constant but we could allow the dynamics on the base to depend on the parameter ε. Also, for the sake of notational symplicity, we will omit the dependence of the evolution operator ϕ_t on the parameter ε and simply write $\phi_{\boldsymbol{t}}(z) = \phi_t(z, \varphi)$.

We now consider the moment map $\mathcal{M}$ composed with Φ of the fiberwise symplectic deformation (C.2). Note that we could simply compute the moment map $\mathcal{M}$, but in order to prove Lemma D.1 we will need precisely the composition.

Lemma C.5. *The moment map of the fiberwise symplectic deformation Φ induced by the flow of H_ε satisfies*

$$\begin{aligned} \mathcal{M}^t \circ \Phi_{\boldsymbol{t}} &= H_\varepsilon \circ \Big(\phi_{\boldsymbol{t}}, \psi_{\boldsymbol{t}}\Big), \\ \mathcal{M}^\varepsilon \circ \Phi_{\boldsymbol{t}} &= \int_0^t \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big)\ ds, \\ \mathcal{M}^\varphi \circ \Phi_{\boldsymbol{t}} &= \int_0^t \mathrm{D}_\varphi H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big)\ ds, \end{aligned}$$

where $\boldsymbol{s} = (s, \varepsilon, \varphi)$.

Proof. For $\mathcal{M}^t \circ \Phi_{\boldsymbol{t}}$, let us use the definition of the moment map and the primitive function of ϕ_t as given in (A.1)—which, for each $(t, \varepsilon, \varphi)$, coincides with the primitive function of the fiberwise Hamiltonian deformation Φ—to obtain

$$\begin{aligned} \mathcal{M}^t \circ \Phi_{\boldsymbol{t}} &= a^\top \circ \phi_{\boldsymbol{t}}\ \partial_t \phi_{\boldsymbol{t}} - \partial_t p_{\boldsymbol{t}} \\ &= H_\varepsilon \circ \Big(\phi_{\boldsymbol{t}}, \psi_{\boldsymbol{t}}\Big). \end{aligned}$$

Note that this result could have been obtained directly from Remark C.2.

For $\mathcal{M}^\varepsilon \circ \Phi_{\boldsymbol{t}}$, let us first differentiate the primitive function of Φ with respect to ε

$$\begin{aligned} \partial_\varepsilon p_{\boldsymbol{t}} &= \partial_\varepsilon \int_0^t \Bigg(a^\top \circ \phi_{\boldsymbol{s}} X \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big) - H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big) \Bigg) ds \\ &= \int_0^t \Bigg(X^\top \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big) \mathrm{D} a \circ \phi_{\boldsymbol{s}}\ \partial_\varepsilon \phi_{\boldsymbol{s}} + a^\top \circ \phi_{\boldsymbol{s}} \Big(\mathrm{D}_z X \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big) \partial_\varepsilon \phi_{\boldsymbol{s}} \\ &\qquad + \partial_\varepsilon X \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big)\Big) - \mathrm{D}_z H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big) \partial_\varepsilon \phi_{\boldsymbol{s}} - \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big) \Bigg)\ ds. \end{aligned}$$

We use that $\mathrm{D}_z H_\varepsilon = -X^\top \Omega$ and (1.4) to obtain

$$\partial_\varepsilon p_{\boldsymbol{t}} = \int_0^t \bigg(X^\top \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big)(\mathrm{D}a\circ\phi_{\boldsymbol{s}})^\top \partial_\varepsilon \phi_{\boldsymbol{s}} + a^\top \circ \phi_{\boldsymbol{s}} \Big(\mathrm{D}_z X \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big) \partial_\varepsilon \phi_{\boldsymbol{s}} + \partial_\varepsilon X \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big)\Big) - \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big)\bigg) ds.$$

We can rewrite it as

$$\partial_\varepsilon p_{\boldsymbol{t}} = \int_0^t \bigg(\frac{d}{ds}\Big(a^\top \circ \phi_{\boldsymbol{s}}\Big) \partial_\varepsilon \phi_{\boldsymbol{s}} + a^\top \circ \phi_{\boldsymbol{s}} \frac{d}{ds}(\partial_\varepsilon \phi_{\boldsymbol{s}}) - \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big)\bigg) ds$$

and obtain

$$\begin{aligned} \partial_\varepsilon p_{\boldsymbol{t}} &= \int_0^t \bigg(\frac{d}{ds} \Big(a^\top \circ \phi_{\boldsymbol{s}} \partial_\varepsilon \phi_{\boldsymbol{s}}\Big) - \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big)\bigg) ds \\ &= a^\top \circ \phi_{\boldsymbol{t}}\ \partial_\varepsilon \phi_{\boldsymbol{t}} - \int_0^t \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big) ds. \end{aligned}$$

Then, by definition, we have

$$\begin{aligned} \mathcal{M}^\varepsilon \circ \Phi_{\boldsymbol{t}} &= a^\top \circ \phi_{\boldsymbol{t}}\ \partial_\varepsilon \phi_{\boldsymbol{t}} - \partial_\varepsilon p_{\boldsymbol{t}} \\ &= \int_0^t \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big)\ ds. \end{aligned}$$

For $\mathcal{M}^\varphi \circ \Phi_{\boldsymbol{t}}$, we will use the the expression of $\mathrm{D}_\varphi p_t$ given by (A.3). Then for the primitive function of Φ, we have

$$\mathrm{D}_\varphi p_{\boldsymbol{t}} = a^\top \circ \phi_{\boldsymbol{t}}\ \mathrm{D}_\varphi \phi_{\boldsymbol{t}} - \int_0^t \mathrm{D}_\varphi H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big)\ ds.$$

Hence,

$$\begin{aligned} \mathcal{M}^\varphi \circ \Phi_{\boldsymbol{t}} &= a^\top \circ \phi_{\boldsymbol{t}} \mathrm{D}_\varphi \phi_{\boldsymbol{t}} - \mathrm{D}_\varphi p_{\boldsymbol{t}} \\ &= \int_0^t \mathrm{D}_\varphi H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}}, \psi_{\boldsymbol{s}}\Big)\ ds. \end{aligned}$$

$\square$

Appendix D

Zero averages

For the continuation of an invariant torus $\mathcal{K}$ with respect to parameters of the Hamiltonian, we need the averages of η^3 given by

$$\eta^3 = \begin{pmatrix} (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \partial_\varepsilon \phi_T \circ \Big(K, \mathrm{id}\Big) \\ \mathcal{X}^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \partial_\varepsilon \phi_T \circ \Big(K, \mathrm{id}\Big) \end{pmatrix}$$

to be zero in order for the small divisors cohomological equations from Section 2.6 to be solvable. Recall the definition $\mathcal{R}_\alpha = \varphi + \alpha$, and let us define

$$\begin{pmatrix} \eta^{31} \\ \eta^{32} \end{pmatrix} := \begin{pmatrix} (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \partial_\varepsilon \phi_T \circ \Big(K, \mathrm{id}\Big) \\ \mathcal{X}^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \partial_\varepsilon \phi_T \circ \Big(K, \mathrm{id}\Big) \end{pmatrix}.$$

Lemma D.1. *The averages of η^{31} and η^{32} are zero.*

Proof. Let us consider the following fiberwise Hamiltonian deformation $\Phi : U \times G \to U \times G$

$$\begin{aligned} \Phi : \quad & U \times G \longrightarrow U \times G \\ & (z, \boldsymbol{t}) \longmapsto \Big(\phi_{\boldsymbol{t}}(z), \tau(\boldsymbol{t})\Big), \end{aligned}$$

with $G \subset \mathbb{R}^{\ell+2}$, $\boldsymbol{t} = (T, \varepsilon, \varphi)$, $\tau(\boldsymbol{t}) = \Big(T, \varepsilon, \psi_{\boldsymbol{t}}\Big)$, and $\psi_{\boldsymbol{t}} = \mathcal{R}_\alpha(\varphi)$, where $\phi_{\boldsymbol{t}}(z) = \phi_T(z, \varphi) = \phi(T, z, \varphi)$ is the evolution operator from (1.2) associated to a Hamiltonian H_ε depending on a parameter ε. See Appendix C. For the sake of notational symplicity we have omited the dependence of ϕ_T on ε. We can then use the definition of the generator $\mathcal{F}^\varepsilon$ of the fiberwise Hamiltonian deformation as given in (C.1), the invariance of $\mathcal{K}$, and Lemma C.4, and obtain for $\langle \eta^{31} \rangle$

$$\begin{aligned} \Big\langle \eta^{31} \Big\rangle &= \Big\langle (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \mathcal{F}^\varepsilon \circ \Big(\phi_{\boldsymbol{t}} \circ K, \tau(\boldsymbol{t})\Big) \Big\rangle \\ &= \Big\langle (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \Omega \circ K \circ \mathcal{R}_{\omega\alpha} \mathcal{F}^\varepsilon \circ \Big(K \circ \mathcal{R}_{\omega\alpha}, \tau(\boldsymbol{t})\Big) \Big\rangle \\ &= \Big\langle (\mathrm{D}_\theta K)^\top \circ \mathcal{R}_{\omega\alpha} \Big(\mathrm{D}_z \mathcal{M}^\varepsilon\Big)^\top \circ \Big(K \circ \mathcal{R}_{\omega\alpha}, \tau(\boldsymbol{t})\Big) \Big\rangle \end{aligned}$$

$$= \left\langle \mathrm{D}_\theta \Big(\mathcal{M}^\varepsilon \circ \big(K \circ \mathcal{R}_{\omega\alpha}, \tau(\boldsymbol{t}) \big) \Big)^\top \right\rangle$$
$$= 0$$

since we are taking averages of derivatives with respect to θ.

For $\langle \eta^{32} \rangle$, let us use Lemma C.4 and that

$$\mathrm{D}_z \phi_T \circ \Big(K, \mathrm{id} \Big) \mathcal{X} = \mathcal{X} \circ \mathcal{R}_{\omega\alpha}$$

to rewrite $\langle \eta^{32} \rangle$ as

$$\left\langle \eta^{32} \right\rangle = \left\langle \mathcal{X}^\top \mathrm{D}_z \Big(\mathcal{M}^\varepsilon \circ \big(\phi_{\boldsymbol{t}} \circ K, \tau(\boldsymbol{t}) \big) \Big)^\top \right\rangle.$$

Since η^{32} is a scalar, let us transpose the previous expression and use the explicit form of $\mathcal{M}^\varepsilon$ as given in Lemma C.5 to obtain

$$\left\langle \eta^{32} \right\rangle = \left\langle \int_0^T \Big(\mathrm{D}_z \partial_\varepsilon H_\varepsilon \circ \big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}} \big) \mathrm{D}_z \phi_{\boldsymbol{s}} \circ K \mathcal{X} \Big) \, ds \right\rangle,$$

where $\boldsymbol{s} = (s, \varepsilon, \varphi)$ and $s \in \mathbb{R}$. We then expand $\mathcal{X}$ to rewrite $\langle \eta^{32} \rangle$ as

$$\left\langle \eta^{32} \right\rangle = \left\langle \int_0^T \Big(\mathrm{D}_z \partial_\varepsilon H_\varepsilon \circ \big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}} \big) \mathrm{D}_z \phi_{\boldsymbol{s}} \circ K \Big(X \circ (K, \mathrm{id}) - \mathrm{D}_\varphi K \hat{\alpha} \Big) \Big) \, ds \right\rangle$$

and we use that

$$X \circ \Big(\phi_t(z, \varphi), \psi_t(\varphi) \Big) = \mathrm{D}_z \phi_t(z, \varphi) X(z, \varphi) + \mathrm{D}_\varphi \phi_t(z, \varphi) \hat{\alpha},$$

see (2.19), to obtain

$$\begin{aligned} \left\langle \eta^{32} \right\rangle = \Big\langle \int_0^T \Big(& \mathrm{D}_z \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}} \Big) \frac{d}{ds} \Big(\phi_{\boldsymbol{s}} \circ K \Big) \\ & - \mathrm{D}_z \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}} \Big) \mathrm{D}_z \phi_{\boldsymbol{s}} \circ K \ \mathrm{D}_\varphi K \hat{\alpha} \\ & - \mathrm{D}_z \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}} \Big) \mathrm{D}_\varphi \phi_{\boldsymbol{s}} \circ K \ \hat{\alpha} \Big) ds \Big\rangle. \end{aligned}$$

We can then rewrite it as

$$\begin{aligned} \left\langle \eta^{32} \right\rangle = \Big\langle \int_0^T \Big(& \mathrm{D}_z \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}} \Big) \frac{d}{ds} \Big(\phi_{\boldsymbol{s}} \circ K \Big) \\ & + \mathrm{D}_\varphi \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}} \Big) \hat{\alpha} + \mathrm{D}_\varphi \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}} \Big) \hat{\alpha} \\ & - \mathrm{D}_z \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}} \Big) \mathrm{D}_z \phi_{\boldsymbol{s}} \circ K \ \mathrm{D}_\varphi K \hat{\alpha} \\ & - \mathrm{D}_z \partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}} \Big) \mathrm{D}_\varphi \phi_{\boldsymbol{s}} \circ K \hat{\alpha} \Big) ds \Big\rangle \end{aligned}$$

to finally obtain

$$\left\langle \eta^{32} \right\rangle = \left\langle \int_0^T - \mathrm{D}_\varphi \Big(\partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}} \Big) \Big) \, ds \right\rangle \hat{\alpha}$$

$$
\begin{aligned}
&+ \left\langle \int_0^T \frac{d}{ds}\Big(\partial_\varepsilon H_\varepsilon \circ \Big(\phi_{\boldsymbol{s}} \circ K, \psi_{\boldsymbol{s}}\Big)\Big)\, ds \right\rangle \\
&= \left\langle -\mathrm{D}_\varphi \Big(\mathcal{M}^\varepsilon \circ \Big(\phi_{\boldsymbol{t}} \circ K, \tau(\boldsymbol{t})\Big)\Big) \right\rangle \hat{\alpha} \\
&\quad + \Big\langle \partial_\varepsilon H_\varepsilon \circ \Big(K \circ \mathcal{R}_{\omega\alpha}, \mathcal{R}_\alpha\Big) - \partial_\varepsilon H_\varepsilon \circ \Big(K, \mathrm{id}\Big)\Big\rangle \\
&= 0,
\end{aligned}
$$

where we used that $\partial_\varepsilon H_\varepsilon \circ \Big(K, \mathrm{id}\Big)$ and $\partial_\varepsilon H_\varepsilon \circ \Big(K \circ \mathcal{R}_{\omega\alpha}, \mathcal{R}_\alpha\Big)$ have the same average and that the average of derivatives with respect to φ is zero. $\square$

Appendix E

Compendium of constants involved in the KAM theorem

Table E.1: Constants in Section 3.2.1

Object	Constant	Label
$\mathcal{X}$	$C_{\mathcal{X}} = c_X + \sigma_{\mathrm{D}_\varphi K}\lvert\hat{\alpha}\rvert$	(3.8)
$\mathcal{X}^\top$	$C_{\mathcal{X}^\top} = c_{J^\top} + \lvert\hat{\alpha}^\top\rvert\sigma_{(\mathrm{D}_\varphi K)^\top}$	(3.8)
L	$C_L = \sigma_{\mathrm{D}_\theta K} + C_{\mathcal{X}} + \sigma_W$	(3.9)
$L^\top$	$C_{L^\top} = \max\left\{\sigma_{(\mathrm{D}_\theta K)^\top}, C_{\mathcal{X}^\top}, \sigma_{W^\top}\right\}$	(3.9)
$\hat{P}$	$C_{\hat{P}} = C_L + \sigma_{\hat{N}}$	(3.10)
$\hat{P}^\top$	$C_{\hat{P}^\top} = \max\left\{C_{L^\top},\ \sigma_{\hat{N}^\top}\right\}$	(3.10)
$\hat{S}$	$C_{\hat{S}} = \sigma_{\hat{N}^\top} c_\Omega c_{\mathrm{D}_z\phi_T}\sigma_{\hat{N}}$	(3.11)
$\hat{S}^\top$	$C_{\hat{S}^\top} = \sigma_{\hat{N}^\top} c_\Omega c_{(\mathrm{D}_z\phi_T)^\top}\sigma_{\hat{N}}$	(3.11)
A	$C_A = \frac{1}{1-(\sigma_\lambda)^2}\left(C_{\hat{S}}\sigma_\lambda + C_{\hat{S}^\top}\right)$	(3.12)
$A^\top$	$C_{A^\top} = \frac{1}{1-(\sigma_\lambda)^2}\left(C_{\hat{S}^\top}\sigma_\lambda + C_{\hat{S}}\right)$	(3.13)
N	$C_N = \sigma_{\hat{N}} + C_L C_A$	(3.14)
$N^\top$	$C_{N^\top} = \sigma_{\hat{N}^\top} + C_{A^\top} C_{L^\top}$	(3.15)
P	$C_P = C_L + C_N$	(3.16)
$P^\top$	$C_{P^\top} = \max\left\{C_{L^\top},\ C_{N^\top}\right\}$	(3.17)

Table E.2: Constants in Lemmas 3.2, 3.3, and 3.4.

Object	Constant	Label
$E_{\mathcal{X}}$	$C_{E_{\mathcal{X}}} = c_{\mathrm{D}_z X}\delta + \ell\lvert\hat{\alpha}\rvert$	(3.25)
$E_{\mathcal{X}}^{\top}$	$C_{E_{\mathcal{X}}^{\top}} = 2n\left(c_{(\mathrm{D}_z X)^{\top}}\delta + \left\lvert\hat{\alpha}^{\top}\right\rvert\right)$	(3.26)
E_L	$C_{E_L}^{K} = d + C_{E_{\mathcal{X}}}, \quad C_{E_L}^{W} = 1$	(3.21)
$E_L^{\top}$	$C_{E_L^{\top}}^{K} = \max\left\{2n,\ C_{E_{\mathcal{X}}^{\top}}\right\}, \quad C_{E_L^{\top}}^{W} = 2n$	(3.22)
$\mathfrak{L}\Omega_{\mathrm{D}K}$	$C_{\mathfrak{L}\Omega_{\mathrm{D}K}} = \sigma_{(\mathrm{D}_\theta K)^{\top}} c_{\mathrm{D}\Omega}\sigma_{\mathrm{D}_\theta K}\delta + \sigma_{(\mathrm{D}_\theta K)^{\top}} c_{\Omega} d + 2n c_{\Omega} c_{\mathrm{D}_z\phi_T}\sigma_{\mathrm{D}_\theta K}$	(3.30)
$\mathfrak{L}\Omega_{\mathrm{D}K\mathcal{X}}$	$C_{\mathfrak{L}\Omega_{\mathrm{D}K\mathcal{X}}} = \sigma_{(\mathrm{D}_\theta K)^{\top}} c_{\mathrm{D}\Omega} C_{\mathcal{X}}\delta + \sigma_{(\mathrm{D}_\theta K)^{\top}} c_{\Omega} C_{E_{\mathcal{X}}} + 2n c_{\Omega} c_{\mathrm{D}_z\phi_T} C_{\mathcal{X}}$	(3.34)
$\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW}$	$C_{\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW}}^{K} = \sigma_{(\mathrm{D}_\theta K)^{\top}} c_{\mathrm{D}\Omega}\sigma_W\sigma_\lambda\delta + 2n c_{\Omega} c_{\mathrm{D}_z\phi_T}\sigma_W, \quad C_{\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW}}^{W} = \sigma_{(\mathrm{D}_\theta K)^{\top}} c_{\Omega}$	(3.35)
$\mathfrak{L}\Omega_{\mathcal{X}\mathrm{D}K}$	$C_{\mathfrak{L}\Omega_{\mathcal{X}\mathrm{D}K}} = C_{\mathcal{X}^{\top}} c_{\mathrm{D}\Omega}\sigma_{\mathrm{D}_\theta K}\delta + C_{\mathcal{X}^{\top}} c_{\Omega} d + C_{E_{\mathcal{X}}^{\top}} c_{\Omega} c_{\mathrm{D}_z\phi_T}\sigma_{\mathrm{D}_\theta K}$	(3.36)
$\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W}$	$C_{\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W}}^{K} = C_{\mathcal{X}^{\top}} c_{\mathrm{D}\Omega}\sigma_W\sigma_\lambda\delta + C_{E_{\mathcal{X}}^{\top}} c_{\Omega} c_{\mathrm{D}_z\phi_T}\sigma_W, \quad C_{\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W}}^{W} = C_{\mathcal{X}^{\top}} c_{\Omega}$	(3.37)
$\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K}$	$C_{\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K}}^{K} = \sigma_\lambda\sigma_{W^{\top}} c_{\mathrm{D}\Omega}\sigma_{\mathrm{D}_\theta K}\delta + \sigma_\lambda\sigma_{W^{\top}} c_{\Omega} d, \quad C_{\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K}}^{W} = 2n c_{\Omega} c_{\mathrm{D}_z\phi_T}\sigma_{\mathrm{D}_\theta K}$	(3.38)
$\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}}$	$C_{\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}}}^{K} = \sigma_\lambda\sigma_{W^{\top}} c_{\mathrm{D}\Omega} C_{\mathcal{X}}\delta + \sigma_\lambda\sigma_{W^{\top}} c_{\Omega} C_{E_{\mathcal{X}}}, \quad C_{\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}}}^{W} = 2n c_{\Omega} c_{\mathrm{D}_z\phi_T} C_{\mathcal{X}}$	(3.39)
$\Omega_{\mathrm{D}K}$	$C_{\Omega_{\mathrm{D}K}} = c_{\mathfrak{R}} C_{\mathfrak{L}\Omega_{\mathrm{D}K}}$	(3.29)
$\Omega_{\mathrm{D}K\mathcal{X}}$	$C_{\Omega_{\mathrm{D}K\mathcal{X}}} = c_{\mathfrak{R}} C_{\mathfrak{L}\Omega_{\mathrm{D}K\mathcal{X}}}$	(3.40)
$\Omega_{\mathrm{D}KW}$	$C_{\Omega_{\mathrm{D}KW}}^{K} = \frac{1}{1-\sigma_\lambda} C_{\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW}}^{K}, \quad C_{\Omega_{\mathrm{D}KW}}^{W} = \frac{1}{1-\sigma_\lambda} C_{\mathfrak{L}^{1\lambda}\Omega_{\mathrm{D}KW}}^{W}$	(3.41)
$\Omega_{\mathcal{X}\mathrm{D}K}$	$C_{\Omega_{\mathcal{X}\mathrm{D}K}} = c_{\mathfrak{R}} C_{\mathfrak{L}\Omega_{\mathcal{X}\mathrm{D}K}}$	(3.42)
$\Omega_{\mathcal{X}W}$	$C_{\Omega_{\mathcal{X}W}}^{K} = \frac{1}{1-\sigma_\lambda} C_{\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W}}^{K}, \quad C_{\Omega_{\mathcal{X}W}}^{W} = \frac{1}{1-\sigma_\lambda} C_{\mathfrak{L}^{1\lambda}\Omega_{\mathcal{X}W}}^{W}$	(3.43)
$\Omega_{W\mathrm{D}K}$	$C_{\Omega_{W\mathrm{D}K}}^{K} = \frac{1}{1-\sigma_\lambda} C_{\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K}}^{K}, \quad C_{\Omega_{W\mathrm{D}K}}^{W} = \frac{1}{1-\sigma_\lambda} C_{\mathfrak{L}^{1\lambda}\Omega_{W\mathrm{D}K}}^{W}$	(3.44)
$\Omega_{W\mathcal{X}}$	$C_{\Omega_{W\mathcal{X}}}^{K} = \frac{1}{1-\sigma_\lambda} C_{\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}}}^{K}, \quad C_{\Omega_{W\mathcal{X}}}^{W} = \frac{1}{1-\sigma_\lambda} C_{\mathfrak{L}^{1\lambda}\Omega_{W\mathcal{X}}}^{W}$	(3.45)
Ω_L	$C_{\Omega_L}^{K} = \max\left\{C_{\Omega_{\mathrm{D}K}} + C_{\Omega_{\mathrm{D}K\mathcal{X}}} + C_{\Omega_{\mathrm{D}KW}}^{K}\gamma\delta^{\tau},\ C_{\Omega_{\mathcal{X}\mathrm{D}K}} + C_{\Omega_{\mathcal{X}W}}^{K}\gamma\delta^{\tau},\ \left(C_{\Omega_{W\mathrm{D}K}}^{K} + C_{\Omega_{W\mathcal{X}}}^{K}\right)\gamma\delta^{\tau}\right\}$ $C_{\Omega_L}^{W} = \max\left\{C_{\Omega_{\mathrm{D}KW}}^{W},\ C_{\Omega_{\mathcal{X}W}}^{W},\ C_{\Omega_{W\mathrm{D}K}}^{W} + C_{\Omega_{W\mathcal{X}}}^{W}\right\}$	(3.32)

Table E.3: Constants in Lemmas 3.5, 3.6, 3.7, 3.8, 3.9, 3.10, and 3.11

Object	Constant	Label
$E_{\mathrm{symp}\hat{P}}$	$C^K_{E_{\mathrm{symp}\hat{P}}} = \max\{1,\ (\sigma_B)^2\}\, C^K_{\Omega_L}, \quad C^W_{E_{\mathrm{symp}\hat{P}}} = \max\{1,\ (\sigma_B)^2\}\, C^W_{\Omega_L}$	(3.48)
$E^{11}_{\mathrm{red}\hat{P}}$	$C^K_{E^{11}_{\mathrm{red}\hat{P}}} = \sigma_{\hat{N}^\top} c_\Omega C^K_{E_L}, \quad C^W_{E^{11}_{\mathrm{red}\hat{P}}} = \sigma_{\hat{N}^\top} c_\Omega C^W_{E_L}$	(3.51)
$E^{21}_{\mathrm{red}\hat{P}}$	$C^K_{E^{21}_{\mathrm{red}\hat{P}}} = C^K_{\Omega_L} + C_{L^\top} c_\Omega C^K_{E_L} \gamma\delta^\tau, \quad C^W_{E^{21}_{\mathrm{red}\hat{P}}} = C^W_{\Omega_L} + C_{L^\top} c_\Omega C^W_{E_L}$	(3.52)
$E^{22}_{\mathrm{red}\hat{P}}$	$C^K_{E^{22}_{\mathrm{red}\hat{P}}} = C_{L^\top} c_{\mathrm{D}\Omega} c_{\mathrm{D}_z\phi_T} \sigma_{\hat{N}} \delta + \sigma_{\lambda^{-1}} C^K_{E_L^\top} c_\Omega c_{\mathrm{D}_z\phi_T} \sigma_{\hat{N}}, \quad C^W_{E^{22}_{\mathrm{red}P}} = \sigma_{\lambda^{-1}} C^W_{E_L^\top} c_\Omega c_{\mathrm{D}_z\phi_T} \sigma_{\hat{N}}$	(3.53)
$E_{\mathrm{red}\hat{P}}$	$C^K_{E_{\mathrm{red}\hat{P}}} = \max\left\{ C^K_{E^{11}_{\mathrm{red}\hat{P}}} \gamma\delta^\tau,\ C^K_{E^{21}_{\mathrm{red}\hat{P}}} + C^K_{E^{22}_{\mathrm{red}\hat{P}}} \gamma\delta^\tau \right\}, \quad C^W_{E_{\mathrm{red}\hat{P}}} = \max\left\{ C^W_{E^{11}_{\mathrm{red}\hat{P}}},\ C^W_{E^{21}_{\mathrm{red}\hat{P}}} + C^W_{E^{22}_{\mathrm{red}\hat{P}}} \right\}$	(3.50)
$E_{\mathrm{inv}\hat{P}}$	$C^K_{E_{\mathrm{inv}\hat{P}}} = \frac{1}{1-\hat{\nu}} C_{\hat{P}} C_{\hat{P}^\top} c_\Omega C^K_{E_{\mathrm{symp}P}}, \quad C^W_{E_{\mathrm{inv}\hat{P}}} = \frac{1}{1-\hat{\nu}} C_{\hat{P}} C_{\hat{P}^\top} c_\Omega C^W_{E_{\mathrm{symp}P}}$	(3.58)
$E_{\mathrm{sym}\Lambda^{-1}\hat{S}}$	$C^K_{\Lambda^{-1}\hat{S}} = (\sigma_B)^2 C^K_{\Omega_L} + \sigma_{\hat{N}^\top} c_{(\mathrm{D}_z\phi_T)^\top} \Big(C_{\hat{S}} \left(c_{\mathrm{D}\Omega} C_L \delta + \sigma_{\lambda^{-1}} c_\Omega C^K_{E_L} \right) \gamma\delta^\tau + c_{\mathrm{D}\Omega} c_{\mathrm{D}_z\phi_T} \sigma_{\hat{N}} \gamma \delta^{\tau+1}$ $+ c_\Omega c_{\mathrm{D}_z\phi_T} \sigma_{\hat{N}} C^K_{E_{\mathrm{inv}\hat{P}}} + c_\Omega C_{\hat{P}} C^K_{E^{22}_{\mathrm{red}P}} \gamma\delta^\tau \Big)$ $C^W_{\Lambda^{-1}\hat{S}} = (\sigma_B)^2 C^W_{\Omega_L} + \sigma_{\hat{N}^\top} c_{(\mathrm{D}_z\phi_T)^\top} c_\Omega \Big(C_{\hat{S}} \sigma_{\lambda^{-1}} C^W_{E_L} + c_{\mathrm{D}_z\phi_T} \sigma_{\hat{N}} C^W_{E_{\mathrm{inv}\hat{P}}} + C_{\hat{P}} C^W_{E^{22}_{\mathrm{red}P}} \Big)$	(3.60)
$E_{\mathrm{sym}A}$	$C^K_{E_{\mathrm{sym}A}} = \frac{1}{1-\sigma_\lambda} C^K_{\Lambda^{-1}\hat{S}}, \quad C^W_{E_{\mathrm{sym}A}} = \frac{1}{1-\sigma_\lambda} C^W_{\Lambda^{-1}\hat{S}}$	(3.63)
$E_{\mathrm{symp}P}$	$C^K_{E_{\mathrm{symp}P}} = (1 + C_{A^\top})(1 + C_A)\, C^K_{E_{\mathrm{symp}\hat{P}}} + C^K_{E_{\mathrm{sym}A}}$ $C^W_{E_{\mathrm{symp}P}} = (1 + C_{A^\top})(1 + C_A)\, C^W_{E_{\mathrm{symp}\hat{P}}} + C^W_{E_{\mathrm{sym}A}}$	(3.65)
$E_{\mathrm{red}P}$	$C^K_{E_{\mathrm{red}P}} = (1 + C_{A^\top})(1 + C_A)\, C^K_{E_{\mathrm{red}\hat{P}}}, \quad C^W_{E_{\mathrm{red}P}} = (1 + C_{A^\top})(1 + C_A)\, C^W_{E_{\mathrm{red}\hat{P}}}$	(3.67)

Table E.4: Constants in Lemma 3.12

Object	Constant	Label
η_K^1	$C_{\eta_K^1} = C_{N^\top} c_\Omega$	(3.74)
η_K^2	$C_{\eta_K^2} = C_{N^\top} c_\Omega$	(3.75)
η_K^3	$C_{\eta_K^3} = \max\left\{\sigma_{(\mathrm{D}_\theta K)^\top},\ C_{\mathcal{X}^\top}\right\}$	(3.76)
$\langle\eta_K^3\rangle$	$C_{\langle\eta_K^{31}\rangle} = 2n c_{\mathrm{D}a} + \sigma_{(\mathrm{D}_\theta K)^\top} c_{\mathrm{D}^2 a} \frac{\delta}{2}, \quad C_{\langle\eta_K^{32}\rangle} = 2n\ell c_{(\mathrm{D}a)^\top} \lvert\hat{\alpha}\rvert + (c_{\mathrm{D}_z^2 H} + 2n\ell c_{\mathrm{D}^2 a} \sigma_{\mathrm{D}_\varphi K} \lvert\hat{\alpha}\rvert) \frac{\delta}{2}$ $C_{\langle\eta_K^3\rangle} = \max\left\{C_{\langle\eta_K^{31}\rangle},\ C_{\langle\eta_K^{32}\rangle}\right\}$	(3.85)-(3.87)
η_K^4	$C_{\eta_K^4} = \sigma_{W^\top} c_\Omega$	(3.77)
$\langle\xi_K^3\rangle$	$C_{\langle\xi_K^3\rangle} = \sigma_{\langle\mathcal{S}\rangle^{-1}} \left(C_{\eta_K^1} \gamma\delta^\tau + C_{\hat{S}} c_{\mathfrak{R}} C_{\eta_K^3}\right)$	(3.78)
ξ_K^1	$C_{\xi_K^1} = c_{\mathfrak{R}} \left(C_{\eta_K^1} \gamma\delta^\tau + C_{\hat{S}} \left(c_{\mathfrak{R}} C_{\eta_K^3} + C_{\langle\xi_K^3\rangle}\right)\right)$	(3.79)
ξ_K^2	$C_{\xi_K^2} = \frac{1}{1-\sigma_\lambda} C_{\eta_K^2}$	(3.80)
ξ_K^3	$C_{\xi_K^3} = c_{\mathfrak{R}} C_{\eta_K^3} + C_{\langle\xi_K^3\rangle}$	(3.81)
ξ_K^4	$C_{\xi_K^4} = \frac{\sigma_\lambda}{1-\sigma_\lambda} C_{\eta_K^4}$	(3.82)
ξ_K	$C_{\xi_K} = \max\left\{C_{\xi_K^1},\ C_{\xi_K^2}\gamma^2\delta^{2\tau},\ C_{\xi_K^3}\gamma\delta^\tau,\ C_{\xi_K^4}\gamma^2\delta^{2\tau}\right\}$	(3.83)
ΔK	$C_{\Delta K} = C_L \max\left\{C_{\xi_K^1},\ C_{\xi_K^2}\gamma^2\delta^{2\tau}\right\} + C_N \max\left\{C_{\xi_K^3}\gamma\delta^\tau,\ C_{\xi_K^4}\gamma^2\delta^{2\tau}\right\}$	(3.84)
E_{lin}^K	$C_{E_{\mathrm{lin}}^K}^K = \left(C_{E_{\mathrm{red}P}}^K + C_{E_{\mathrm{symp}P}}^K\right) C_{\xi_K} + C_{\langle\eta_K^3\rangle}\gamma^3\delta^{3\tau}, \quad C_{E_{\mathrm{lin}}^K}^{KW} = \left(C_{E_{\mathrm{red}P}}^W + C_{E_{\mathrm{symp}P}}^W\right) C_{\xi_K}$	(3.88)
$E_{\bar{K}}$	$C_{E_{\bar{K}}}^{KK} = \frac{C_P}{1-\nu} C_{E_{\mathrm{lin}}^K}^K \gamma\delta^{\tau-1} + \frac{1}{2} c_{\mathrm{D}_z^2\phi_T} (C_{\Delta K})^2, \quad C_{E_{\bar{K}}}^{KW} = \frac{C_P}{1-\nu} C_{E_{\mathrm{lin}}^K}^{KW}$	(3.90)

Table E.5: Constants in Lemma 3.14

Object	Constant	Label
$\widetilde{E_W}$	$C^K_{\widetilde{E_W}} = c_{\mathrm{D}^2_z\phi_T} C_{\Delta K}\sigma_W, \quad C^W_{\widetilde{E_W}} = 1$	(3.94)
η^1_W	$C^K_{\eta^1_W} = C_{N^\top} c_\Omega C^K_{\widetilde{E_W}}, \quad C^W_{\eta^1_W} = C_{N^\top} c_\Omega$	(3.95)
η^2_W	$C^K_{\eta^2_W} = C_{N^\top} c_\Omega C^K_{\widetilde{E_W}}, \quad C^W_{\eta^2_W} = C_{N^\top} c_\Omega$	(3.96)
η^3_W	$C^K_{\eta^3_W} = \max\left\{\sigma_{(\mathrm{D}_\theta K)^\top},\ C_{\mathcal{X}^\top}\right\} c_\Omega C^K_{\widetilde{E_W}}, \quad C^W_{\eta^3_W} = \max\left\{\sigma_{(\mathrm{D}_\theta K)^\top},\ C_{\mathcal{X}^\top}\right\} c_\Omega$	(3.97)
η^4_W	$C^K_{\eta^4_W} = \sigma_{W^\top} c_\Omega C^K_{\widetilde{E_W}}, \quad C^W_{\eta^4_W} = \sigma_{W^\top} c_\Omega$	(3.98)
ξ^1_W	$C^K_{\xi^1_W} = \frac{1}{1-\sigma_\lambda}\left(C^K_{\eta^1_W} + \frac{C_{\hat{S}} C^K_{\eta^3_W}}{1-\sigma_\lambda}\right), \quad C^W_{\xi^1_W} = \frac{1}{1-\sigma_\lambda}\left(C^W_{\eta^1_W} + \frac{C_{\hat{S}} C^W_{\eta^3_W}}{1-\sigma_\lambda}\right)$	(3.99)
ξ^2_W	$C^K_{\xi^2_W} = c_{\mathfrak{R}\sigma_{\lambda^{-1}}} C^K_{\eta^2_W}, \quad C^W_{\xi^2_W} = c_{\mathfrak{R}\sigma_{\lambda^{-1}}} C^W_{\eta^2_W}$	(3.100)
ξ^3_W	$C^K_{\xi^3_W} = \frac{1}{1-\sigma_\lambda} C^K_{\eta^3_W}, \quad C^W_{\xi^3_W} = \frac{1}{1-\sigma_\lambda} C^W_{\eta^3_W}$	(3.101)
ξ^4_W	$C^K_{\xi^4_W} = \frac{\sigma_\lambda}{1-(\sigma_\lambda)^2} C^K_{\eta^4_W}, \quad C^W_{\xi^4_W} = \frac{\sigma_\lambda}{1-(\sigma_\lambda)^2} C^W_{\eta^4_W}$	(3.102)
ξ_W	$C^K_{\xi_W} = \max\left\{C^K_{\xi^1_W}\gamma\delta^\tau,\ C^K_{\xi^2_W},\ C^K_{\xi^3_W}\gamma\delta^\tau,\ C^K_{\xi^4_W}\gamma\delta^\tau\right\}$ $C^W_{\xi_W} = \max\left\{C^W_{\xi^1_W}\gamma\delta^\tau,\ C^W_{\xi^2_W},\ C^W_{\xi^3_W}\gamma\delta^\tau,\ C^W_{\xi^4_W}\gamma\delta^\tau\right\}$	(3.103)
ΔW	$C^K_{\Delta W} = C_L \max\left\{C^K_{\xi^1_W}\gamma\delta^\tau,\ C^K_{\xi^2_W}\right\} + C_N \max\left\{C^K_{\xi^3_W},\ C^K_{\xi^4_W}\right\}\gamma\delta^\tau$ $C^W_{\Delta W} = C_L \max\left\{C^W_{\xi^1_W}\gamma\delta^\tau,\ C^W_{\xi^2_W}\right\} + C_N \max\left\{C^W_{\xi^3_W},\ C^W_{\xi^4_W}\right\}\gamma\delta^\tau$	(3.104)
$\Delta\lambda$	$C^K_{\Delta\lambda} = C^K_{\eta^2_W}, \quad C^W_{\Delta\lambda} = C^W_{\eta^2_W}$	(3.105)
E^W_{lin}	$C^{KK}_{E^W_{\mathrm{lin}}} = C^K_{E_{\mathrm{red}P}} C^K_{\xi_W} + C^K_{E_{\mathrm{symp}P}} C^K_{\xi_W}\sigma_\lambda + C^K_{E_{\mathrm{symp}P}} C^K_{\Delta\lambda}\gamma\delta^\tau$ $C^{KW}_{E^W_{\mathrm{lin}}} = C^W_{E_{\mathrm{red}P}} C^K_{\xi_W} + C^W_{E_{\mathrm{symp}P}} C^K_{\xi_W} + \left(C^K_{E_{\mathrm{red}P}} C^W_{\xi_W} + C^K_{E_{\mathrm{symp}P}} C^W_{\xi_W}\right)\gamma\delta^{\tau-1}$ $\qquad + C^W_{E_{\mathrm{symp}P}} C^K_{\Delta\lambda}\gamma\delta^\tau + C^K_{E_{\mathrm{symp}P}} C^W_{\Delta\lambda}\gamma^2\delta^{2\tau-1}$ $C^{WW}_{E^W_{\mathrm{lin}}} = C^W_{E_{\mathrm{red}P}} C^W_{\xi_W} + C^W_{E_{\mathrm{symp}P}} C^W_{\xi_W}\sigma_\lambda + C^W_{E_{\mathrm{symp}P}} C^W_{\Delta\lambda}\gamma\delta^\tau$	(3.106)
$\Delta \mathrm{D}_z\phi_T[\Delta W]$	$C^{KK}_{\Delta \mathrm{D}_z\phi_T[\Delta W]} = c_{\mathrm{D}^2_z\phi_T} C_{\Delta K} C^K_{\Delta W}, \quad C^{KW}_{\Delta \mathrm{D}_z\phi_T[\Delta W]} = c_{\mathrm{D}^2_z\phi_T} C_{\Delta K} C^W_{\Delta W}$	(3.107)
$\Delta W\Delta\lambda$	$C^{KK}_{\Delta W\Delta\lambda} = C^K_{\Delta W} C^K_{\Delta\lambda}, \quad C^{KW}_{\Delta W\Delta\lambda} = C^K_{\Delta W} C^W_{\Delta\lambda} + C^W_{\Delta W} C^K_{\Delta\lambda}, \quad C^{WW}_{\Delta W\Delta\lambda} = C^W_{\Delta W} C^W_{\Delta\lambda}$	(3.108)
$E_{\bar{W}}$	$C^{KK}_{E_{\bar{W}}} = \frac{C_P}{1-\nu} C^{KK}_{E^W_{\mathrm{lin}}}\gamma\delta^{\tau-1} + C^{KK}_{\Delta \mathrm{D}_z\phi_T[\Delta W]} + C^{KK}_{\Delta W\Delta\lambda}$ $C^{KW}_{E_{\bar{W}}} = \frac{C_P}{1-\nu} C^{KW}_{E^W_{\mathrm{lin}}} + C^{KW}_{\Delta \mathrm{D}_z\phi_T[\Delta W]} + C^{KW}_{\Delta W\Delta\lambda}, \quad C^{WW}_{E_{\bar{W}}} = \frac{C_P}{1-\nu} C^{WW}_{E^W_{\mathrm{lin}}} + C^{WW}_{\Delta W\Delta\lambda}$	(3.109)

Table E.6: Constants in Lemma 3.15

Object	Constant	Label
$E_{\mathrm{symp}\hat{P}}$	$C_{E_{\mathrm{symp}\hat{P}}} = C^K_{E_{\mathrm{symp}\hat{P}}}\gamma\delta^{\tau-1} + C^W_{E_{\mathrm{symp}\hat{P}}}$	(3.137)
$E_{\mathrm{symp}P}$	$C_{E_{\mathrm{symp}P}} = C^K_{E_{\mathrm{symp}P}}\gamma\delta^{\tau-1} + C^W_{E_{\mathrm{symp}P}}$	(3.138)
$\bar{W} - W$	$C_{\Delta W} = C^K_{\Delta W} + C^W_{\Delta W}$	(3.139)
$\bar{L} - L$	$C^K_{\Delta L} = (d + \ell\lvert\hat{\alpha}\rvert + c_{\mathrm{D}_z X}\delta)\, C_{\Delta K}\gamma\delta^{\tau-1} + C^K_{\Delta W}, \quad C^W_{\Delta L} = C^W_{\Delta W}, \quad C_{\Delta L} = C^K_{\Delta L} + C^W_{\Delta L}$	(3.140)
$\bar{L}^\top - L^\top$	$C^K_{\Delta L^\top} = 2n\left(\max\left\{1, c_{(\mathrm{D}_z X)^\top}\delta + \lvert\hat{\alpha}^\top\rvert\right\} C_{\Delta K}\gamma\delta^{\tau-1} + C^K_{\Delta W}\right), \quad C^W_{\Delta L^\top} = 2nC^W_{\Delta W}$ $C_{\Delta L^\top} = C^K_{\Delta L^\top} + C^W_{\Delta L^\top}$	(3.141)
$G_{\bar{L}} - G_L$	$C^K_{\Delta G_L} = C_{L^\top} c_{\mathrm{D}G} C_L C_{\Delta K}\gamma\delta^\tau + C_{L^\top} c_G C^K_{\Delta L} + c_G C_L C^K_{\Delta L^\top}$ $C^W_{\Delta G_L} = C_{L^\top} c_G C^W_{\Delta L} + c_G C_L C^W_{\Delta L^\top}, \quad C_{\Delta G_L} = C^K_{\Delta G_L} + C^W_{\Delta G_L}$	(3.142)
$\bar{B} - B$	$C^K_{\Delta B} = (\sigma_B)^2 C^K_{\Delta G_L}, \quad C^W_{\Delta B} = (\sigma_B)^2 C^W_{\Delta G_L}, \quad C_{\Delta B} = C^K_{\Delta B} + C^W_{\Delta B}$	(3.143)
$\hat{\bar{N}} - \hat{N}$	$C^K_{\Delta\hat{N}} = c_J\left(C_L C^K_{\Delta B} + C^K_{\Delta L}\sigma_B\right) + c_{\mathrm{D}J} C_{\Delta K} C_L \sigma_B \gamma\delta^\tau$ $C^W_{\Delta\hat{N}} = c_J\left(C_L C^W_{\Delta B} + C^W_{\Delta L}\sigma_B\right), \quad C_{\Delta\hat{N}} = C^K_{\Delta\hat{N}} + C^W_{\Delta\hat{N}}$	(3.144)
$\hat{\bar{N}}^\top - \hat{N}^\top$	$C^K_{\Delta\hat{N}^\top} = c_{J^\top}\left(C_{L^\top} C^K_{\Delta B} + \sigma_B C^K_{\Delta L^\top}\right) + \sigma_B C_{L^\top} c_{(\mathrm{D}J)^\top} 2n C_{\Delta K}\gamma\delta^\tau$ $C^W_{\Delta\hat{N}^\top} = c_{J^\top}\left(C_{L^\top} C^W_{\Delta B} + \sigma_B C^W_{\Delta L^\top}\right), \quad C_{\Delta\hat{N}^\top} = C^K_{\Delta\hat{N}^\top} + C^W_{\Delta\hat{N}^\top}$	(3.145)
$\hat{\bar{S}} - \hat{S}$	$C^K_{\Delta\hat{S}} = \sigma_{\hat{N}^\top} c_\Omega c_{\mathrm{D}_z\phi_T} C^K_{\Delta\hat{N}} + \sigma_{\hat{N}^\top}\sigma_{\hat{N}}\left(c_{\mathrm{D}\Omega} c_{\mathrm{D}_z\phi_T} + c_\Omega c_{\mathrm{D}^2_z\phi_T}\right) C_{\Delta K}\gamma\delta^\tau + C^K_{\Delta\hat{N}^\top} c_\Omega c_{\mathrm{D}_z\phi_T}\sigma_{\hat{N}}$ $C^W_{\Delta\hat{S}} = \sigma_{\hat{N}^\top} c_\Omega c_{\mathrm{D}_z\phi_T} C^W_{\Delta\hat{N}} + C^W_{\Delta\hat{N}^\top} c_\Omega c_{\mathrm{D}_z\phi_T}\sigma_{\hat{N}}, \quad C_{\Delta\hat{S}} = C^K_{\Delta\hat{S}} + C^W_{\Delta\hat{S}}$	(3.146)
$\left\langle\bar{\mathcal{S}}\right\rangle^{-1} - \langle\mathcal{S}\rangle^{-1}$	$C^K_{\Delta\langle\mathcal{S}\rangle^{-1}} = \left(\sigma_{\langle\mathcal{S}\rangle^{-1}}\right)^2 C^K_{\Delta\hat{S}}, \quad C^W_{\Delta\langle\mathcal{S}\rangle^{-1}} = \left(\sigma_{\langle\mathcal{S}\rangle^{-1}}\right)^2 C^W_{\Delta\hat{S}}, \quad C_{\Delta\langle\mathcal{S}\rangle^{-1}} = C^K_{\Delta\langle\mathcal{S}\rangle^{-1}} + C^W_{\Delta\langle\mathcal{S}\rangle^{-1}}$	(3.148)
$\Delta\lambda$	$C_{\Delta\lambda} = C^K_{\Delta\lambda} + C^W_{\Delta\lambda}$	(3.149)
$\Delta\lambda^{-1}$	$C_{\Delta\lambda^{-1}} = (\sigma_{\lambda^{-1}})^2 C_{\Delta\lambda}$	(3.150)
$E_{\bar{K}}$	$C_{E_{\bar{K}}} = C^{KK}_{E_{\bar{K}}} + C^{KW}_{E_{\bar{K}}}$	(3.151)
$E_{\bar{W}}$	$C_{E_{\bar{W}}} = C^{KK}_{E_{\bar{W}}} + C^{WW}_{E_{\bar{W}}} + C^{KW}_{E_{\bar{W}}}$	(3.152)
$\mathcal{E}$	$C_\Delta = \max\Bigg\{ C_{E_{\mathrm{symp}P}}\gamma\delta^\tau, \ \frac{\max\left\{\gamma^2\delta^{2\tau}, C_{\Delta K}\right\}}{R - \lVert K - K_0\rVert_\rho}\gamma\delta^\tau, \ \frac{dC_{\Delta K}}{\sigma_{\mathrm{D}_\theta K} - \lVert\mathrm{D}_\theta K\rVert_\rho}\gamma\delta^{\tau-1}, \ \frac{\ell C_{\Delta K}}{\sigma_{\mathrm{D}_\varphi K} - \lVert\mathrm{D}_\varphi K\rVert_\rho}\gamma\delta^{\tau-1},$ $\frac{2nC_{\Delta K}}{\sigma_{(\mathrm{D}_\theta K)^\top} - \lVert(\mathrm{D}_\theta K)^\top\rVert_\rho}\gamma\delta^{\tau-1}, \ \frac{2nC_{\Delta K}}{\sigma_{(\mathrm{D}_\varphi K)^\top} - \lVert(\mathrm{D}_\varphi K)^\top\rVert_\rho}\gamma\delta^{\tau-1}, \ \frac{C_{\Delta W}}{\sigma_W - \lVert W\rVert_\rho}, \ \frac{2nC_{\Delta W}}{\sigma_{W^\top} - \lVert W^\top\rVert_\rho},$ $\frac{C_{\Delta B}}{\sigma_B - \lVert B\rVert_\rho}, \ \frac{C_{\Delta\hat{N}}}{\sigma_{\hat{N}} - \lVert\hat{N}\rVert_\rho}, \ \frac{C_{\Delta\hat{N}^\top}}{\sigma_{\hat{N}^\top} - \lVert\hat{N}^\top\rVert_\rho}, \ \frac{C_{\Delta\langle\mathcal{S}\rangle^{-1}}}{\sigma_{\langle\mathcal{S}\rangle^{-1}} - \left\lvert\langle\mathcal{S}\rangle^{-1}\right\rvert}, \ \frac{C_{\Delta\lambda}}{\sigma_\lambda - \lvert\lambda\rvert}\gamma\delta^\tau, \ \frac{C_{\Delta\lambda^{-1}}}{\sigma_{\lambda^{-1}} - \lvert\lambda^{-1}\rvert}\gamma\delta^\tau \Bigg\}$	(3.111)

Table E.7: Constants in Theorem 3.1 from Section 3.4.2

Constant	Label
$C_{\mathcal{E}} = \max\left\{ C_{E_{\bar{K}}} a^{2\tau}, C_{E_{\bar{W}}} \gamma\delta \right\}$	(3.153)
$\mathfrak{C}_{\Delta K} = \frac{a^{2\tau}}{a^{2\tau}-\kappa} C_{\Delta K}$	(3.155)
$\mathfrak{C}_{\tilde{\mathcal{U}}} = \gamma^2\delta^{2\tau} + \mathfrak{C}_{\Delta K}$	(3.156)
$\mathfrak{C}_{\Delta \mathrm{D}_\theta K} = \frac{a^{2\tau-1}}{a^{2\tau-1}-\kappa} d C_{\Delta K}$	(3.156)
$\mathfrak{C}_{\Delta \mathrm{D}_\varphi K} = \frac{a^{2\tau-1}}{a^{2\tau-1}-\kappa} \ell C_{\Delta K}$	(3.156)
$\mathfrak{C}_{\Delta (\mathrm{D}_\theta K)^\top} = \frac{a^{2\tau-1}}{a^{2\tau-1}-\kappa} 2n C_{\Delta K}$	(3.156)
$\mathfrak{C}_{\Delta (\mathrm{D}_\varphi K)^\top} = \frac{a^{2\tau-1}}{a^{2\tau-1}-\kappa} 2n C_{\Delta K}$	(3.156)
$\mathfrak{C}_{\Delta W} = \frac{a^{\tau}}{a^{\tau}-\kappa} C_{\Delta W}$	(3.156)
$\mathfrak{C}_{\Delta W^\top} = \frac{a^{\tau}}{a^{\tau}-\kappa} 2n C_{\Delta W}$	(3.156)
$\mathfrak{C}_{\Delta B} := \frac{a^{\tau}}{a^{\tau}-\kappa} C_{\Delta B}$	(3.156)
$\mathfrak{C}_{\Delta \hat{N}} = \frac{a^{\tau}}{a^{\tau}-\kappa} C_{\Delta \hat{N}}$	(3.156)
$\mathfrak{C}_{\Delta \hat{N}^\top} = \frac{a^{\tau}}{a^{\tau}-\kappa} C_{\Delta \hat{N}^\top}$	(3.156)
$\mathfrak{C}_{\Delta \langle \mathcal{S} \rangle^{-1}} = \frac{a^{\tau}}{a^{\tau}-\kappa} C_{\Delta \langle \mathcal{S} \rangle^{-1}}$	(3.156)
$\mathfrak{C}_{\Delta \lambda} = C_\lambda \frac{a^{2\tau}}{a^{2\tau}-\kappa} C_{\Delta \lambda}$	(3.156)
$\mathfrak{C} = \max\left\{ a^{2\tau} C_{\mathcal{E}},\ \mathfrak{C}_\Delta \gamma\delta^\tau \right\}$	(3.156)
$\mathfrak{C}_\Delta = \max\left\{ C_{E_{\mathrm{symp}P}} \gamma\delta^\tau,\ \frac{\mathfrak{C}_{\tilde{\mathcal{U}}}}{R}\gamma\delta^\tau,\ \frac{\mathfrak{C}_{\Delta \mathrm{D}_\theta K}}{\sigma_{\mathrm{D}_\theta K} - \lVert \mathrm{D}_\theta K \rVert_\rho}\gamma\delta^{\tau-1},\ \frac{\mathfrak{C}_{\Delta \mathrm{D}_\varphi K}}{\sigma_{\mathrm{D}_\varphi K} - \lVert \mathrm{D}_\varphi K \rVert_\rho}\gamma\delta^{\tau-1}, \frac{\mathfrak{C}_{\Delta (\mathrm{D}_\theta K)^\top}}{\sigma_{(\mathrm{D}_\theta K)^\top} - \lVert (\mathrm{D}_\theta K)^\top \rVert_\rho}\gamma\delta^{\tau-1},\ \frac{\mathfrak{C}_{\Delta (\mathrm{D}_\varphi K)^\top}}{\sigma_{(\mathrm{D}_\varphi K)^\top} - \lVert (\mathrm{D}_\varphi K)^\top \rVert_\rho}\gamma\delta^{\tau-1},\ \frac{\mathfrak{C}_{\Delta W}}{\sigma_W - \lVert W \rVert_\rho},\ \frac{\mathfrak{C}_{\Delta W^\top}}{\sigma_{W^\top} - \lVert W^\top \rVert_\rho}, \frac{\mathfrak{C}_{\Delta B}}{\sigma_B - \lVert B \rVert_\rho},\ \frac{\mathfrak{C}_{\Delta \hat{N}}}{\sigma_{\hat{N}} - \lVert \hat{N} \rVert_\rho},\ \frac{\mathfrak{C}_{\Delta \hat{N}^\top}}{\sigma_{\hat{N}^\top} - \lVert \hat{N}^\top \rVert_\rho},\ \frac{\mathfrak{C}_{\Delta \langle \mathcal{S} \rangle^{-1}}}{\sigma_{\langle \mathcal{S} \rangle^{-1}} - \left\lvert \langle \mathcal{S} \rangle^{-1} \right\rvert},\ \frac{\mathfrak{C}_{\Delta \lambda}}{\sigma_\lambda - \lvert \lambda \rvert}\gamma\delta^\tau,\ \frac{\mathfrak{C}_{\Delta \lambda^{-1}}}{\sigma_{\lambda^{-1}} - \lvert \lambda^{-1} \rvert}\gamma\delta^\tau, \right\}$	(3.156)

Bibliography

[AG03] E. L. Allgower and K. Georg. *Introduction to numerical continuation methods*, volume 45 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2003. Reprint of the 1990 edition Springer-Verlag, Berlin;.

[And98] M. A. Andreu. *The quasi-bicircular problem*. PhD thesis, Universitat de Barcelona, 1998.

[And03] M. A. Andreu. Preliminary study on the translunar halo orbits of the real earth–moon system. *Celestial Mechanics and Dynamical Astronomy*, 86(2):107–130, 2003.

[Arn62] V. I. Arnold. On the classical perturbation theory and the stability problem of planetary systems. In *Doklady Akademii Nauk*, volume 145, pages 487–490. Russian Academy of Sciences, 1962.

[Arn63a] V. I. Arnold. Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian. *Uspehi Mat. Nauk*, 18(5 (113)):13–40, 1963.

[Arn63b] V. I. Arnold. Small denominators and problems of stability of motion in classical and celestial mechanics. *Russ. Math. Surveys*, 18:85–192, 1963.

[Arn64] V. I. Arnold. Instability of dynamical systems with many degrees of freedom. *Dokl. Akad. Nauk SSSR*, 156:9–12, 1964.

[Arn89] V. I. Arnold. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989.

[BHM24] M. Barcelona, À. Haro, and J. M. Mondelo. Semianalytical computation of heteroclinic connections between center manifolds with the parameterization method. *SIAM Journal on Applied Dynamical Systems*, 23(1):98–126, 2024.

[BMO09] E. Barrabés, J. M. Mondelo, and M. Ollé. Numerical continuation of families of homoclinic connections of periodic orbits in the RTBP. *Nonlinearity*, 22(12):2901–2918, 2009.

[BMO13] E. Barrabés, J. M. Mondelo, and M. Ollé. Numerical continuation of families of heteroclinic connections between periodic orbits in a hamiltonian system. *Nonlinearity*, 26(10):2747, 2013.

[CCdlL13] R. Calleja, A. Celletti, and R. de la Llave. A KAM theory for conformally symplectic systems: efficient algorithms and their validation. *J. Differential Equations*, 255(5):978–1049, 2013.

[CCGdlL22] R. Calleja, A. Celletti, J. Gimeno, and R. de la Llave. Efficient and accurate KAM tori construction for the dissipative spin-orbit problem using a map reduction. *J. Nonlinear Sci.*, 32(1):1–40, 2022.

[CCH21] R. Calleja, M. Canadell, and À. Haro. Non-twist invariant circles in conformally symplectic systems. *Communications in Nonlinear Science and Numerical Simulation*, 96:105695, 2021.

[CdS01] A. Cannas da Silva. *Lectures on symplectic geometry*, volume 1764 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001.

[CFdlL03a] X. Cabré, E. Fontich, and R. de la Llave. The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces. *Indiana Univ. Math. J.*, 52(2):283–328, 2003.

[CFdlL03b] X. Cabré, E. Fontich, and R. de la Llave. The parameterization method for invariant manifolds. II. Regularity with respect to parameters. *Indiana Univ. Math. J.*, 52(2):329–360, 2003.

[CFdlL05] X. Cabré, E. Fontich, and R. de la Llave. The parameterization method for invariant manifolds. III. Overview and applications. *J. Differential Equations*, 218(2):444–515, 2005.

[CH14] M. Canadell and À. Haro. Parameterization method for computing quasi-periodic reducible normally hyperbolic invariant tori. In *F. Casas, V. Martínez (eds.),* Advances in Differential Equations and Applications, volume 4 of *SEMA SIMAI Springer Series*. Springer, 2014.

[CH17a] M. Canadell and À. Haro. Computation of Quasi-Periodic Normally Hyperbolic Invariant Tori: Algorithms, Numerical Explorations and Mechanisms of Breakdown. *J. Nonlinear Sci.*, 27(6):1829–1868, 2017.

[CH17b] M. Canadell and À. Haro. Computation of Quasiperiodic Normally Hyperbolic Invariant Tori: Rigorous Results. *J. Nonlinear Sci.*, 27(6):1869–1904, 2017.

[Chi79] B. V. Chirikov. A universal instability of many-dimensional oscillator systems. *Phys. Rep.*, 52(5):264–379, 1979.

[CHP25] R. Calleja, À. Haro, and P. Porras. Constructive approaches to qp-time-dependent kam theory for lagrangian tori in Hamiltonian systems. *arXiv preprint arXiv:2503.09740*, 2025.

[CJ00] E. Castellà and À. Jorba. On the vertical families of two-dimensional tori near the triangular points of the bicircular problem. *Celestial Mech. Dynam. Astronom.*, 76(1):35–54, 2000.

[CM08] E. Canalias and J. Masdemont. Computing natural transfers between Sun-Earth and Earth-Moon Lissajous libration point orbits. *Acta Astronautica*, 63(1-4):238–248, 2008.

[DdlLS06] A. Delshams, R. de la Llave, and T. M. Seara. A geometric mechanism for diffusion in hamiltonian systems overcoming the large gap problem: Heuristics and rigorous verification on a model: Heuristics and rigorous verification on a model. *Mem. Amer. Math. Soc.*, 179(844), 2006.

[DDP03] D. J. Dichmann, E. J. Doedel, and R. C. Paffenroth. The computation of periodic solutions of the 3-body problem using the numerical continuation software AUTO. In G. Gómez, M. W. Lo, and J. Masdemont, editors, *Libration Point Orbits and Applications*, pages 489–528. World Scientific, 2003.

[dlL01] R. de la Llave. A tutorial on KAM theory. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 175–292. Amer. Math. Soc., Providence, RI, 2001.

[dlLGJV05] R. de la Llave, A. González, À. Jorba, and J. Villanueva. KAM theory without action-angle variables. *Nonlinearity*, 18(2):855–895, 2005.

[dlLO06] R. de la Llave and A. Olvera. The obstruction criterion for non-existence of invarian circles and renormalization. *Nonlinearity*, 19(8):1907–1937, 2006.

[DTT17] D. A. Dei Tos and F. Topputo. On the advantages of exploiting the hierarchical structure of astrodynamical models. *Acta Astronautica*, 136:236–247, 2017.

[Dum14] H. S. Dumas. *The KAM story.* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2014. A friendly introduction to the content, history, and significance of classical Kolmogorov-Arnold-Moser theory.

[FdlLS09] E. Fontich, R. de la Llave, and Y. Sire. Construction of invariant whiskered tori by a parameterization method. I. Maps and flows in finite dimensions. *J. Differential Equations*, 246(8):3136–3213, 2009.

[FH24] J.-Ll Figueras and À. Haro. A modified parameterization method for invariant lagrangian tori for partially integrable hamiltonian systems. *Physica D: Nonlinear Phenomena*, 462:134127, 2024.

[FHL17] J.-Ll. Figueras, À. Haro, and A. Luque. Rigorous Computer-Assisted Application of KAM Theory: A Modern Approach. *Found. Comput. Math.*, 17(5):1123–1193, 2017.

[FMHdlLM24] Á. Fernández-Mora, À. Haro, R. de la Llave, and J. M. Mondelo. Whiskered tori in Hamiltonian systems: a unified approach. 2024. Under preparation.

[FMHM24a] Á. Fernández-Mora, À. Haro, and J. M. Mondelo. On the convergence of flow map parameterization methods in hamiltonian systems. *arXiv preprint arXiv:2411.11772*, 2024.

[FMHM24b] Á. Fernández-Mora, À. Haro, and J. M. Mondelo. Flow map parameterization methods for invariant tori in quasi-periodic hamiltonian systems. *SIAM Journal on Applied Dynamical Systems*, 23(1):127–166, 2024.

[GHdlL14] A. González, À. Haro, and R. de la Llave. Singularity theory for non-twist KAM tori. *Mem. Amer. Math. Soc.*, 227(1067):vi+115, 2014.

[GHdlL22] A. González, À. Haro, and R. de la Llave. Efficient and reliable algorithms for the computation of non-twist invariant circles. *Foundations of Computational Mathematics*, 22(3):791–847, 2022.

[GJJC+23] J. Gimeno, À. Jorba, M. Jorba-Cuscó, N. Miguel, and M. Zou. Numerical integration of high-order variational equations of odes. *Applied Mathematics and Computation*, 442:127743, 2023.

[GJNO22] J. Gimeno, À. Jorba, B. Nicolás, and E. Olmedo. Numerical computation of high-order expansions of invariant manifolds of high-dimensional tori. *SIAM Journal on Applied Dynamical Systems*, 21(3):1832–1861, 2022.

[GJSM01a] G. Gómez, À. Jorba, C. Simó, and J. Masdemont. *Dynamics and mission design near libration points. Vol. III*, volume 4 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2001. Advanced methods for collinear points.

[GJSM01b] G. Gómez, À. Jorba, C. Simó, and J. Masdemont. *Dynamics and mission design near libration points. Vol. IV*, volume 5 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2001. Advanced methods for triangular points.

[GKL+04] G. Gómez, W. S. Koon, M. W. Lo, J. E. Marsden, J. Masdemont, and S. D. Ross. Connecting orbits and invariant manifolds in the spatial restricted three-body problem. *Nonlinearity*, 17(5):1571–1606, 2004.

[GLMS01] G. Gómez, J. Llibre, R. Martínez, and C. Simó. *Dynamics and mission design near libration points. Vol. I*, volume 2 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2001. Fundamentals: the case of collinear libration points, With a foreword by Walter Flury.

[GM01] G. Gómez and J. M. Mondelo. The dynamics around the collinear equilibrium points of the RTBP. *Phys. D*, 157(4):283–321, 2001.

[GMM02] G. Gómez, J. Masdemont, and J. M. Mondelo. Solar system models with a selected set of frequencies. *Astronomy & Astrophysics*, 390(2):733–749, 2002.

[Gre75] J. M. Greene. A method for determining a stochastic transition. *J. Math. Phys*, 20(6):1183–1201, 1975.

[GSLM01] G. Gómez, C. Simó, J. Llibre, and R. Martínez. *Dynamics and mission design near libration points. Vol. II*, volume 3 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2001. Fundamentals: the case of triangular libration points.

[HBW97] K. C. Howell, R. S. Barden, and R. S. Wilson. Trajectory design using a dynamical systems approach with application to genesis. *Adv. Astron. Sci.*, 1997.

[HCF+16] À. Haro, M. Canadell, J.-Ll. Figueras, A. Luque, and J. M. Mondelo. *The parameterization method for invariant manifolds*, volume 195 of *Applied Mathematical Sciences*. Springer, [Cham], 2016. From rigorous results to effective computations.

[HdlL06a] À. Haro and R. de la Llave. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: numerical algorithms. *Discrete Contin. Dyn. Syst. Ser. B*, 6(6):1261–1300, 2006.

[HdlL06b] À. Haro and R. de la Llave. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: rigorous results. *J. Differential Equations*, 228(2):530–579, 2006.

[HdlL07] À. Haro and R. de la Llave. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: explorations and mechanisms for the breakdown of hyperbolicity. *SIAM J. Appl. Dyn. Syst.*, 6(1):142–207 (electronic), 2007.

[HdlLS12] G. Huguet, R. de la Llave, and Y. Sire. Computation of whiskered invariant tori and their associated manifolds: new fast algorithms. *Discrete Contin. Dyn. Syst.*, 32(4):1309–1353, 2012.

[Hén66] M. Hénon. Explorationes numérique du probleme restreint iv: Masses egales, orbites non periodique. *Bullettin Astronomique*, 3(1):49–66, 1966.

[Hen79] P. Henrici. Fast fourier methods in computational complex analysis. *SIAM Review*, 21(4):481–527, 1979.

[HM21] À. Haro and J. M. Mondelo. Flow map parameterization methods for invariant tori in hamiltonian systems. *Communications in Nonlinear Science and Numerical Simulation*, 101:105859, 2021.

[HPS70] M. W. Hirsch, C. C. Pugh, and M. Shub. Invariant manifolds. *Bull. Amer. Math. Soc.*, 76(5):1015–1019, 1970.

[JM99] À. Jorba and J. Masdemont. Dynamics in the center manifold of the collinear points of the restricted three body problem. *Phys. D*, 132(1-2):189–213, 1999.

[JZ05] À. Jorba and M. Zou. A software package for the numerical integration of ODEs by means of high-order Taylor methods. *Experiment. Math.*, 14(1):99–117, 2005.

[KAdlL21a] B. Kumar, R. L. Anderson, and R. de la Llave. High-order resonant orbit manifold expansions for mission design in the planar circular restricted 3-body problem. *Communications in Nonlinear Science and Numerical Simulation*, 97:105691, 2021.

[KAdlL21b] B. Kumar, R. L. Anderson, and R. de la Llave. Using GPUs and the parameterization method for rapid search and refinement of connections between tori in periodically perturbed planar circular restricted 3-body problems. *arXiv preprint arXiv*, 2109, 2021.

[KAdlL22] B. Kumar, R. L. Anderson, and R. de la Llave. Rapid and accurate methods for computing whiskered tori and their manifolds in periodically perturbed planar circular restricted 3-body problems. *Celestial Mech. Dynam. Astronom.*, 134(1):1–38, 2022.

[Kol54] A. N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR (N.S.)*, 98:527–530, 1954. Translated in p. 51–56 of *Stochastic Behavior in Classical and Quantum Hamiltonian Systems, Como 1977* (eds. G. Casati and J. Ford) Lect. Notes Phys. 93, Springer, Berlin, 1979.

[Lio55] J. Liouville. Note sur l'intégration des équations différentielles de la dynamique, présentée au bureau des longitudes le 29 juin 1853. *Journal de Mathématiques pures et appliquées*, 20:137–138, 1855.

[LV11] A. Luque and J. Villanueva. A KAM theorem without action-angle variables for elliptic lower dimensional tori. *Nonlinearity*, 24(4):1033–1080, 2011.

[Mañ78] R. Mañé. Persistent manifolds are normally hyperbolic. *Trans. Amer. Math. Soc.*, 246:261–283, 1978.

[Mat68] J. N. Mather. Characterization of Anosov diffeomorphisms. *Nederl. Akad. Wetensch. Proc. Ser. A 71 = Indag. Math.*, 30:479–483, 1968.

[MHO09] K. R. Meyer, G. R. Hall, and D. Offin. *Introduction to Hamiltonian Dynamical Systems and the N–Body Problem.* Springer–Verlag, 2nd edition, 2009.

[Mos62] J. Moser. On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II*, 1962:1–20, 1962.

[OS87] A. Olvera and C. Simó. An obstruction method for the destruction of invariant curves. *Phys. D*, 26(1-3):181–192, 1987.

[Pér15] F. Pérez. *Dynamical transport mechanisms in celestial mechanics and astrodynamics problems.* PhD thesis, Universitat de Barcelona, 2015.

[PHB19] J. A. Pérez-Hernández and L. Benet. PerezHz/TaylorIntegration.jl: TaylorIntegration v0.4.1, February 2019.

[Poi90] H. Poincaré. Sur le problème des trois corps et les équations de la dynamique. *Acta mathematica*, 13(1):A3–A270, 1890.

[Poi87] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste. Tome III.* Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics]. Librairie Scientifique et Technique Albert Blanchard, Paris, 1987.

[Rüs75] H. Rüssmann. On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In *Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*, pages 598–624. Lecture Notes in Phys., Vol. 38. Springer, Berlin, 1975.

[SBA+14] T. H. Sweetser, S. B. Broschart, V. Angelopoulos, G. J. Whiffen, D. C. Folta, M. K. Chung, S. J. Hatch, and M. A. Woodard. Artemis mission design. *The ARTEMIS Mission*, pages 61–91, 2014.

[Sie42] C. L. Siegel. Iterations of analytic functions. *Ann. of Math.*, 43:807–812, 1942.

[SM71] C. L. Siegel and J. Moser. *Lectures on Celestial Mechanics.* Springer-Verlag, New York, 1971. Translation by C. I. Kalme, Die Grundlehren der mathematischen Wissenschaften, Band 187.

[Sma65] S. Smale. *Diffeomorphisms with Many Periodic Points*, pages 63–80. Princeton University Press, Princeton, 1965.

[Sma67] S. Smale. Differentiable dynamical systems. *Bulletin of the American mathematical Society*, 73(6):747–817, 1967.

[SS74] R. J. Sacker and G. R. Sell. Existence of dichotomies and invariant splittings for linear differential systems. I. *J. Differential Equations*, 15:429–458, 1974.

[Vic67] S. Victor. *Theory of orbits. The Restricted Problem of Three Bodies.* Academic Press, 1967.

[Vil17] J. Villanueva. A new Approach to the Parameterization Method for Lagrangian Tori of Hamiltonian Systems. *J. Nonlinear Science*, 27:495–530, 2017.

[WZ03] D. Wilczak and P. Zgliczyński. Heteroclinic connections between periodic orbits in planar restricted circular three-body problem–a computer assisted proof. *Communications in mathematical physics*, 234(1):37–75, 2003.

[WZ05] D. Wilczak and P. Zgliczyński. Heteroclinic connections between periodic orbits in planar restricted circular three body problem. Part II. *Communications in mathematical physics*, 259:561–576, 2005.