HAL
open science

Hamiltonian Perturbation Theory on a Poisson Algebra. Application to a Throbbing Top and to Magnetically Confined Particles

Lorenzo Valvo

To cite this version:

Lorenzo Valvo. Hamiltonian Perturbation Theory on a Poisson Algebra. Application to a Throbbing Top and to Magnetically Confined Particles. Mathematical Physics [math-ph]. Aix Marseille University, 2019. English. NNT : . tel-02442078

HAL Id: tel-02442078

https://theses.hal.science/tel-02442078v1

Submitted on 16 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Aix-Marseille Université
Faculté des Sciences de Luminy
Ecole doctorale 352 - Physique et Sciences de la Matière
Centre de Physique Théorique - UMR 7332

Thèse de Doctorat
Mention Physique Théorique et Mathématique

Defended by

Lorenzo Valvo

Hamiltonian Perturbation Theory on a Poisson Algebra. Application to a Throbbing Top and to Magnetically Confined Particles

---

Thesis referred by:

Prof. Ugo LOCATELLI	Università degli Studi di Roma "Tor Vergata"
Prof. Philip MORRISON	The University of Texas at Austin

Defended on December 12th, 2019, before a jury composed of:

Prof. Dominique ESCANDE	PIIM, CNRS	
Prof. Daniela GRASSO	ISC-CNR and Politecnico di Torino	
Prof. Xavier LEONCINI	CPT, Aix-Marseille Université	
Prof. Ugo LOCATELLI	Università degli Studi di Roma "Tor Vergata"	
Prof. Philip MORRISON	The University of Texas at Austin	
Prof. Michel VITTOT	CPT, CNRS	**Thesis Supervisor**

Resumé de la These

La plupart des problèmes de physique sont formulés en termes de système dynamiques, et beaucoup de ces systèmes dynamiques ont une structure dite *de Poisson*. Deux exemples, considérés dans ce travail, sont un corps rigide non autonome (auquel nous nous référons comme "Toupie Pulsante"), et une particule classique relativiste chargée. Nous allons proposer une généralisation de la théorie de perturbation Hamiltonienne pour les algebres de Lie, qui donc peut être appliquée à l'étude de ce type de systèmes dynamiques.

Une *algebre de Lie*[1] est un espace vectoriel $\mathbb{V}$ sur un corps $\mathbb{K}$ et avec une opération bilinéaire, ici notée par un crochet [2] $\{\,\}$ (autour du premier argument), et satisfaisant deux propriétés (ici $A, B, C \in \mathbb{V}$),

$$\{A\}B = -\{B\}A \qquad \text{(alternance)} \tag{1}$$

$$\{\{A\}B\}C + \{\{B\}C\}A + \{\{C\}A\}B = 0 \qquad \text{(identité de Jacobi)} \tag{2}$$

Une *algèbre de Poisson* $\mathbb{V}$ est une algèbre de Lie avec aussi un produit associatif

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

ayant la propriété distributive par rapport au crochet de Lie

$$\{A \cdot B\}\, C = A \cdot \big(\{B\}\, C\big) + \big(\{A\}\, C\big) \cdot B$$

Les algèbres de Lie sont omniprésentes en physique, et dans le cas de la physique classique, elles sont pour la plupart des algèbres de Poisson: c'est le cas de la mécanique classique, Newtonienne et aussi relativiste, électrodynamique, relativité générale, mécanique des fluides et théorie cinétique.

En general, un système dynamique est un *flot* sur un ensemble, ce qui pour nous est un algèbre de Lie $\mathbb{V}$. Un flot est un groupe de transformations à un paramètre qui associe, à un élément donné $F \in \mathbb{V}$ (la condition initiale), un autre élément $F(t) \in \mathbb{V}$, pour toute valeur du paramètre t.

Un concept clé pour définir un flot sur $\mathbb{V}$ est la *derivation*: une application linéaire $\mathcal{D}\colon \mathbb{V} \to \mathbb{V}$ satisfaisant la propriété de Leibnitz:

$$\mathcal{D}\{A\}B = \{\mathcal{D}A\}\, B + \{A\}\mathcal{D}B$$

On peut construire une dérivation par le crochet de Lie: si on fixe $D \in \mathbb{V}$, alors $\{D\}$ est une dérivation.

Après avoir choisi une dérivation $\mathcal{H}$, on construit le flot

$$A(t) = e^{t\mathcal{H}} A_0 \equiv \sum_{n=0}^{\infty} \frac{t^n \mathcal{H}^n}{n!} A_0$$

[1] Dans ce qui suit, la référence principale est toujours le livre d'Arnold [5].

[2] Habituellement, dans la littérature, le crochet englobe les deux arguments: $\{A, B\} \in \mathbb{V}$. Notre notation insiste sur ce point: une fois le premier élément (ici A) fixé, on obtient une opération linéaire, communément notée dans la littérature par $\{A, \}$.

qui est la solution de

$$\begin{cases} \dfrac{dA}{dt} = \mathcal{H}A \\ A(0) = A_0 \end{cases}$$

Si $\mathcal{H} \equiv \{H\}$, alors H est appelé l'"Hamiltonien". On appelle un système de ce type "système Hamiltonien".

A titre d'exemple, on considère une particule chargée non-relativiste dans un champ magnétique statique. Ce problème est defini sur l'algèbre de Poisson $\mathcal{C}^\infty(\mathbb{X})$, ou $\mathbb{X}$ est une variété à 6 dimensions (l'*espace de phase*) avec coordonnées $\{\overline{\boldsymbol{x}}, \overline{\boldsymbol{v}}\}$. Si on denote par e la charge électrique et par $\boldsymbol{B}$ le champ magnétique, le crochet est[3]

$$\{A\}B = \frac{\partial A}{\partial \boldsymbol{v}}\frac{\partial B}{\partial \overline{\boldsymbol{x}}} - \frac{\partial A}{\partial \boldsymbol{x}}\frac{\partial B}{\partial \overline{\boldsymbol{v}}} + \frac{e}{c}\frac{\partial A}{\partial \boldsymbol{v}}[\![\boldsymbol{B}]\!]\frac{\partial B}{\partial \overline{\boldsymbol{v}}}$$

et la Hamiltonienne est $H_m \stackrel{\text{def}}{=} \frac{1}{2}m\overline{\boldsymbol{v}}\boldsymbol{v}$, où m est la masse de la particule. Les équations de mouvement résultantes sont

$$\dot{\boldsymbol{x}} = \{H_m\}\boldsymbol{x} \equiv \boldsymbol{v}\,, \qquad \dot{\boldsymbol{v}} = \{H_m\}\boldsymbol{v} \equiv \frac{e}{m}[\![\boldsymbol{v}]\!]\boldsymbol{B}$$

Un autre exemple d'intérêt pour cette thèse est la *Toupie* (également connue sous le nom de *Corps Rigide*). L'algèbre est $\{f\colon \mathbb{R}^3 \to \mathbb{R}\}$ et le crochet est defini par

$$\{F\}\,G = \overline{\boldsymbol{M}}\,[\![\partial_{\overline{\boldsymbol{M}}}F]\!]\,\partial_{\overline{\boldsymbol{M}}}G\,, \qquad \boldsymbol{M} \in \mathbb{R}^3$$

Ce crochet a une propriété intéressante: il donne zero quand il est evalué sur la fonction $\overline{\boldsymbol{M}}\boldsymbol{M}$ (appelé un *élément de Casimir*). Vu que le Casimir represente geometriquement la coordonnée radiale d'un système de coordonnées sphériques, il s'agit en fait d'un crochet entre des fonctions sur $\mathbb{S}^2$.

Un système Hamiltonien est dit *canonique* s'il est ecrit dans un ensemble de coordonnées de dimension pair, généralement désigné par $(p_i, q_i)_{i=1}^n \equiv (\overline{\boldsymbol{p}}, \overline{\boldsymbol{q}})$, tel que on peut ecrire le crochet comme

$$\{A\}B = \frac{\partial A}{\partial \boldsymbol{p}}\frac{\partial B}{\partial \overline{\boldsymbol{q}}} - \frac{\partial B}{\partial \boldsymbol{q}}\frac{\partial A}{\partial \overline{\boldsymbol{p}}}$$

Les coordonnées canoniques ont été considérées depuis longtemps la seule option pour les systèmes Hamiltoniens, donc la plupart des méthodes et des théorèmes de la mécanique classique (surtout la théorie des perturbations) nécessite une structure canonique. Un théorème de Darboux affirme qu'on peut introduire, sur une algèbre de Poisson quelconque, au moins localement, un ensemble des coordonnées canoniques, éventuellement apres avoir quotienté les Casimirs.

Mais les coordonnées non canoniques ont aussi leurs avantages.

[3]Les lettres en gras indiquent les vecteurs sur $\mathbb{R}^3$; nous éviterons plutôt la notation avec les indices.

Une surligne dénote une transposition par rapport à la métrique euclidienne, et le résultat est un covecteur. Lorsqu'un covecteur est suivi d'un vecteur, la contraction est implicite. La dérivation par un covecteur donne un vecteur, tandis que la dérivation par un vecteur résulte dans un covecteur.

On utilise des covecteurs plutôt que des vecteurs pour dénoter les coordonnées. La raison est liée au Théorème de Lie-Poisson.

Un double crochet carré $[\![\,]\!]$ indique le produit vectoriel sur $\mathbb{R}^3$; il s'agit en fait d'un crochet de Lie.

Si le système Hamiltonien ne vient pas de la transformation de Legendre d'un système lagrangien (comme c'est le cas dans la mécanique classique), la forme fonctionnelle des coordonnées canoniques peut ne pas être évidente. Ça arrive souvent dans les théories fluides, qui ont une structure de Poisson comme conséquence du théorème de Lie-Poisson.

On a déjà introduit la Toupie: c'est une structure de Poisson sur l'espace de phase $\mathbb{R}^3$ qui a un Casimir. Une fois que ce dernier est quotienté, on obtient une espace de phase reduit à $\mathbb{S}^2$ (il faut imaginer d'introduire des coordonnées spheriques sur $\mathbb{R}^3$); là, il est possible d'introduire des coordonnées canoniques. Cependant, elles ne sont pas définies sur toute la sphère (la géométrie différentielle nous montre qu'il est impossible de couvrir la sphère avec une seule carte).

L'utilisation de coordonnées canoniques peut impliquer "*the representation of physically interesting quantities by means of awkward mathematical constructions*", mais aussi introduire une dépendance de jauge non désirée. Ces deux aspects sont évidents dans l'exemple de la particule chargée. Lorsque dans la description non canonique de ce système on utilise le champ magnétique pour construire le crochet, la description canonique implique plutot le potentiel magnétique $\boldsymbol{A}$ (lié au champ magnétique par $\boldsymbol{B} = \mathbf{curl}\boldsymbol{A}$) dans l'Hamiltonien $H_{mc} = |\boldsymbol{p} - e\boldsymbol{A}|^2/(2m)$.

En fait, si on calcule $\{H_{mc}\}$ avec le crochet canonique on trouve

$$\dot{\boldsymbol{x}} = \{H_{mc}\}\boldsymbol{x} = \frac{1}{m}\Big(\boldsymbol{p} - e\boldsymbol{A}\Big) \equiv \boldsymbol{v}\,, \qquad \dot{\boldsymbol{p}} = \{H_{mc}\}\boldsymbol{p} \implies \dot{\boldsymbol{v}} = \frac{e}{m}[\![\boldsymbol{v}]\!]\mathbf{curl}\boldsymbol{A}$$

Ici on a introduit une liberté de jauge dans le système, et les calculs sont plus compliqués.

Ceci nous ramène au but de cette thèse: construire une théorie Hamiltonienne pour les Algèbres de Lie, qui peuvent donc être appliquées également aux systèmes Hamiltoniens non canoniques. Voici un bref aperçu des trois chapitres composant la thèse: la theorie des perturbations; son application à un Toupie non-autonome; la reduction de la dynamique d'une particule chargée.

1. Theorie Algèbrique des Perturbations Hamiltoniennes.

Une approche très efficace a la théorie des perturbations en mécanique classique est la théorie KAM. En general, un Hamiltonien H_0 est dit *integrable* s'il détermine une foliatione de l'espace de phase en tores invariants. Dans la théorie des perturbations, on cherches le tores d'un Hamiltonien quelconque H en le décomposant en $H = H_0 + V$, où H_0 est intégrable et V est une autre fonction (la perturbation), "plus petite" (par rapport à une quelque norme) que H_0. Le théorème KAM affirme qu'une ensemble de tores de H_0 n'est pas détruit, mais seulement déformé, par la perturbation (sous des hypothèses appropriées sur H). En fait il est possible de conjuger l'Hamiltonien H à un autre Hamiltonien, disons $\tilde{H}$, qui préserve un tore prescrit de H_0. La première étape est de construire une "transformation principale", pour supprimer la partie la plus forte de la perturbation: en effet, V est divisé en un une partie "bonne" (qui conserve le tore) et une partie "mauvaise", qui est quadratique en V. La deuxième étape, c'est l'iteration de cette transformation, dont on obtient un algorithme superconvergent pour construire $\tilde{H}$. Et on peut montrer que l'algorithme est bien défini pour un ensemble dense des tores.

Plusieurs variations du théorème de KAM ont été proposée: pour les difféomorphismes de l'anneau, pour les systèmes présympletiques, pour les champs de vecteurs; cependant, une version pour les systèmes de Poisson (qui ne nécessite pas des coordonnées angle-action) n'a pas encore été proposée. Dans le premier chapitre, on propose une version de la "transformation principale" pour une algèbre de Lie $\mathbb{V}$. Comme d'habitude, un système

dynamique est associé à une dérivation $\mathcal{H}$, qui est ensuite divisé en

$$\mathcal{H} \,=\, \mathcal{H}_0 \,+\, \{V\}$$

Le système non perturbé est associé à $\mathcal{H}_0$, qui peut être une dérivation interne ou externe. La perturbation par contre doit être interne, en fait elle est $\{V\}$. Comme il nous manque la notion de «tore», on la remplace par une sous-algèbre $\mathbb{B} \subset \mathbb{V}$, invariante par $\mathcal{H}_0$. Ensuite, on montre comment construire une transformation pour diviser la perturbation V en une partie qui préserve $\mathbb{B}$ et une autre, disons V_*, qui est quadratique en V. On a besoin d'un projecteur $\mathcal{R}\colon \mathbb{V} \to \mathbb{B}$ et un opérateur $\Gamma\colon \mathbb{V} \to \operatorname{der}\mathbb{V}$ tel que

$$[\mathcal{H}](\Gamma f) \,=\, \{(1-\mathcal{R})f\}\,, \forall f \in \mathbb{V}$$

Cette équation peut être appelée une «équation homologique» par analogie avec la théorie des perturbations en mécanique classique. Ensuite, on montre que [4]

$$e^{[\Gamma V]}(\mathcal{H} + \{V\}) = \mathcal{H}_* + \{V_*\}\,, \qquad \mathcal{H}_* \stackrel{\text{def}}{=} \mathcal{H} + \{\mathcal{R}V\}$$

Donc, la transformation $e^{\Gamma V}$ conjugue la (derivation) Hamiltonienne perturbée $\mathcal{H}$ à $\mathcal{H}_*$, modulo des corrections quadratique dans la norme de V. L'amélioration est que $\mathcal{H}_*$ préserve $\mathbb{B}$, comme $\mathcal{H}_0$.

Maintenant, $\Gamma V \in \operatorname{der}\mathbb{V}$, ce qui peut poser problèmes par rapport à la convergence de la série de Lie. On demande que l'algèbre $\mathbb{V}$ soit dotée d'une échelle de normes de Banach: une famille de normes labellisées par un parametre (disons s); alors, il est possible de borner Γ "avec perte", c'est-à-dire qu'on change s, et on perd un peu de régularité.

Cette approche à la théorie des perturbations est basée sur une methode introduite par Vittot dans [92]. Le développement le plus important de cette partie serait d'identifier un algorithme pour resoudre l'equation homologique avec $\mathcal{H}_*$ en place de $\mathcal{H}$; c'est à dire, de construire deux operateurs Γ_* et $\mathcal{R}_*$, et l'on pourrait appliquer iterativement la transformation induite. Une deuxième amélioration serait de fournir une échelle des normes de Banach générique pour les algèbres de Poisson. Avec ces éléments, il devrait être possible de construire quantitativement l'itération dans le cadre générique.

2. La Toupie Pulsante.

Dans le deuxième chapitre, on applique notre theorie algebrique des perturbations pour étudier la dynamique d'une Toupie non-autonome (ou *Toupie Pulsante*). On a dejà discuté la structure de Poisson non-canonique de la Toupie. La Toupie Pulsante a la meme structure, et elle peut étre étudiée comme une perturbation de la Toupie statique, qui est un système intégrable. Ce systeme est l'un des exemples les plus basiques de systeme de Poisson non-canonique, affichant des caractéristiques typiques (comme des Casimirs) même sur un espace de phase de basse dimension.

L'algèbre et le crochet de la Toupie sont

$$\mathbb{V}_{\text{Top}} \stackrel{\text{def}}{=} \mathcal{C}^\infty(\mathbb{R}^{3*} \to \mathbb{R}) \ni f, g \qquad \Longrightarrow \qquad \{f,g\}(\overline{\boldsymbol{M}}) = \overline{\boldsymbol{M}}\,[\![\partial_{\overline{\boldsymbol{M}}} f]\!]\,\partial_{\overline{\boldsymbol{M}}} g\,,$$

où $\boldsymbol{M} \in \mathbb{R}^3$ and $\overline{\boldsymbol{M}} \in \mathbb{R}^{3*} \equiv \overline{\mathbb{R}^3}$. Le Casimir est $\rho^2 \stackrel{\text{def}}{=} M_1^2 + M_2^2 + M_3^2$ et evidemment il a la proprieté $\{\rho\}F = 0\,, \forall F \in \mathbb{V}_{\text{Top}}$. L'energie (Hamiltonien) est $E = \frac{1}{2}\,\overline{\boldsymbol{M}}\mathsf{L}\boldsymbol{M}$, où $\mathsf{L} = \operatorname{diag}(I_1^{-1}, I_2^{-1}, I_3^{-1})$ une matrice symétrique appellée *tenseur d'inertie*. Physiquement, les trois valeurs propres I_i contiennent l'information sur la géométrie du corps.

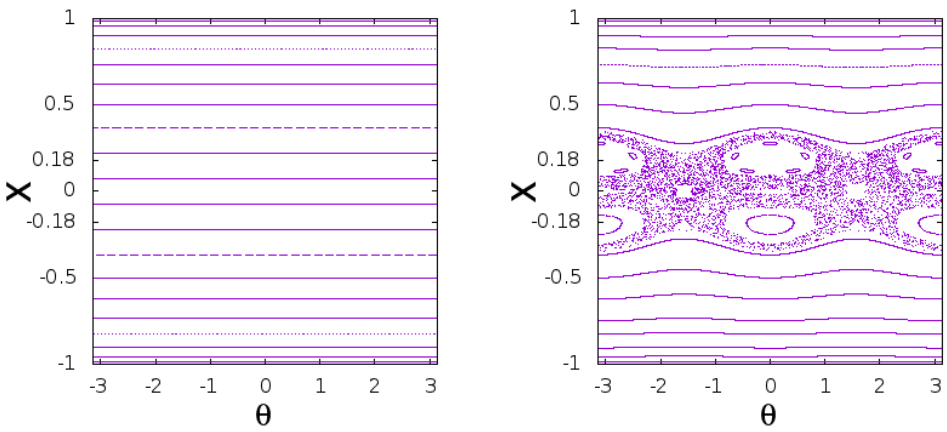
[4] $[\quad]$ est le commutateur entre les dérivations, qui est aussi un crochet de Lie.

Dans une Toupie Pulsante, les quantités I_i deviennent des fonctions du temps; physiquement, on peut imaginer un mécanisme qui bouge des morceaux internes du corps dans le temps. Pour décrire la Toupie pulsante, on considére l'algèbre

$$\mathbb{V}_{TT} \stackrel{\text{def}}{=} \mathcal{C}^\infty\Big(\begin{pmatrix}\mathbb{T}\\ \mathbb{R}\end{pmatrix} \to \mathbb{V}_{\text{Top}}\Big) \ni f = f\big(\overline{\boldsymbol{M}}, t\big)$$

avec le même crochet de la Toupie statique. On choisirait $\mathbb{T}$ (un tore) pour un système périodique (comme dans la première partie de ce chapitre), ou $\mathbb{R}$, pour un dépendance temporelle générique (comme dans la deuxième partie du chapitre).

2.1. On a etudié en première lieu une Toupie symétrique (dont $I_1 = I_2$) et périodique. Une Toupie statique symétrique admet une description canonique en coordonnées sphériques (ρ, φ, θ) (l'action étant $\cos(\varphi)$). La Toupie Pulsante est alors analogue à un système canonique *dégénéré* (un système canonique ne dépendant pas de toutes les actions). On prouve, pour ce système, une variante du théorème KAM, en appliquant de manière itérative notre formule de perturbation. Dans la figure suivante on montre deux portrait de phase pour une Toupie Pulsante symétrique. La figure sur la gauche se réfère au système sans perturbation, par contre, dans celle à droite, on ajoute une perturbation $V \propto M_2^2 \cos(t)$. On voit qu'une partie des orbites sur la gauche est detruit dans la figure à droite, d'autres ne sont que déformées: elles correspondent aux données initiales pour lesquelles la transformation $e^{\Gamma V}$ converge.



(a) $\epsilon = 0$ **(b)** $\epsilon = 0.1$

Figure 1: *Portrait de phase du type "time-2π" pour une Toupie Pulsante. Cet à dire qu'on represente ici seulement les iterations aux temps multiples entier de 2π. Ici $X = \cos(\varphi)$. A gauche, une Toupie imperturbé (et donc statique) avec $I_i = 1/i, \quad i = 1, 2, 3$. A droite, une Toupie avec perturbation, donc $I_2 = 1/\big(2\,(1 + 0.1\cos(t))\big)$*

2.2. Dans un deuxième moment, on a considéré une Toupie non symétrique et non périodique. Dans ce cas, nous n'avons pas introduit des coordonnées sur la sphere, mais on a travaillé directement avec le crochet sur $\mathbb{R}^3$, et ciblé la sous-algèbre $\mathbb{B} = \ker(\partial_t)$ de fonctions indépendantes du temps. Dans ce cas on voit une analogie aussi avec la théorie de Floquet, ce qui le rend encore plus intéressant.Cette fois on a montre que, pour des perturbations avec support compact $[0, T]$ dans le temps, notre "transformation principale" peut étre applique iterativement. Ça signifie que si la perturbation est suffisamment faible

ou elle agit pour un temps suffisamment court T, les valeurs des observables à l'avenir ne diffère pas beaucoup de leur valeur au moment $t = 0$.

3. La Particule Chargée.

Le problème de l'électrodynamique classique est de déterminer la dynamique des particules chargées sujettes à des champs électriques et magnétiques, selon les équations de Newton (en fait, leur version relativiste), et en même temps l'évolution des champs qui ont les particules comme sources, selon les équations de Maxwell.

Les équations de Maxwell-Vlasov sont un modèle mathématique pour ce problème; elles ont une structure de Poisson (non canonique), comme a eté montré par Morrison [63]. On pourrait donc résoudre le problème de l'électrodynamique classique par la théorie des perturbations (qu'on a construit) aux equations de Maxwell-Vlasov. Alors, la dynamique des particules dans des champs donnés est le système non perturbé, tandis que leur rétroaction sur les champs serait décrite de manière perturbative. Nous sommes guidé par l'observation empirique: en physique des plasmas, la dynamique d'un ensemble de particules chargées est dominée par le champ magnétique. En fait, on observe un mouvement giratoire s'enroulant autour des lignes de champ magnétique, avec une fréquence qui définit l'échelle de temps la plus élevée du problème. Le reste de la dynamique est beaucoup plus lent.

La première étape de ce plan ambitieux est la caractérisation des trajectoires non perturbées, pour definir la sous-algebre invariante $\mathbb{B}$. À cette fin, on considère deux "théories de réduction"[5] pour la dynamique des particules chargées.

3.1. La première est présentée ici pour la première fois: on l'appelle une "réduction Eulerienne". Dans cette approche, on essaye de reduire de moitié le nombre de degres de liberté d'une particule chargée; en particulier, par l'integration exacte des degrees de liberté de la velocité. C'est equivalent à donner la description fluide du mouvement d'une seule particule, d'où le nom «Eulerien». En effet, on calcule le champ de vitesse $\boldsymbol{V}$ associé à la force de Lorentz, de sorte que la dynamique complète est réduite à $\dot{\boldsymbol{x}} = \boldsymbol{V}(t, \boldsymbol{x})$. Si le champ obéit à une équation différentielle partielle, qu'on écrive sous la forme

$$\boldsymbol{E_c}(t, \boldsymbol{x}) \, + \, \big[\!\big[\boldsymbol{V}(t, \boldsymbol{x})\big]\!\big] \boldsymbol{B_c}(t, \boldsymbol{x}) \, = \, 0$$

alors l'équation de Newton pour la vitesse est satisfaite identiquement. Les champs $\boldsymbol{E_c}, \boldsymbol{B_c}$ sont les champs électromagnétiques avec des corrections liées à la dynamique de la particule. Cette description est équivalente à donner un changement de variables, où les degrés de vitesse de liberté sont remplacés par trois paramètres constants (disons $\boldsymbol{\alpha}$) qui détermine le champ de vitesse (donc il faudrait plutot ecrire $\boldsymbol{V}(t, \boldsymbol{x}, \boldsymbol{\alpha})$). Trouver des solutions pour $\boldsymbol{V}$ est important car, bien que les équations de Newton+Lorentz décrivent un problème fondamental de la physique, peu de résultats analytiques sont disponibles dans la littérature.

3.2. La deuxième méthode de réduction est la célèbre theorie du Centre Guide[6] qui est utilisée en physique des plasmas pour introduire un ensemble de coordonnées adaptée à la description des trajectoires «hélicoïdales» des particules qui tourne autour des lignes de champ magnétique. Avec ces coordonnées la dynamique est découplée dans le mouvement lent d'un point (le Centre Guide) qui suit les lignes de champ, et une rotation rapide de la particule réelle autour de ce point. Cependant, dans l'approche commune, les coordonnées du Centre Guide sont définies de manière perturbative Vu qu'il manque la forme fermée

[5] C'est à dire, une méthode pour réduire le nombre de degrés de liberté d'un système dynamique

[6] introduite en origine par Littlejohn [54], voir aussi la revue par Cary et Brizard [15]

de la transformation, on n'a pas de moyennes pour etudier sur quelles regions de l'espace de phase elle est effectivement bien definie.

On préfère suivre une nouvelle approche non perturbative à cette theoire, introduit recemment par Di Troia [22], [23]. Dans ce nouvelle methode on introduit un ensemble de coordonnées du Centre-Guide $\{T, \boldsymbol{X}, \zeta, \mu\, \varepsilon\}$, definies exactement (aussi dans le cadre relativiste):

- Le *Centre Guide* est le cadre de référence, ayant coordonnées $(T, \boldsymbol{X})$ dans le système du laboratoire, où une particule chargée se déplace sur une trajectoire périodique;
- La gyrophase ζ est l'angle dont cet trajectoire periodique est parametrisée;
- Le moment magnetique μ est propotionnel par (m/e) à l'action canonique associée à la gyrophase;
- La transformation de Centre-Guide $\mathcal{T}_{\mathrm{GC}}$ est choisie de façon que la nouvelle forme de Poincaré-Cartan est $\varpi_{GC} = P^0\, dT - \overline{\boldsymbol{P}}\, d\boldsymbol{X} + \frac{m}{e}\, \mu\, d\zeta$, avec champs $P^0(T, \boldsymbol{X}, \varepsilon, \mu)$ et $\boldsymbol{P}(T, \boldsymbol{X}, \varepsilon, \mu)$;

La velocité du Centre-Guide est encore un champ, donné par

$$m\gamma \boldsymbol{V}(T, \boldsymbol{X}, \varepsilon, \mu) = \boldsymbol{P}(T, \boldsymbol{X}, \varepsilon, \mu) - e\boldsymbol{A}(T, \boldsymbol{X})$$

et il se trouve qu'elle obeit encore à

$$\boldsymbol{E_c}(T, \boldsymbol{X}) + [\![\boldsymbol{V}]\!]\, \boldsymbol{B_c}(T, \boldsymbol{X}) = 0$$

La velocité du Centre-Guide est alors definie exactement comme la solution de cette equation. Dans ce travail on a derivé pour la première fois les equations du mouvement relativistes, et leur structure de Poisson. On trouve aussi des proprietées géometriques de ces coordonnées: le Centre Guide n'est pas une particule, dans le sens qu'il ne satisfait pas $W^0 = 1/\sqrt{1 - \overline{\boldsymbol{V}}\boldsymbol{V}}$ (où W^0 est l'energie cinetique), sauf si $\mu \to 0$.

L'inconvénient principal de ce nouvelle approche est que la forme explicite de la transformation $\mathcal{T}_{GC}$ n'a pas encore été fournie, de sorte que la relation entre les nouvelles coordonnées et les anciennes n'est pas connue; un développement intéressant serait de pouvoir construire cette carte, au moins dans certains exemples. Dans le même temps, la vitesse du Centre Guide est la solution d'une équation différentielle aux derivées partielles impliquant les expressions explicites des champs électromagnétiques: serait il possible d'élaborer des coordonnées du Centre Guide directement à partir de cette solution, de sorte que leur relation avec les coordonnées des particules pourrait être déduite a posteriori, ou même ignorée? De nouveau, le développement le plus intéressant serait de fournir des exemples explicites.

Contents

1 Introduction **1**

2 Perturbation Theory on a Lie Algebra **7**
2.1 Introduction . . . 7
2.2 The Algebraic "Main Step" Map . . . 8
2.2.1 A hint of the Iteration Procedure . . . 13
2.3 A brief historical survey . . . 13

3 The Throbbing Top **15**
3.1 About the Rigid Body (or Top) . . . 16
3.1.1 Lie-Poisson structure for the Rigid Body . . . 17
3.1.2 The Throbbing Top . . . 19
3.2 The symmetric case . . . 19
3.2.1 A scale of Banach norms for $\mathbb{V}_{\text{symm}}$ . . . 22
3.2.2 Application of Proposition 1 . . . 24
3.2.3 A KAM theorem for the Symmetric Throbbing Top . . . 31
3.2.4 Numerical experiments . . . 34
3.2.5 About the Diophantine condition . . . 36
3.3 The case of the Non-Symmetric Throbbing Top . . . 38
3.3.1 A new scale of norms . . . 38
3.4 Conclusion . . . 41

4 The Charged Particle **43**
4.1 The dynamical system . . . 44
4.2 The Eulerian reduction . . . 47
4.2.1 The Guiding Particle Solution . . . 49
4.2.2 The Poisson structure after the Eulerian Reduction . . . 49
4.2.3 Example: a constant uniform magnetic field . . . 51
4.2.4 Another example: the toroidal field . . . 52
4.2.5 Euler angles decomposition . . . 53
4.3 Non-perturbative Guiding Centre Theory . . . 55
4.3.1 The Guiding Centre Transform . . . 57
4.3.2 From Lagrangian to Hamiltonian Formulation . . . 59
4.3.3 Conservation of the Lagrangian . . . 62
4.3.4 A hint for future developments . . . 63
4.4 Conclusions . . . 63

5 Conclusions and Future Perspectives **65**

A Some mathematical tools **67**
A.1 About Lie Algebras 67
A.1.1 Metric structure on the algebra 69
A.1.2 Non-autonomous Hamiltonian Systems 70
A.1.3 The Lie-Poisson theorem 70
A.2 Complexification of a Lie Algebra 71
A.3 About $\mathbb{R}^n$ 72
A.3.1 Vector fields 72
A.3.2 Coordinate systems on $\mathbb{R}^n$ 73
A.3.3 A special case: $n = 3$ 73

B A WKB ansatz for equation (3.20) **75**

C The Lagrangian description of dynamics **79**
C.1 Relation with Hamiltonian formulation 79
C.2 Example: Relativistic Mechanics 80
C.3 The Poincaré-Cartan 1-form 81

1

Introduction

Most problems in Physics are set in terms of dynamical systems, and many of these dynamical systems have a so-called Poisson structure. Two examples, considered in this work, are a non-autonomous Rigid Body (to which we refer as "Throbbing Top"), and a classical relativistic charged particle. We will propose a generalization of Hamiltonian perturbation theory for a Lie algebraic setting, which can be applied to this type of dynamical systems.

A *Lie algebra*[1] is a vector space $\mathbb{V}$ over a field $\mathbb{K}$ and with a bilinear operation, here denoted by a bracket[2] $\{\,\}$ (embracing the first argument), and satisfying two properties (here $A, B, C \in \mathbb{V}$),

$$\{A\}B = -\{B\}A \tag{1.1}$$

$$\{\{A\}B\}C + \{\{B\}C\}A + \{\{C\}A\}B = 0 \tag{1.2}$$

As a consequence of (1.1) we say that the bracket is *alternating*; the second property (1.2) is instead called *Jacobi identity*.

A *Poisson algebra* $\mathbb{V}$ is a Lie algebra with also an associative product

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

having the distributive propert with respect to the Lie bracket

$$\{A \cdot B\}\, C = A \cdot \big(\{B\}\, C\big) + \big(\{A\}\, C\big) \cdot B$$

Lie algebras are ubiquitous in physics, and in the case of classical physics, they are most of times Poisson algebras: this is the case of classical mechanics [5], fluid dynamics and plasma physics [64], kinetic theory [76], special and general relativity [58].

Broadly speaking, a dynamical system is a *flow* on some set, which for us is a Lie algebra $\mathbb{V}$. A flow is a one-parameter group of applications associating, to a given element $F \in \mathbb{V}$ (the initial condition), another element $F(t) \in \mathbb{V}$, for any value of the parameter t. If $\mathbb{V} = \mathcal{C}^\infty(\mathbb{X} \to \mathbb{R})$, then the set $\mathbb{X}$ is called the *phase space*.

[1] In what follows, the main reference is always understood to be the book of Arnold [5].

[2] Usually in the literature the bracket embraces both arguments: $A, B \in \mathbb{V}$. Our notation emphatizes that, once the first element (here A) is fixed, we get a linear operation, commonly denoted in the literature by $\{A, \cdot\}$.

A key concept to define a flow on $\mathbb{V}$ is the *derivation*: a linear application $\mathcal{D}\colon \mathbb{V} \to \mathbb{V}$ which satisfy the Leibnitz property:

$$\mathcal{D}\{A\}B \;=\; \{\mathcal{D}A\}\,B \;+\; \{A\}\mathcal{D}B$$

A possible way of building a derivation is by the Lie bracket: once $D \in \mathbb{V}$ is fixed, $\{D\}$ is a derivation.

After choosing a derivation $\mathcal{H}$, we build the flow

$$A(t) \;=\; e^{t\mathcal{H}}A_0 \;\equiv\; \sum_{n=0}^{\infty} \frac{t^n \mathcal{H}^n}{n!} A_0 \tag{1.3}$$

which is the solution of

$$\begin{cases} \dfrac{dA}{dt} \;=\; \mathcal{H}A \\[2ex] A(0) \;=\; A_0 \end{cases} \tag{1.4}$$

in the sense that we can plug $A(t)$ from (1.3) into the system (1.4) and get an identity. If $\mathcal{H} \equiv \{H\}$, then H is called the "Hamiltonian". We call the system (1.4) a "Hamiltonian system".

As an example let us consider the charged particle in a static magnetic field, in the non-relativistic approximation. The phase space is the Poisson algebra $\mathcal{C}^\infty(\mathbb{X})$, $\mathbb{X}$ being a 6-dimensional manifold with coordinates $\{\overline{\boldsymbol{x}}, \overline{\boldsymbol{v}}\}$. Denoting by e the electric charge and by $\boldsymbol{B}$ the magnetic field, the bracket is[3]

$$\{F\}G \;=\; \frac{\partial F}{\partial \boldsymbol{v}}\frac{\partial G}{\partial \overline{\boldsymbol{x}}} \;-\; \frac{\partial F}{\partial \boldsymbol{x}}\frac{\partial G}{\partial \overline{\boldsymbol{v}}} \;+\; \frac{e}{c}\frac{\partial F}{\partial \boldsymbol{v}}\big[\!\big[\boldsymbol{B}\big]\!\big]\frac{\partial G}{\partial \overline{\boldsymbol{v}}} \tag{1.5}$$

and the Hamiltonian is $H_m \stackrel{\text{def}}{=} \frac{1}{2}m\overline{\boldsymbol{v}}\boldsymbol{v}$, where m is the mass of the particle. The resulting equations of motion are

$$\dot{\boldsymbol{x}} \;=\; \{H_m\}\boldsymbol{x} \;\equiv\; \boldsymbol{v}\,, \qquad \dot{\boldsymbol{v}} \;=\; \{H_m\}\boldsymbol{v} \;\equiv\; \frac{e}{m}\big[\!\big[\boldsymbol{v}\big]\!\big]\boldsymbol{B} \tag{1.6}$$

Another example of interest for this dissertation is the Top (also known as the Rigid Body). The phase space is $\{f\colon \mathbb{R}^3 \to \mathbb{R}\}$ and the bracket is defined by

$$\{F\}\,G \;=\; \overline{\boldsymbol{M}}\,\big[\!\big[\partial_{\overline{\boldsymbol{M}}}F\big]\!\big]\,\partial_{\overline{\boldsymbol{M}}}G\,, \qquad \boldsymbol{M} \in \mathbb{R}^3 \tag{1.7}$$

This bracket has an interesting property: it preserves the function $\overline{\boldsymbol{M}}\boldsymbol{M}$ (called a *Casimir element*), independently of the second argument. As the Casimir has the meaning of "radial coordinate" for a spherical coordinate system, we may in fact think of a bracket between functions on $\mathbb{S}^2$.

[3] Bold letters denote vectors on $\mathbb{R}^n$; we will rather avoid index notation.

An overline denotes transposition with respect to the Euclidean metric, and the result is a covector. When a covector is followed by a vector, contraction is intended. Derivation by a covector gives a vector, while derivation by a vector results in a covector.

There's a deep reason, related to the Lie-Poisson theorem, for which we use a covector rather than a vector to denote a set of coordinates; this is discussed in the appendix A.1.3.

A double square bracket $\big[\!\big[\,\cdot\,\big]\!\big]$ denotes the cross vector product. We denote this operation by a bracket because it is, in fact, a Lie bracket.

We give a resume of the notation and of this topics in appendix A.3.

A Hamiltonian system is called *canonical* if it has an even-dimensional set of coordinates, usually denoted by $(p_i, q_i)_{i=1}^n \equiv (\overline{\boldsymbol{p}}, \overline{\boldsymbol{q}})$, which put the bracket in the form

$$\{A\}B = \frac{\partial A}{\partial \boldsymbol{p}}\frac{\partial B}{\partial \overline{\boldsymbol{q}}} - \frac{\partial B}{\partial \boldsymbol{q}}\frac{\partial A}{\partial \overline{\boldsymbol{p}}} \tag{1.8}$$

Canonical coordinates have been considered for long time the only option for Hamiltonian systems, so that many efficient methods and theorems of classical mechanics (above all, perturbation theory) require a canonical structure. A theorem by Darboux [5], [53] states that any Poisson algebra can be reduced, at least locally, to a canonical one, eventually by quotienting out some dimensions[4].

But non-canonical coordinates have also their advantages.

If the dynamical system doesn't come from the Legendre transform of a Lagrangian one (is the case for classical mechanics), the functional form of canonical coordinates may not be evident. This happens often in fluid mechanics, where the Poisson structure is a consequence of the Lie-Poisson theorem (see for instance [65]).

We already mentioned the Top (that will also be the subject of chapter 3): it is a Poisson structure for functions on $\mathbb{R}^3$ which has a Casimir. Once the latter is quotienting out, we get a bracket for functions on $\mathbb{S}^2$, and it is possible to introduce canonical coordinates on this set. However, they are not defined over the whole $\mathbb{S}^2$ at once (it is known from differential geometry that it is not possible to cover the whole sphere with a single chart).

The use of canonical coordinates may involve "*the representation of physically interesting quantities by means of awkward mathematical constructions*" [55], or it may introduce an undesired gauge dependence. Both of these aspects are evident in the example of the charged particle. Indeed, while the non-canonical description of this system involves the magnetic field in the bracket (as seen in equation (1.5)), the canonical description involves the magnetic potential $\boldsymbol{A}$ (related to the magnetic field by $\boldsymbol{B} = \mathbf{curl}\boldsymbol{A}$) in the Hamiltonian $H_{mc} = |\boldsymbol{p} - e\boldsymbol{A}|^2/(2m)$. Indeed by computing $\{H_{mc}\}$ with the bracket (1.8), then

$$\dot{\boldsymbol{x}} = \{H_{mc}\}\boldsymbol{x}\,, \qquad \dot{\boldsymbol{p}} = \{H_{mc}\}\boldsymbol{p}$$

are equivalent to equations (1.6); however, the system is no more gauge free, and the computations are rather cumbersome.

This leads us back to the purpose of this thesis: to build a Hamiltonian perturbation theory for Lie algebras, which can thus be applied also to non-canonical Hamiltonian systems.

A very efficient approach to perturbation theory in classical mechanics is KAM theorem. In classical mechanics, a Hamiltonian H_0 is called *integrable* if it determines a foliation of the phase space into invariant tori. In perturbation theory one aims at finding the tori of a generic Hamiltonian H by decomposing it as $H = H_0 + V$, where H_0 is integrable, and V is another function (the perturbation), "smaller" (with respect to some norm) than H_0. The KAM theorem states that a set of tori of H_0 are not destroyed but only deformed by the perturbation (under suitable hypothesis on H). This is shown by conjugating the Hamiltonian H to another Hamiltonian, say $\tilde{H}$, which preserves some prescribed torus of H_0. A "main step map" is built, to delete the strongest part of the perturbation: indeed, V is split into a "nice" part (which preserves the torus) and a "bad"

[4] We are referring to the Casimir elements.

part, which is of order $|V|^2$. By iterating this map we get superconvergent algorithm to build $\tilde{H}$. And it can be shown that the tori for which the algorithm is well defined form a dense set.

Adapted versions of the KAM theorem have been provided for classical mechanics without action-angle variables [20], presymplectic mechanics [2], vector fields [8]; however, a version adapted for Poisson systems is still missing. In **chapter 2** we propose a version of the "main step map" set on a Lie-algebra $\mathbb{V}$. The unperturbed system is associated to a derivation $\mathcal{H}_0$. As we are missing the notion of "torus", we replace it with the more general one of a subalgebra $\mathbb{B} \subset \mathbb{V}$, invariant by $\mathcal{H}_0$. Then we add a perturbation in the form of an inner derivation, $\mathcal{H}_0 \mapsto \mathcal{H} = \mathcal{H}_0 + \{V\}$, for some $V \in \mathbb{V}$. We then show how to build a map to split the perturbation V into a part which preserves $\mathbb{B}$ and another, say V_*, which is quadratic in the magnitude of V. The map is defined by a Lie series and we pose some minimal hypothesis to achieve its convergence. This is a generalization of an approach introduced by Vittot in [92].

In chapter 3 we apply our perturbation formula to study the dynamics of a non autonomous Top: we call this system a *Throbbing Top*. As already discussed, the Top has a non-canonical Poisson structure; however, the Top is also an integrable system, and so out of scope for perturbation theory. This is no more true if a parametrical time-dependence is added: indeed in this case the integrability is lost. We consider two particular cases:

1. a symmetric[5] Top with periodic time-dependence. In this case a chart on $\mathbb{S}^2$ can be introduced, in which the Lie bracket is reduced to a nearly canonical form. Moreover the system is described by three coordinates: two of them are angles and don't appear in the stationary Hamiltonian. Then we are nearly in a classical mechanical system; indeed we show that our perturbation formula can be iteratively applied to provide a full KAM theorem. This shows also that our method is compatible with the existing theory;

2. a non-symmetric Top with non-periodic time-dependence: in this case we use the non-canonical structure (1.7) and target the algebra of time-independent functions. This case is more analogous to Floquet theory rather than KAM theory, which makes it even more interesting because, to our knowledge, a Floquet theory for non-periodic dynamical systems has not yet been proposed. We show that, for perturbations with compact support $[0, T]$ in time, the algorithm converges. This means that if the perturbation is sufficiently weak or applied for a sufficiently short time T, the values of the observables in the future do not differ much from their value at time $t = 0$.

Another physical problem to which we may apply perturbation theory is classical electrodynamics. We already discussed the non-canonical Poisson structure of the charged particle. This represents half of "the problem" of classical electrodynamics: the charged particles evolve according to Newton equations, with forces given by the fields; the second part of the problem is to determine the evolution of the fields which have the perticles themselves as sources, according to Maxwell equations.

[5]For those acquainted with the topic, we call symmetric a Top with two equals momenta of inertia.

One mathematical model for this problem are the Maxwell-Vlasov equations. This dynamical system has a (non-canonical) Poisson structure (this was shown by Morrison: see [63] and also [64]). So we may attempt to achieve the solution of the problem of classical electrodynamics by applying perturbation theory to the Maxwell-Vlasov equations. The dynamics of the particles in given fields is the unperturbed system, while field fluctuations may be described perturbatively. We are led by the empirical observation that the dynamics of a set of charged particles is always dominated by the magnetic field: we observe a gyratory motion wrapping around the magnetic field lines, with a frequency that defines the highest time-scale of the problem. The rest of the dynamics is much slower.

The first step of this ambitious plan, is the characterization of the unperturbed trajectories, which would define the invarian sualgebra $\mathbb{B}$. For this purpose, **in chapter 4** we consider two reduction theories[6] for the charged particle dynamics:

- The first one is presented here for the first time: we call it "Eulerian reduction". In this approach we try to fix half of the degrees of freedom of the charged particle motion. The resulting picture is like the fluid description of the single particle motion, hence the name "Eulerian". Indeed, we propose the velocity field $\boldsymbol{V}(\boldsymbol{x}, t)$ associated to the Lorentz force, so that the full dynamics is reduced to $\dot{\boldsymbol{x}} = \boldsymbol{V}(\boldsymbol{x}, t)$. This picture may be thought as a change of variables to a new system of coordinates where the velocity degrees of freedom have been replaced by three constants parameters determining the field $\boldsymbol{V}$. We computed the Poisson structure in the new Picture. Finally, we tried also to figure out some method to compute $\boldsymbol{V}$. These results are important because, despite the Newton+Lorentz equations describe a very fundamental problem of physics, nearly no analytical results are available in the literature;

- The second reduction method is the well known Guiding Centre theory[7] which is used in plasma physics to introduce a set of coordinates adapted to describe the "helicoidal" trajectories winding around the magnetic field lines, as commonly observed both in nature and in laboratory experiments. With these coordinates the dynamics is decoupled in the slow motion of a point (the Guiding Centre) stick to the field lines, and a fast rotation around it. However, in the common approach, the Guiding Centre coordinates are defined perturbatively and only up to a finite number of orders. As a consequence, they are not suited for an *exact* characterization of the trajectories. Instead we will consider a new non-perturbative approach to the theory, which has been proposed by Di Troia [22], [23]. In this approach the new coordinates are defined exactly, also in the relativistic setting; however, the equations of motion have been derived only in the Newtonian approximation. We derive the relativistic equations of motion as well as their Poisson structure. We also observe some geometrical consequences of the result: in particular, the gyrophase appears as a sort of "internal clock", in the same way as proper time.

[6] By the term "reduction theory", we refer to a method to reduce the number of degrees of freedom of a dynamical system.

[7] see for instance the recent review by Cary and Brizard [15] for a large overview on both the established version, originally proposed by Littlejohn [54], and its historical development.

About the notation

We will try to keep using a coherent notation throughout the manuscript.

In footnote 3 we've given an overview on our notation on vectors. It's rather unusual but very practical, and we beg pardon for the time that the reader will spend on getting used to it. We recall that **B**old **L**etters denote vectors.

As a general rule, $\mathbb{D}$oubled $\mathbb{L}$etters will denote sets, while $\mathcal{C}$alligrafic $\mathcal{L}$etters will denote linear applications (with the exception of the set $\mathcal{C}^\infty$, which could be confused with the set of complex numbers $\mathbb{C}$)

Instead Sans Serif letters will denote matrices; however, skew-symmetric matrices can be also denoted by calligrafic letters as they are derivations if built as Lie brackets with respect to the cross product.

The use of the letters in chapter 3 is indepedent of that of chapter 4; this means that, for instance, a quantity called ω in one chapter won't have anything to do with another quantity called ω in the other chapter. Instead the letters introduced in the first chapter 2 on the general perturbation theory will have the same meaning in the whole manuscript.

2

Perturbation Theory on a Lie Algebra

" Striclty speaking, there does not exists a KAM Theorem ("KAM" standing for the initials of A.N. Kolmogorv, V.I. Arnold and J.K. Moser), however, normally, it refers to (variations of) Kolmogorov's theorem. Here, we follow this tradition. " [17]

In this chapter we propose a method to perform perturbation theory for dynamical systems on Lie algebras. We call them Hamiltonian systems, as they are a generalization of Hamilton equations of classical mechanics. A perturbation theory for Hamiltonian systems was proposed by Vittot in [92]: in his method, a perturbed Hamiltonian is conjugated to a new one, preserving the kernel of the original Hamiltonian. The conjugation was built by a single Lie transform, but the formal proof of its convergence required some very strong hypothesis (the boundedness, in the sense of Banach, of some derivation operators). Here we develop further this approach, with two main novelties. First, we target any subalgebra of the phase space, invariant by the flow of the unperturbed Hamiltonian. The conjugation is built by a kind of superconvergent algorithm[1], as is commonly done in KAM theory.

This chapter is organized in three sections[2]: an informal introduction, which follows; a more technical description of our method, in section 2.2, and finally a (very succint) historical survey on KAM theory, in section 2.3.

2.1 Introduction

In classical mechanics, the phase space is a symplectic manifold $\mathbb{X}$, and we consider the Lie algebra of analytical functions defined on it[3]. An Hamiltonian H_0 is called integrable if it determines a foliation of phase space into invariant tori. In perturbation theory, one looks for the tori of $H = H_0 + V$, where V is called a perturbation of H_0. In the KAM method, the problem is tackled in an indirect way: after choosing one torus of H_0, the perturbed Hamiltonian H is conjugated to a new one, say $\tilde{H}$, preserving the chosen torus. Moreover it can be shown that the tori of H_0 for which the conjugation exists is dense.

[1]Even though, unfortunately, we were not able to push the formal proof of convegence to all orders, as we will see.

[2]We think that the original material in this chapter is presented in a sufficiently compact way that a concluding section won't be needed.

[3]The bracket being given by (1.8)

The conjugation is built algorithmically, by iteration of a well chosen canonical transform. The transform isn't required to delete all the perturbation at once. Instead, at each step, the perturbation is split into what preserves the original torus, and a "remainder". Then, only part of this "remainder" is deleted: the part linear in the perturbation strength[4]. Terms of higher order are kept: they will be deleted in the following steps, when the procedure is iterated.

Once the canonical transform has been applied and the perturbation has been reduced in strength, one ends up with a new (*reduced*) dynamical system: and this reduced dynamical system satisfies again the orginal hypothesis; however, if the original magnitude of V was ϵ, the new perturbation has magnitude ϵ^2. So the "main step map" can be applied again; after this passage, the perturbation will be of order ϵ^4. The iteration of this process generates the new Hamiltonian $\tilde{H}$.

A key element to this theory are the action-angle coordinates: here denoted by $\{A_i, \vartheta_i\}_{i=1}^n$, they are a preferred set of canonical coordinates for the phase space $\mathbb{X}$, in which the flow takes the simple form

$$\dot{A}_i \,=\, 0\,, \qquad \dot{\vartheta}_i \,=\, \frac{\partial H(A)}{\partial A_i} \equiv \omega_i\,, \qquad i = 1, \ldots, n$$

So, these coordinates are needed to define a notion of integrability, or equivalently, the foliation in tori (a torus is a level surface of the A_is). We may say that action-angle variables encode the informations we have on the unperturbed flow. To avoid the use of these coordinates, we will ask for a subalgebra $\mathbb{B}$ preserved by the unperturbed flow. This replaces the notion of integrability[5]

2.2 The Algebraic "Main Step" Map

Given a Lie algebra $\mathbb{V}$ with bracket $\{\quad\}$, to define a dynamical system on it we introduce the space $\operatorname{der}\mathbb{V}$ of *derivations* of $\mathbb{V}$,

$$\operatorname{der}\mathbb{V} \stackrel{\text{def}}{=} \Big\{ \mathcal{D}\colon \mathbb{V} \to \mathbb{V} \quad \text{s.t.} \quad \mathcal{D}\{F\}G \,=\, \{\mathcal{D}F\}G + \{F\}\mathcal{D}G, \quad \forall F, G \in \mathbb{V} \Big\} \tag{2.1}$$

Inner derivations are those for which $\mathcal{D} \equiv \{D\}$. Derivations which are not inner are called outer.

Once a derivation $\mathcal{H}$ is fixed, a dynamical system is given by

$$\dot{f} \,=\, \mathcal{H}f \tag{2.2}$$

A perturbation of this dynamical system is

$$\mathcal{H} \,\to\, \mathcal{H} + \{V\}$$

In our perturbation theory, the original derivation $\mathcal{H}$ can be of any type, also outer, but the perturbation must be inner.

[4]Sometimes a parameter ϵ is added in front of the perturbation; then the "perturbation strength" is ϵ. Instead one may think of ϵ simply as the norm of V; we follow this second approach.

[5]When action-angles are available, we may of course use of them to define the invariant subalgebra

$$\mathbb{B} \,=\, \big\{ f \in \mathcal{C}^\infty(\mathcal{M}),\, \{f\}A_i \,=\, A_i\,,\, i = 1, \ldots, n \big\}$$

Another tool that will be needed is a scale of norms (described in detail in A.1.1). A norm is a function $\mathbb{V} \to \mathbb{R}_+$, with two properties of triangular inequality and linearity:

$$\|F+G\| \leq \|F\| + \|G\|\,, \qquad \|\lambda F\| = |\lambda|\|F\|\,; \qquad \forall F \in \mathbb{V}, \lambda \in \mathbb{K}$$

A scale of norms is a sequence of norms $\| \quad \|_{r\, r\in\mathbb{I}}$, where $\mathbb{I}$ is an ordered set, such that

$$\|F\|_{r_1} \leq \|F\|_{r_2}\,, \quad \forall r_1, r_2 \in \mathbb{I}\,,\ r_1 \leq r_2$$

In the next proposition, we propose a formula, henceforth called "the main step formula", to conjugate the perturbed derivation $\mathcal{H} + \{V\}$ to another one, where the perturbation V is split into a term preserving the same subalgebra of $\mathcal{H}$, plus a new perturbation V_*, smaller than V. The conjugation is defined by a Lie series. By assuming that there exist a scale of Banach norms (see appendix A.1.1) defined on $\mathbb{V}$, we give minimal hypothesis to ensure the convergence of the series. In figure 2.1 we provide a scheme of all the quantities involved in our formula.

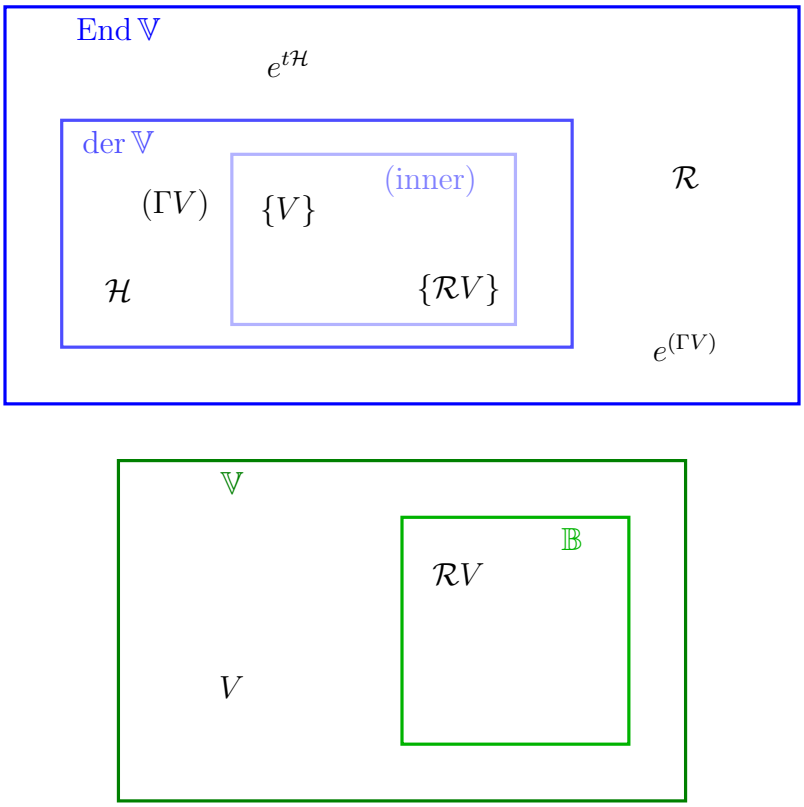


Figure 2.1: *A scheme of all the elements of our perturbation formula, and their domain of definition. The two most important sets are the Lie algebra $\mathbb{V}$, in green, and the space of endomorphisms $\operatorname{End}\mathbb{V} \cong \mathbb{V} \otimes \mathbb{V}^*$. $\mathbb{B}$ is a sub-Lie algebra of $\mathbb{V}$. Inside the endomorphisms there are many subspaces, and we are interested in particular in two of them, which are the derivations and the projectors. The latter contains $\mathcal{R}\colon \mathbb{V} \to \mathbb{B}$. The derivations are those endormorphisms which satisfy* (2.1)*, and they contain the subset of inner derivations, which are of type $\{f\}, f \in \mathbb{V}$. Derivations which are not inner are called outer: for instance we have $\mathcal{H}$, which may be equal to $\{H\}$ (as in the case of classical mechanics), for some $H \in \mathbb{V}$, but not necessarily. The application Γ takes an element of $\mathbb{V}$ and maps it to an element of $\operatorname{End}\mathbb{V}$ and it is a key element in our KAM formula. Finally, the operator $\mathcal{G}$ is an Endormorphism and it can replace Γ in the sense that it may be possible to have $\Gamma f \equiv \{\mathcal{G}f\}, f \in \mathbb{B}$.*

Proposition 1. *Let $\mathbb{V}$ be a Lie algebra with bracket $\{\quad\}$ and with a scale of Banach norms $\{\|\quad\|_r\}_{r\in\mathbb{I}}$. Let $\mathcal{H}\in\operatorname{der}\mathbb{V}$ and $\|V\|_s<\infty$ for some $s\in\mathbb{I}$.*

Assume that there exist

(i) a Lie sub-algebra $\mathbb{B}\subseteq\mathbb{V}$ such that $\mathcal{H}\mathbb{B}\subseteq\mathbb{B}$, and [6] $V\notin\mathbb{B}$;

(ii) a projector $\mathcal{R}\colon\mathbb{V}\to\mathbb{B}$ and a function $\Xi\colon\mathbb{I}\to\mathbb{R}_+$ such that $\forall\delta\in\mathbb{I}$

$$\|\mathcal{N}V\|_{s-\delta}\;<\;\Xi(\delta)\|V\|_s\qquad\mathcal{N}\stackrel{def}{=}\mathbb{I}-\mathcal{R}\tag{2.3}$$

(iii) an operator $\Gamma\colon\mathbb{V}\to\operatorname{der}\mathbb{V}$ and a function $\Lambda\colon\mathbb{I}\times\mathbb{I}\to\mathbb{R}_+$ such that $\forall d,\delta\in\mathbb{I}$, $(d+\delta)<s$ and $\forall F\in\mathbb{V}_{s-\delta}$

$$[\mathcal{H},\Gamma F]=\{\mathcal{N}F\}\tag{2.4}$$

$$\|(\Gamma V)F\|_{s-\delta-d}\;\le\;\Lambda(d,\delta)\,\|V\|_s\,\|F\|_{s-\delta}\tag{2.5}$$

$$\epsilon_\delta=\frac{1}{2}\sup_{n\in\mathbb{N}}\left(\frac{1}{n!}\prod_{j=1}^{n}\Lambda\Big(\frac{\delta}{n},\,\frac{(j-1)\delta}{n}\Big)\right)^{-1/n}\;\in\;(0,\infty)\tag{2.6}$$

Then

$$e^{[\Gamma V]}(\mathcal{H}+\{V\})=\mathcal{H}_*+\{V_*\}\qquad\mathcal{H}_*\stackrel{def}{=}\mathcal{H}+\{\mathcal{R}V\}\tag{2.7}$$

and for any $\mathbb{I}\ni\mu<s/3$ we have

$$\|V\|_s\le\epsilon_\mu\;\Longrightarrow\;\|V_*\|_{s-3\mu}\;\le\;\kappa\,{\epsilon_\mu}^2\tag{2.8}$$

for some real positive constant κ.

Remark 1. We recall that $\mathcal{R}$ is a projector if $\mathcal{R}^2=\mathcal{R}$ or equivalently $\mathcal{N}\mathcal{R}=0$

Remark 2. If we can write $\Gamma F=\{\mathcal{G}F\}$, then by the Jacobi identity

$$[\mathcal{H},\Gamma F]=[\{H\},\{\mathcal{G}F\}]=\{\mathcal{H}\mathcal{G}F\}$$

So the fourth hypothesis can be rewritten in a much simpler way,

$$\mathcal{H}\mathcal{G}F=\mathcal{N}F,\quad\forall F\in\mathbb{V}\tag{2.9}$$

Proof. We start by expanding the l.h.s. of equation (2.7),

$$e^{[\Gamma V]}(\mathcal{H}+\{V\})\;=\;\mathcal{H}\;+\;[\Gamma V]\mathcal{H}\;+\;\sum_{l=2}^{\infty}\frac{[\Gamma V]^l}{l!}\mathcal{H}\;+\;\{V\}\;+\;(e^{[\Gamma V]}-1)\{V\}$$

By hypothesis (2.4), $[\Gamma V]\mathcal{H}=-[\mathcal{H}]\Gamma V=-\{\mathcal{N}V\}$ so that

$$\sum_{l=2}^{\infty}\frac{[\Gamma V]^l}{l!}\mathcal{H}\;=\;-\sum_{l=2}^{\infty}\frac{[\Gamma V]^{l-1}}{l!}[\mathcal{H}]\Gamma V\;=$$

$$=\;-\sum_{l=2}^{\infty}\frac{[\Gamma V]^{l-1}}{l!}\{\mathcal{N}V\}\;=\;-\frac{e^{[\Gamma V]}-1-[\Gamma V]}{[\Gamma V]}\{\mathcal{N}V\}$$

[6] The case $V\in\mathbb{B}$ is trivial.

the latter expression being formal. Then

$$e^{[\Gamma V]}(\mathcal{H}+\{V\}) = \mathcal{H} - \{\mathcal{N}V\} - \frac{e^{[\Gamma V]}-1-[\Gamma V]}{[\Gamma V]}\{\mathcal{N}V\} + \{V\} + (e^{[\Gamma V]}-1)\{V\} =$$

$$= \mathcal{H} + \{\mathcal{R}V\} + (e^{[\Gamma V]}-1)\{V\} - \frac{e^{[\Gamma V]}-1-[\Gamma V]}{[\Gamma V]}\{\mathcal{N}V\} \tag{2.10}$$

Now consider the following identity in $\mathbb{V}$,

$$[\Gamma V]\{F\} = \{(\Gamma V)F\} \tag{2.11}$$

which holds because ΓV is a derivation, by definition. In fact, $\forall l \in \mathbb{N}$,

$$[\Gamma V]^l \mathcal{F} = [\Gamma V]^{l-1}[\Gamma V]\mathcal{F} \;=\; [\Gamma V]^{l-1}\{(\Gamma V)F\} \;=\; [\Gamma V]^{l-2}\{(\Gamma V)^2 F\} \;=\; \ldots \;=\; \{(\Gamma V)^l F\}$$

so that

$$(e^{[\Gamma V]}-1)\{V\} \;=\; \sum_{l=1}^{\infty}\frac{[\Gamma V]^l}{l!}\{V\} = \sum_{l=1}^{\infty}\frac{\{(\Gamma V)^l V\}}{l!} \;=\; \{\sum_{l=1}^{\infty}\frac{(\Gamma V)^l V}{l!}\} \;=\; \{(e^{\Gamma V}-1)V\}$$

One can proceed analogously to prove that

$$\frac{e^{[\Gamma V]}-1-[\Gamma V]}{[\Gamma V]}\{\mathcal{N}V\} = \{\frac{e^{\Gamma V}-1-\Gamma V}{\Gamma V}\mathcal{N}V\}$$

Now, if we inject

$$V_* = (e^{\Gamma V}-1)V - \frac{e^{\Gamma V}-1-\Gamma V}{\Gamma V}\mathcal{N}V \tag{2.12}$$

into equation (2.10) we recover the thesis (2.7).

To make the previous formal manipulation meaningful, it's sufficient to show that $e^{\Gamma V}$ is bounded with loss from $\mathbb{V}_s$ to $\mathbb{V}_{s-\mu}$ (it's easier to study convergence on an algebra rather than on the space of its derivations). Then $e^{[\Gamma V]}$ can be computed by the following relation,

$$e^{[\mathcal{A}]}\,\mathcal{B} = \; e^{\mathcal{A}}\,\mathcal{B}\;e^{-\mathcal{A}}, \quad \forall \mathcal{A},\mathcal{B} \in \operatorname{der}\mathbb{V} \tag{2.13}$$

which is readily proven by using a series expansion on both sides,

$$\sum_{N\geq 0}\frac{[\mathcal{A}]^N}{N!}\,\mathcal{B} \;=\; \sum_{n\geq 0, m\geq 0}\frac{\mathcal{A}^n}{n!}\,\mathcal{B}\,\frac{(-\mathcal{A})^m}{m!}$$

On the r.h.s. we use a change of variable, $m \mapsto N-n$,

$$\sum_{N\geq 0}\sum_{n=0}^{N}\frac{\mathcal{A}^n\,\mathcal{B}\,(-\mathcal{A})^{N-n}}{n!(N-n)!} \tag{2.14}$$

Instead by the relation

$$[\mathcal{A}]^N\mathcal{B} = \sum_{k=0}^{N}\binom{N}{k}\mathcal{A}^k\,\mathcal{B}\,(-\mathcal{A})^{N-k}$$

the l.h.s of (2.13) can be rewritten as

$$\sum_{N\geq 0}\sum_{k=0}\frac{\mathcal{A}^k\,\mathcal{B}\,(-\mathcal{A})^{N-k}}{k!\,(N-k)!}$$

By renaming an index, the above is equal to expression (2.14).

Now consider the expression $(\Gamma V)^n F$, as n varies. For $n = 1$ we can apply hypothesis 2.5 with $\delta = 0$ and $d = \mu$ to get

$$\|(\Gamma V)F\|_{s-\mu} \leq \Lambda(\mu, 0)\, \|V\|_s\, \|F\|_s$$

Now let $n \geq 1$ and for any $1 \leq j \leq n$, consider the operator

$$(\Gamma V)^j : \mathbb{V}_{s-(j-1)\mu/n} \to \mathbb{V}_{s-j\mu/n}$$

By applying hypothesis 2.5 with $d = \mu/n$ and $\delta = (j-1)\mu/n$ we get

$$\|(\Gamma V)^j F\|_{s-\mu} \leq \Lambda(\tfrac{\mu}{n}, \tfrac{(j-1)\mu}{n})\, \|V\|_s\, \|(\Gamma V)^{j-1} F\|_{s-(j-1)\mu/n}$$

and, iterating the above n times,

$$\left\|(\Gamma V)^n F\right\|_{s-\mu} \leq \prod_{j=1}^{n} \Lambda(\tfrac{\mu}{n}, \tfrac{(j-1)\mu}{n}) \|V\|_s{}^n \|F\|_s$$

We can finally bound $e^{\Gamma V}$ with loss,

$$\begin{aligned}\left\|e^{\Gamma V} F\right\|_{s-\mu} &\leq \sum_{n=0}^{\infty} \frac{1}{n!} \left\|(\Gamma V)^n F\right\|_{s-\mu} \leq \sum_{n=0}^{\infty} \frac{1}{n!} \prod_{j=1}^{n} \Lambda(\tfrac{\mu}{n}, \tfrac{(j-1)\mu}{n}) \|V\|_s{}^n \|F\|_s \leq \\ &\leq \sum_{n=0}^{\infty} \left(\frac{\|V\|_s}{2\epsilon_\mu}\right)^n \|F\|_s \leq 2\|F\|_s \qquad (2.15)\end{aligned}$$

where we also used equation (2.6) and the hypothesis that $\|V\|_s < \epsilon_\mu$.

To bound the norm of V_* we use a similar technique,

$$\begin{aligned}\|V_*\|_{s-3\mu} &= \Big\| \sum_{l\geq 1} \frac{(\Gamma V)^l}{l!} V - \sum_{l\geq 2} \frac{(\Gamma V)^{l-1}}{l!} \mathcal{N} V \Big\|_{s-3\mu} \leq \\ &\leq \sum_{l\geq 1} \frac{1}{l!} \Big\| (\Gamma V)^l V - \frac{(\Gamma V)^{l-1}}{l+1} \mathcal{N} V \Big\|_{s-3\mu} \leq \qquad (2.16) \\ &\leq \sum_{l\geq 0} \frac{1}{l!} \Big\| (\Gamma V)^l \frac{1}{l+1} \Big((\Gamma V) V + \frac{1}{l+1} (\Gamma V) \mathcal{N} V \Big) \Big\|_{s-3\mu} \leq \qquad (2.17) \\ &\leq \sum_{n=0}^{\infty} \left(\frac{\|V\|_s}{2\epsilon_\mu}\right)^n \Big(\left\|(\Gamma V) V\right\|_{s-2\mu} + \|(\Gamma V) \mathcal{N} V\|_{s-2\mu} \Big) \leq \qquad (2.18) \\ &\leq 2 \Big(\Lambda(2\mu, 0) \|V\|_s^2 + \Lambda(\mu, \mu) \|V\|_s \underbrace{\|\mathcal{N} V\|_{s-\mu}}_{\leq \Xi(\mu)\|V\|_s} \Big) \leq \kappa\, \epsilon_{\mu/3}^2 \qquad (2.19)\end{aligned}$$

where we used hypothesis 2.5, 2.3 and also that, for any positive integer l, $1/(l+1) < 1$, $1/(l+2) < 1$. So we proved equation (2.8). □

2.2.1 A hint of the Iteration Procedure

The second part of the proof of KAM theorem would work in the following way. Suppose that we can apply formula (2.7) to the dynamical system $\mathcal{H} + \{\mathcal{R}V\}$, being V_* the new perturbation. We would need two operators $\mathcal{R}_*$ and Γ_* which satisfy

$$\mathcal{R}_*^2 = \mathcal{R}\,, \quad \operatorname{ran}\mathcal{R} \subseteq \mathbb{B}\,, \quad [\mathcal{H} + \{\mathcal{R}V\}]\Gamma_* F \;=\; \{(\mathbb{I} - \mathcal{R}_*)F\}\,, \forall F \in \mathbb{V}$$

It may happen that $\mathcal{R}_* = \mathcal{R}$, but it is not granted. Then we have:

$$e^{[\Gamma_* V_*]}(\mathcal{H} + \{\mathcal{R}V\} + \mathcal{V}_*) \;=\; \mathcal{H} + \{\mathcal{R}V\} + \{\mathcal{R}_* V_*\} + \mathcal{V}_{**}$$

where now $\mathcal{V}_{**} = O(V_*^2) = O(V^4)$. By iterating infinitely many times one would eventually get to

$$\mathcal{H}_\infty = \mathcal{H} + \sum_{i=1}^{\infty} \{\mathcal{R}\underbrace{*\ldots*}_{i-\text{times}} V \underbrace{*\ldots*}_{i-\text{times}}\} \qquad \text{so that} \qquad \mathcal{H}_\infty \mathbb{B} \subseteq \mathbb{B} \tag{2.20}$$

So that the perturbed system $\mathcal{H} + \{V\}$ has been conjugated to $\mathcal{H}_\infty$.

Equivalently, define a map $\mathcal{T}$ by

$$\mathcal{T} \equiv \cdots \circ e^{\Gamma_{**} V_{**}} \circ e^{\Gamma_* V_*} \circ e^{\Gamma V} \implies \mathcal{T}(\mathcal{H} + \mathcal{V})\mathcal{T}^{-1} = \mathcal{H}_\infty$$

Then

$$\mathcal{T}^{-1}\,\mathcal{H}_\infty\,\mathcal{T}\,\mathcal{T}^{-1}\,\mathbb{B} \;\subseteq\; \mathcal{T}^{-1}\,\mathbb{B}$$

By defining $\tilde{\mathbb{B}} := \mathcal{T}^{-1}\mathbb{B}$ the above equation can be rewritten as

$$(\mathcal{H} + \{V\})\tilde{\mathbb{B}} \;\subseteq\; \tilde{\mathbb{B}}$$

where we read that the algebra $\mathbb{B}$ is structurally stable under perturbation, but it get deformed into $\tilde{\mathbb{B}}$, like the tori of classical mechanics. Still in other words, if $\mathcal{H}$ preserves some subalgebra $\mathbb{B}$, then it is possible to build a new subalgebra $\tilde{\mathbb{B}}$ preserved by the perturbed system $\mathcal{H} + \{V\}$. This is possible if $\mathbb{B}$ satisfies some conditions that ensure the existence of $\mathcal{T}$.

Unfortunately, up to now we have not been able to provide an algorithm to build the operators $\mathcal{R}_*$ and Γ_* in this generical setting. In the following chapter, we will make this iteration mechanism quantitative in two examples.

2.3 A brief historical survey

The literature on KAM theory is immense; here we propose a modest review without any claim of completeness, focusing on those developments we met in developing our work.

In 1954 Kolmogorov proposed a new way to overcome the convergence issues of classical perturbation theory [47]. The main novelty of this approach was to focus on a single torus, and to build a superconvergent algorithm which conjugates the perturbed Hamiltonian to a new one, preserving the prescribed torus. Convergence was assured by the "Diophantine condition" and by non-degeneracy of the Hamiltonian. Also, the Hamiltonian was required to be analytic. The proof was so concise that it was accused to be incomplete [3]. However, a new proof of the theorem, following the original guidelines of Kolmogorov, was provided by Benettin, Galgani, Giogilli and Strelcyn in 1984 [7].

In 1963 Arnol'd [3], [4] gave a new proof of the theorem, with a few technical differencies. Also, he dropped the non-degeneracy condition assumed by Kolmogorov, in place of another one, called the isoenergetic non-degeneracy. In 1962 also Moser provided a result analogous to Kolmogorov's theorem in the framework of area-preserving twist mappings of an annulus ([68], [71]), dropping the analiticity condition for finite differentiability (up to order 333). In the context of Hamiltonian mechanics the requirement for analiticity of the perturbation was dropped by Poschel [78]; later by Salamon dropped also the analiticity of the Hamiltonian [83]. The Diophantine condition was relaxed by Russman [81] and even further by Bruno [94].

One line of research in KAM theory has been the development of an inversion theorem for Frechet spaces; it began with the works of Moser [70], [69], introducing the theorem nowadays called the Nash-Moser theorem. On this topic, we mention also the works of Zehnder [95], [96], Hamilton [45], Herman [8] and more recently Fejoz [37]. For a nice introduction see Raymond [79].

Another line of research was the relation between KAM and Renormalization theory: this was studied by Greene [42] and his student MacKay [57] for twist maps of the annulus, and by Gallavotti (see for instance [38]) and Doveil and Escande [29], [30], [27] for Hamiltonians. For the relation between these lines of research and plasma physics, see also [28].

Among the developments of KAM theory we mention also: a lagrangian formulation of the theorem proposed by Zehnder and Salamon [84]; a version adapted to nearly integrable systems on Poisson manifolds (based on action angle coordinates, but admitting a degeneracy in the symplectic form) by Li and Yi [52]; De La Llave, Gonzales, Jorba and Villanueva who proposed a KAM algorithm for Hamiltonian systems without action-angle variables [20]; De La Llave and Alishah for a KAM theorem for presymplectic systems [2]; Bounemura and Fischler, who recently proposed a new proof where diophantine estimates for small divisors are replaced by continued fractions approximations [9].

Finally we mention the works by Vittot [24], [92], [93]. He proposed an algebraic perturbation theory, which required an operator $\mathcal{G}$ such that $\{H\}^2\mathcal{G} = \{H\}$; the reader can recognize a particular case of theorem 1, in the case of inner derivations (see remark 2), and having chosen $\mathbb{B} = \ker\{H\}$. However, instead of proposing an iteration mechanism à la Kolmogorov, the author proposes a formula to perform the perturbation transform in one step: given a perturbed system $\{H + V\}$, define $W \in \mathbb{V}$ by

$$V \;=\; F(W), \quad F(W) \;=\; e^{-\{\mathcal{G}W\}}\mathcal{R}W \;+\; \frac{1 - e^{-\{\mathcal{G}W\}}}{\{\mathcal{G}W\}}\mathcal{N}W$$

Then,

$$e^{\{\mathcal{G}W\}}(H \;+\; V) \;=\; H \;+\; \mathcal{R}W$$

Unfortunately, the method was proven to work only if $\{\mathcal{G} \quad \}$ is bounded in the sense of Banach.

3

The Throbbing Top

In this chapter we apply the perturbation theorem of chapter 2 to study the dynamics of a Throbbing Top (that is, a Top which moments of inertia are given functions of time).

Physically, the name "Top", or "Rigid Body" denotes a set of points (in the physical space) which mutual distances remain constant in time. In the absence of external forces, the dynamics of a Rigid Body is a rotation around an axis (hence the picture of a Top). The axis is in the direction of the angular momentum vector $\boldsymbol{M}$, while its modulus determines the spinning frequency: to determine the dynamics of the Top we need to solve a dynamical system for its angular momentum $\boldsymbol{M}$.

This dynamical system has been deeply studied in the literature, under different points of view: as an integrable system, as a dynamical system on a Lie group, and as a Poisson system; see for instance [5], [21], [32], [31], [51], [89], [91], [59], [61]. The Top has the advantage of being a low dimensional example, involving less severe calculations that sometimes may be even performed explicitly. And a Throbbing Top can be described as a perturbation of the stationary Top, retaining the same bracket (with all the related features)[1]

The chapter is organized in three main sections:

- the first section 3.1, is an introduction in which we recall the equations of motion for the Top, discuss their conserved quantities, and their solution, their Lie structure, and how these properties change when we switch to a non-autonomous Top;

- in the second section 3.2 we study the dynamics of symmetric Throbbing Top: we introduce a chart on the sphere, in which the bracket is redressed to a canonical one, and the Hamiltonian is diagonalized; we introduce the operators needed to apply our perturbation theory, and a scale of Banach norms. By iteratively applying our "main step formula" we recover a full KAM theorem. We also show the results of some numerical simulations;

- in the third section 3.3 we study the dynamics of a non-symmetric and non-periodic Throbbing Top; in this case we don't introduce any chart and adopt a novel approach to define an invariant algebra and a new scale of Banach norms. Again, our perturbation formula can be applied and also iterated.

[1] this is the reason why we use the name *Top* rather than *Rigid Body*: as the properties of the body are changing in time, it's not "rigid".

As we are dealing with a dynamical system in $\mathbb{R}^3$, for the rest of this chapter we will use the notations introduced in section A.3.

3.1 About the Rigid Body (or Top)

Here we propose a basic review on the rigid body dynamics; for more details see for instance [59], [5], [31], [51]. The properties of the Top (its shape, mass distribution ...) are "summarized" by a matrix, called the "inertia tensor". This matrix is symmetric, so there exist a reference frame in which it is diagonal; the coordinates of this reference frame will be called M_1, M_2, M_3 because they represent the 3 components of the angular momentum vector $\boldsymbol{M}$. The equations of motion of the rigid body in these coordinates read

$$\begin{cases} \dot{M}_1 = \Big(\dfrac{1}{I_3} - \dfrac{1}{I_2}\Big) M_2 M_3 \\ \dot{M}_2 = \Big(\dfrac{1}{I_1} - \dfrac{1}{I_3}\Big) M_3 M_1 \\ \dot{M}_3 = \Big(\dfrac{1}{I_2} - \dfrac{1}{I_1}\Big) M_1 M_2 \end{cases} \tag{3.1}$$

The numbers $\{I_i\}_{i=1}^3$ are the eigenvalues of the inertia tensor, called "moments of inertia"; they are always positive. In general an ordering like $I_1 > I_2 > I_3$ or the opposite $I_1 < I_2 < I_3$ is assumed. The special cases $I_1 = I_2 = I_3$ and $I_1 = I_2$ (or $I_2 = I_3$) are respectively known as the "spherical top", and as the "symmetric top".

Equations (3.1) have two conserved quantities, namely

$$\rho^2 \stackrel{\text{def}}{=} M_1^2 + M_2^2 + M_3^2 \tag{3.2}$$

$$E = \frac{1}{2} \overline{\boldsymbol{M}} \mathsf{L} \boldsymbol{M}, \quad \mathsf{L} = \begin{pmatrix} I_1^{-1} & 0 & 0 \\ 0 & I_2^{-1} & 0 \\ 0 & 0 & I_3^{-1} \end{pmatrix} \tag{3.3}$$

The first quantity is the modulus of $\boldsymbol{M}$: its conservation implies that the motion of the angular momentum is itself a rotation. The second one has the meaning of (kinetic) energy. The two conserved quantities determine two surfaces in $\mathbb{R}^3$, a sphere and an ellipsoid, so that the point of the vector $\boldsymbol{M}$ lies at each instant on their intersection, or, equivalently, the trajectories of the system (3.1) are given by the intersections of the two objects. As a consequence, there exists a set of accessible values for the energy: given ρ and the moments of inertia, the system will have a solution only for

$$\rho^2/(2I_3) \le E \le \rho^2/(2I_1) \tag{3.4}$$

(if $I_1 > I_3$). Outside of this range of values, the ellipsoid and the sphere have no intersection, so the system (3.1) has no solution.

For completeness, we give a description of the trajectories. In the case of a spherical top, the only solution is the identity, $M_i(t) = M_i(0)$. In the case of a symmetric top, the solution is a uniform rotation around the third axis (M_3 is constant in time), if $I_1 = I_2$; or around the first axis (M_1 is constant in time), if $I_2 = I_3$. Now let $I_1 > I_2 > I_3$, (but the description would be analogous with the opposite ordering). The phase space is partioned by the dynamics into four different open and disjoint subsets. Following the

notation of [89], they are

$$
\begin{aligned}
U_1^- &= \left\{(M_1, M_2, M_3) \in \mathbb{R}^3 \text{ s.t. } I_1 < 2\,E\,\rho^{-2} < I_2,\ M_1 < 0\right\} \\
U_1^+ &= \left\{(M_1, M_2, M_3) \in \mathbb{R}^3 \text{ s.t. } I_1 < 2\,E\,\rho^{-2} < I_2,\ M_1 > 0\right\} \\
U_3^- &= \left\{(M_1, M_2, M_3) \in \mathbb{R}^3 \text{ s.t. } I_2 < 2\,E\,\rho^{-2} < I_3,\ M_3 < 0\right\} \\
U_3^+ &= \left\{(M_1, M_2, M_3) \in \mathbb{R}^3 \text{ s.t. } I_2 < 2\,E\,\rho^{-2} < I_3,\ M_3 > 0\right\}
\end{aligned}
\tag{3.5}
$$

To each set belongs one of the four elliptic equilibrium points, $\rho^2 = E/(2I_{1,3})$. Instead for $\rho^2 = E/(2I_2)$ we get the four separatrices joining the two hyperbolic equilibrium points of coordinates $M_1 = M_3 = 0$. Inside each of the charts the trajectory can be explicitly expressed in terms of Jacobi elliptic functions. For instance, in U_3^+,

$$
\begin{cases}
M_1(t) = \operatorname{sign}(M_1^0)\sqrt{\dfrac{I_1}{\rho}}\sqrt{\dfrac{\rho^2 - 2I_3E}{I_1 - I_3}}\ \operatorname{cn}\big(\omega(t-t_0)|m\big) \\
M_2(t) = -\operatorname{sign}(M_1^0)\sqrt{\dfrac{I_2}{\rho}}\sqrt{\dfrac{\rho^2 - 2I_3E}{I_2 - I_3}}\ \operatorname{sn}\big(\omega(t-t_0)|m\big) \\
M_3(t) = \sqrt{\dfrac{I_3}{\rho}}\sqrt{\dfrac{2I_1E - \rho^2}{I_1 - I_3}}\ \operatorname{dn}\big(\omega(t-t_0|m)\big)
\end{cases}
$$

$$
\omega = -\sqrt{\frac{(\rho^2 - 2I_1E)(I_3 - I_2)}{I_1I_2I_3}} \tag{3.6}
$$

$$
m = \frac{(\rho^2 - 2I_3E)(I_1 - I_2)}{(\rho^2 - 2I_1E)(I_3 - I_2)}
$$

The trajectories on U_3^- will have the same analitical form, but with opposite sign of M_3; those on $U_1^\pm$ are found by exchanging the roles of M_1 and M_3.

In figure 3.1 we also plot a few trajetories for a rigid body with moments $I_i = i$.

3.1.1 Lie-Poisson structure for the Rigid Body

The space $\mathbb{R}^3$ is a Lie algebra with the bracket $[\![\ \ ]\!]$ (the vector product). As a consequence of the Lie-Poisson theorem (which is discussed in section A.1.3, see also [59]), the set

$$
\mathbb{V}_{\text{Top}} \stackrel{\text{def}}{=} \mathcal{C}^\infty(\mathbb{R}^{3*} \to \mathbb{R}) \tag{3.7}
$$

is a Poisson algebra with bracket

$$
\{F, G\}\big(\overline{\boldsymbol{M}}\big) = \overline{\boldsymbol{M}}\big[\!\big[\partial_{\overline{\boldsymbol{M}}}F\big]\!\big]\partial_{\overline{\boldsymbol{M}}}G, \quad \forall F, G \in \mathbb{V}_{\text{Top}} \tag{3.8}
$$

The operator $\partial_{\overline{\boldsymbol{M}}}$ on $\mathbb{V}$ is defined by

$$
\overline{\boldsymbol{N}}\partial_{\overline{\boldsymbol{M}}}f = \lim_{\eta\to 0}\frac{f(\overline{\boldsymbol{M}} + \eta\overline{\boldsymbol{N}}) - f(\overline{\boldsymbol{M}})}{\eta}
$$

and it takes elements of $\mathbb{V}$ into elements of $\mathbb{R}^3$. This is evident from the definition: when we act on $\partial_{\overline{\boldsymbol{M}}}f$ with an element $\overline{\boldsymbol{N}} \in \mathbb{R}^{3*}$, we get a scalar.

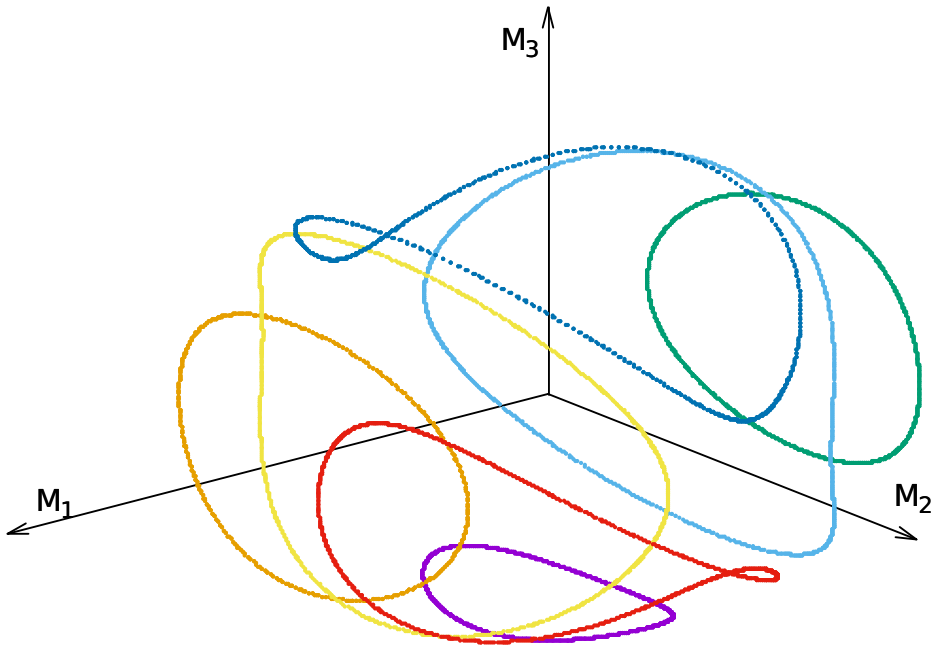


Figure 3.1: *A few trajectories of the system* (3.1) *with parameters* $I_1 = 1$, $I_2 = 2$, $I_3 = 3$. *The trajectories were generated by a code emplying a Runge-Kutta 4th order integration scheme and step* $h = 0.001$. *The initial data were randomly generated with the unique constraint of having all the same value of* $\rho = 2$. *Conservation of the energy and of* ρ *was achieved up to numerical precision.*

If we consider as Hamiltonian the function E of equation (4.41) then by $\dot{M} = \{E\}M$ we recover the system (3.1).

The function ρ defined by equation (3.2) has the property $\{\rho\}F = 0\,,\ \forall F \in \mathbb{V}_{\text{Top}}$; we call it a *Casimir element* [59]. A Casimir element is constant under the flow determined by any Hamiltonian; in fact, it is a property of the algebra, not of the flow. Each Casimir element determines one dimension of a Poisson algebra that is frozen by any dynamics, for any Hamiltonian. Thus the dynamics of a Top takes place in a two dimensional space: a sphere of radius ρ. It is possible to show that, given a dynamical system on a Poisson algebra, by quotienting away the Casimir elements, we get a canonical system. And a two dimensional, autonomous canonical system is integrable[2], and so is the case for the Top.

[2]In the context of sympletic mechanics, a dynamical system of dimension $2n$ is called integrable if it has n quantities in involution (i.e. having zero bracket) among themselves and with the Hamiltonian. As an obvious consequence, a canonical Hamiltonian system is always integrable for $n = 1$, which is the case of the Top.

3.1.2 The Throbbing Top

We switch to a non-autonomous Top following the method described in the appendix A.1.2. So, from the algebra (3.7) we pass to the algebra

$$\mathbb{V}_{TT} \stackrel{\text{def}}{=} \mathcal{C}^\infty\Big(\begin{pmatrix}\mathbb{T}\\ \mathbb{R}\end{pmatrix} \to \mathbb{V}_{\text{Top}}\Big) \ni f = f\big(\overline{\boldsymbol{M}}, t\big) \tag{3.9}$$

again with the bracket (3.8). We will choose $\mathbb{T}$ (a torus) for a periodic system (as in section 3.2), or $\mathbb{R}$, for a generical time-dependence (as in section 3.3). In particular, as the time variable doesn't enter in the bracket, ρ is still a Casimir. This means that the energy, even if it is fluctuating, has to respect the bound (3.4). And the bound itself may start fluctuating, if the fluctuation is added to I_1 or I_3.

The unperturbed Hamiltonian is still given by (3.3). We are interested in perturbations of type

$$V = \frac{1}{2}\,\overline{\boldsymbol{M}}\mathsf{A}(t)\boldsymbol{M} \tag{3.10}$$

where $\mathsf{A}(t)$ is a 3×3 diagonal matrix with time dependent coefficients. Physically, this will represent a top for which the moments of inertia are changing in time.

The new dynamical system is

$$\dot{F} \;=\; \mathcal{H}F \;+\; \{V\}F\,, \quad \mathcal{H} \;=\; \{E\} \;+\; \partial_t \tag{3.11}$$

In particular, for $F = M_i$, $i = 1, 2, 3$ and V given by (3.10) with

$$\mathsf{A}(t) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \epsilon\cos(\nu t) & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.12}$$

we recover the Euler-Poinsot equations with

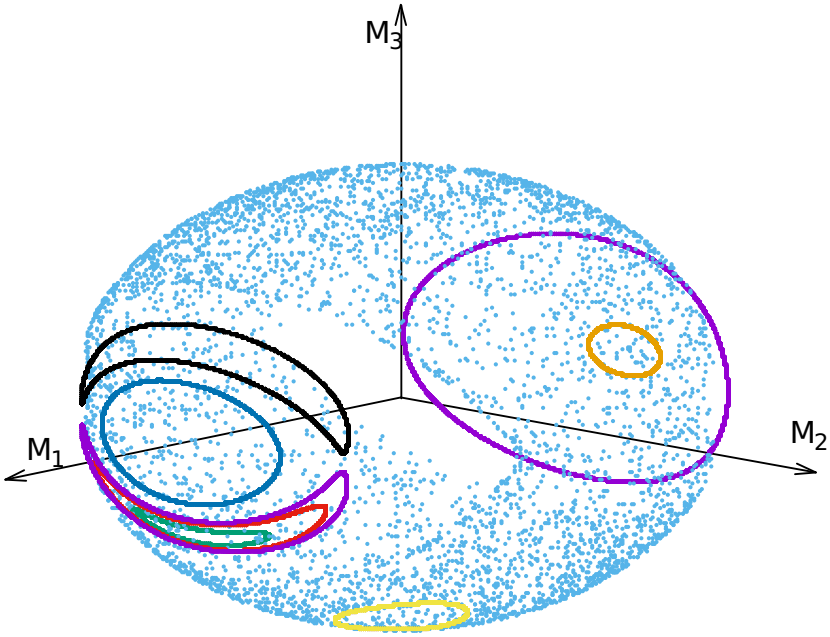
$$I_2 \;\mapsto\; I_2/(1 + I_2^{\emptyset}\epsilon\cos(\nu t))$$

In figure 3.1 we plot the trajectories of this dynamical system, and also in two other cases, when a similar perturbation is applied to I_1 and to I_3 (see also the captions of the subplots). We observe the typical features observed in dynamical systems with coexistence of order and chaos. The separatrices disappear, and are replaced by orbits spanning a two-dimensional area. Around the elliptic equilibrium points, some of the original trajectories are only deformed, some others are lost and leave place to a set of new equilibrium points; some of the new equilibrium points are elliptic, and new closed orbits appear around them.

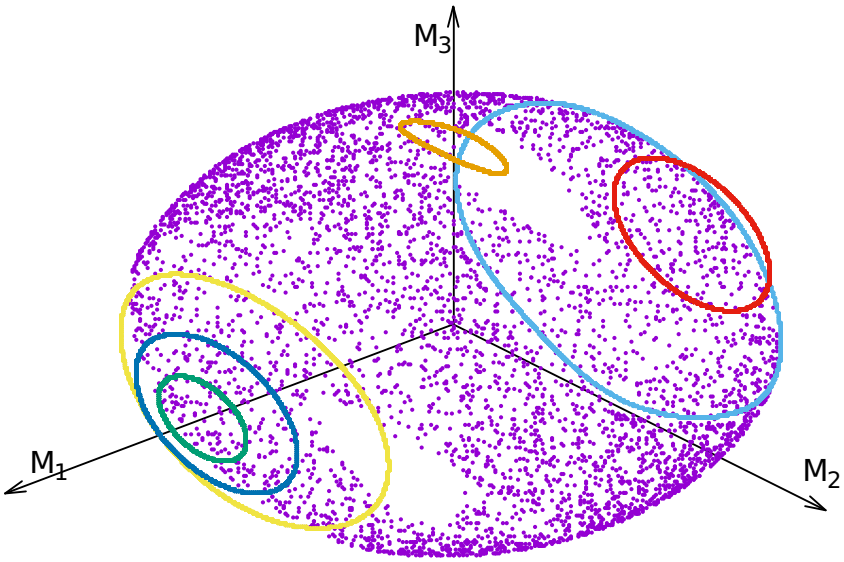
3.2 The symmetric case

By definition a Top is symmetric if two moments of inertia are equal; here we fix $I_1 = I_2 \equiv I_\perp$. In this case we consider the change of coordinates (see [21], [44]),

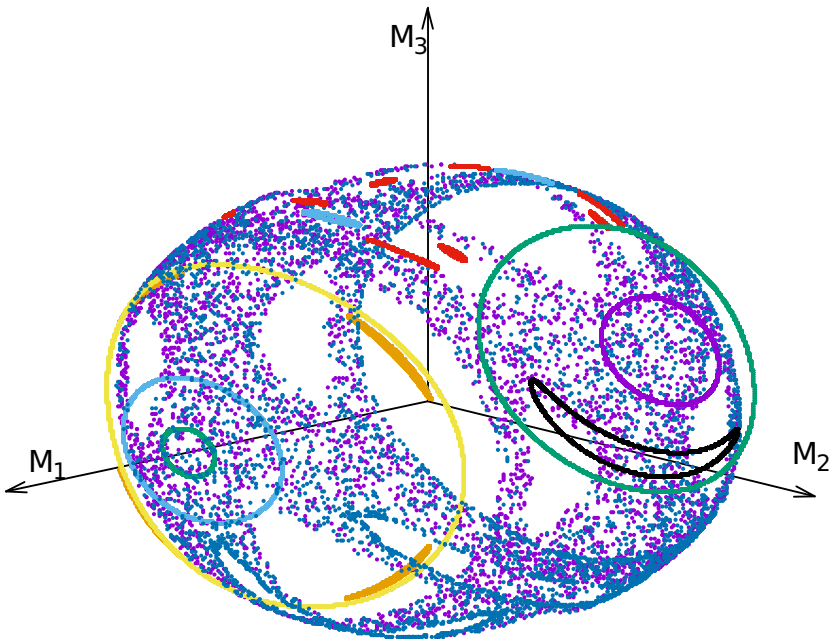
$$\{M_1, M_2, M_3\} \mapsto \{\rho, X, \theta\} : \begin{cases} M_1 = \rho\sqrt{1-X^2}\cos(\theta) \\ M_2 = \rho\sqrt{1-X^2}\sin(\theta) \\ M_3 = \rho X \end{cases}$$



(a) $I_1 = 1$, $I_2 = 2/(1 + 0.2\epsilon\cos(t))$, $I_3 = 3$



(b) $I_1 = 1$, $I_2 = 2$, $I_3 = 3/(1 + 0.3\epsilon\cos(t))$



(c) $I_1 = 1/(1 + 0.1\cos(t))1$, $I_2 = 2$, $I_3 = 3$

Figure 3.1: *A few trajectories of the system* (3.1) *with time-dependent moments, as denoted in the subcaptions. The trajectories were generated by a code employing a Runge-Kutta 4th order integration scheme and step* $h = 2$*. The initial data were randomly generated with the unique constraint of having all the same value of* $\rho = 2$*; conservation of the latter along the simulation was also checked.*

with $X \in (-1,1)$ and $\theta \in [0, 2\pi)$. These coordinates don't cover the whole sphere, as the north and south poles are excluded. However, in the stationary case the poles are elliptic equilibria, so they are not very interesting for the dynamics. In the non-autonomous case the energy is still subject to the bound (3.4). If the bound is strengthened to strict inequalities then the dynamics will never reach the poles.

Remark 3. The functions $f(X, \theta, t)$ are only a subalgebra of the whole algebra $\mathbb{V}_{TT}$ (defined in (3.9)); after making a further restriction to analytic functions, we are left with a new subalgebra $\mathbb{V}_{\text{symm}} \subseteq \mathbb{V}_{TT}$. The bracket (3.8) restricted to $\mathbb{V}_{\text{symm}}$ becomes[3]

$$\{F\}G = \frac{1}{\rho}\big(\partial_X F\, \partial_\theta G \,-\, \partial_\theta F\, \partial_X G\big) \tag{3.13}$$

The new bracket contains no derivatives in ρ, coherently with the definition of a Casimir. The Hamiltonian (3.3) becomes

$$E_{\text{symm}} \,=\, \frac{\rho^2}{2}\left(\frac{1-X^2}{I_\perp} + \frac{X^2}{I_3}\right) \,\equiv\, \frac{\rho^2}{2}\Delta\, X^2 \,+\, \frac{\rho^2}{2I_\perp} \tag{3.14}$$

where we have set $\Delta = \frac{1}{I_3} - \frac{1}{I_\perp}$.

We see that the coordinates X, θ both redress the bracket and diagonalize the Hamiltonian, like action-angle coodinates in classical mechanics. Another similarity is that any Hamiltonian will be at most quadratic in X, as seen by equations (4.41) and (3.10). However, at variance with classical mechanics, the phase space $\mathbb{S}_2 \times \mathbb{T}$ is not the cotangent bundle of anything.

3.2.1 A scale of Banach norms for $\mathbb{V}_{\text{symm}}$

Functions in $\mathbb{V}_{\text{symm}}$ admit the Fourier representation

$$F(X, \theta, t) \,=\, \sum_{l,m\in\mathbb{Z}} F_{l,m}(X)\, e^{ilt+im\theta} \tag{3.15}$$

In analogy with classical mechanics, we build a complex extension of the domain of X. The procedure for the complexification of a Lie algebra is described in section A.2, however, we don't need it as it's evident that the complexification of $\mathbb{R}$ is $\mathbb{C}$. For any $X \in (-1,1)$, we consider a ball $\mathbb{B}_r(x) \subseteq \mathbb{C}$ of radius[4] $r \in \mathbb{R}_+$ centered at X. The radius r has to be sufficiently small so that $|X \pm r| < 1$. Then we define the set

$$\mathbb{A}_r \stackrel{\text{def}}{=} \bigcup_{X\in(-1,1)} \mathbb{B}_r(X)$$

The algebra $\mathbb{V}_{\text{symm}}$ is a subalgebra of

$$\mathbb{V}_r \stackrel{\text{def}}{=} \mathcal{C}^\infty(\mathbb{A}_r \otimes \mathbb{T}^2)$$

for any r. Each space $\mathbb{V}_r$ is endowed with the Banach norm

$$\big\|f\big\|_r \stackrel{\text{def}}{=} \sum_{l,m\in\mathbb{Z}} \big|f_{l,m}(X)\big|_r\, e^{r(|l|+|m|)}\,, \quad \big|f_{l,m}\big|_r \stackrel{\text{def}}{=} \sup_{X\in\mathbb{A}_r} |f(X)| \tag{3.16}$$

So we have a scale of Banach norms $\{\big\| \quad \big\|_r\}_{r\in\mathbb{R}_+}$ and a scale of Banach spaces $\{\mathbb{V}_r\}_{r\in\mathbb{R}_+}$. The properties of this norm are collected in the following proposition.

[3] by abuse of notation, we use the same symbol $\{\quad\}$ as before

[4] we denote by $\mathbb{R}_+$ the set of positive reals

Proposition 2. *Let $r, \delta, d \in \mathbb{R}_+$ with $d + \delta < r$. Let also $W \in \mathbb{V}_r$, $Z \in \mathbb{V}_{r-\delta}$. Then*

$$\|\partial_X W\|_{r-d} \leq \frac{1}{d}\|W\|_r \tag{3.17}$$

$$\|\partial_\theta W\|_{r-d} \leq \frac{1}{ed}\|W\|_r \tag{3.18}$$

$$\|\{W\}Z\|_{r-d-\delta} \leq \frac{2}{\rho e d(d+\delta)}\|W\|_r\|Z\|_{r-\delta} \tag{3.19}$$

Proof. As $\mathbb{A}_r \subseteq \mathbb{C}$, the Cauchy inequality (see for instance [41], or also appendix A.1.1)

$$|\partial_X f|_{r-\delta} \leq \frac{1}{\delta}\,|f|_r$$

holds, and it implies that

$$\|\partial_X W\|_{r-d} = \sum_{l,m} e^{(r-d)(|l|+|m|)}\big|\partial_X W_{l,m}\big|_{r-d} \leq \sum_{l,m} e^{(r-d)(|l|+|m|)}\frac{|W_{l,m}|_r}{d} \leq \frac{\|W\|_r}{d}$$

To prove formula (3.18) we need the following identity,

$$x, \alpha, \beta \in \mathbb{R}_+ \implies x^\alpha e^{-\beta x} \leq \left(\frac{\alpha}{e\beta}\right)^\alpha \tag{3.20}$$

Then

$$\|\partial_\theta W\|_{r-d} = \sum_{l,m} |i\,m\,W_{l,m}|_{r-d}\,e^{(r-d)(|l|+|m|)} =$$

$$= \sum_{l,m} |m|\,|W_{l,m}|_{r-d}\,e^{(r-d)(|l|+|m|)} \leq \frac{1}{ed}\sum_{l,m}|W_{l,m}|_{r-d}\,e^{r(|l|+|m|)} \leq \frac{1}{ed}\|W\|_r$$

Finally, to prove formula (3.19) we need the two previous ones, and eq. (3.20) as well.

$$\|\{W\}Z\|_{r-\delta-d} = \frac{1}{\rho}\big\|\partial_X W\,\partial_\theta Z - \partial_\theta W\partial_X Z\big\|_{r-\delta-d} =$$

$$= \Big\| \sum_{L,M\in\mathbb{Z};l,m\in\mathbb{Z}_0} e^{iLt+iM\theta} i\big(\partial_X W_{L-l,M-m}\times$$

$$\times\, m\,Z_{l,m} - (M-m)\,W_{L-l,M-m}\,\partial_X f_{l,m}\big)\Big\|_{r-\delta-d} =$$

$$= \sum_{L,M\in\mathbb{Z}} e^{(r-\delta-d)(|L|+|M|)}\Big|\sum_{l,m\in\mathbb{Z}_0}\big(\partial_X W_{L-l,M-m}\times$$

$$\times\, m\,Z_{l,m} - (M-m)\,W_{L-l,M-m}\,\partial_X Z_{l,m}\big)\Big|_{r-\delta-d}$$

$$\leq \frac{1}{\rho}\sum_{L,M\in\mathbb{Z}} e^{(r-d-\delta)(|L|+|M|)}\sum_{l,m\in\mathbb{Z}}\Big(|m|\,\big|\partial_X W_{L-l,M-m}\big|_{r-\delta-d}\,|Z_{l,m}|_{r-\delta-d} +$$

$$+\,|M-m|\,|W_{L-l,M-m}|_{r-\delta-d}\,\big|\partial_X f\big|_{r-\delta-d}\Big)$$

$$\leq \frac{1}{\rho} \sum_{L,M,l,m\in\mathbb{Z}} \Bigg(e^{(r-d-\delta)(|L-l|+|M-m|)}\, e^{(r-\delta)(|L-l|+|M-m|)}\, |Z_{l,m}|_{r-\delta-d}\, \underbrace{e^{-d(|l|+|m|)}|m|}_{\leq 1/(e(d+\delta))} \times$$

$$\times \frac{|W_{L-l,M-m}|_r}{d+\delta} \;+\; e^{r(|L-l|+|M-m|)}\, |W_{L-l,M-m}|_{r-\delta-d}\, e^{(r-\delta-d)(|l|+|m|)} \frac{|Z|_{r-\delta}}{d} \times$$

$$\times \underbrace{e^{-(\delta+d)(|L-l|+|M-m|)}(|M-m|+|L-l|)}_{\leq 1/(e(\delta+d))} \Bigg) \leq \frac{2}{\rho e d(d+\delta)} \|W\|_r \|Z\|_{r-\delta}$$

□

3.2.2 Application of Proposition 1 to the Symmetric Throbbing Top

Before proceeding we make a further change of coordinates

$$X \;=\; x_0 + x \;\;\Longrightarrow\;\; \partial_X \mapsto \partial_x$$

where $x_0 \in (-1,1)$ is fixed and x is sufficiently small so that $x_0 + x \in (-1,1)$. This change of variables is simply a translation and it doesn't affect the algebraic and metric properties that we introduced up to now. Let us also define

$$Q \stackrel{\text{def}}{=} \partial^2_{xx} H|_{x=0} \tag{3.21}$$

so that, for instance, $\{E_{\text{symm}}\} \;=\; \rho x_0 \Delta \partial_\theta + \{\frac{1}{2} Q x^2\}$

Now we will show that all the hypothesis of Proposition 1 are satisfied for the symmetric Throbbing Top, that is, by system (3.11) on the algebra $\mathbb{V}_{\text{symm}}$ with $E = E_{\text{symm}}$.

1. First we look for a subalgebra $\mathbb{B}$ of $\mathbb{V}_{\text{symm}}$, invariant by $\mathcal{H}$. Lead again by analogy with classical mechanics, we choose

$$\mathbb{B} = \big\{ F(\rho,x,\theta,t) \in \mathbb{V} \quad \text{s.t.} \quad F(\rho,0,\theta,t) = 0,\, \partial_x F(\rho,0,\theta,t) = 0 \big\} \tag{3.22}$$

By definition we have $F \in \mathbb{B} \iff \mathcal{P}_{\geq 2} F_2 = F_2$, but neither $\{H\}$ nor ∂_t can decrease the degree in x of a polynomial. So $\mathcal{H}\mathbb{B} \subseteq \mathbb{B}$.

2. As a second step we build the projector $\mathcal{R}$ (and thus $\mathcal{N} = \mathbb{I} - \mathcal{R}$). We choose

$$\mathcal{R} \stackrel{\text{def}}{=} \mathcal{R}_s - \mathcal{K}, \quad \mathcal{N} \stackrel{\text{def}}{=} \mathcal{N}_s + \mathcal{K} \tag{3.23}$$

where $\mathcal{R}_s, \mathcal{N}_s$ and $\mathcal{K}$ are defined in table 3.1. It's evident that $\mathcal{R}_s$ takes values in $\mathbb{B}$, and then $\mathcal{R} = \mathcal{R}_s - \mathcal{K} \equiv \mathcal{R}_s(1-\mathcal{K})$. In point (**4**) we show that $\mathcal{N}\mathcal{R} = 0$ so we can conclude that they are both projectors.

3. The third step is to build the operator Γ. Still making reference to table 3.1, consider the following operator:

$$\mathcal{G} = \mathcal{G}_s + \rho\theta\mathcal{A} - x\mathcal{G}_s Q(\mathcal{A} + \partial_\theta \mathcal{G}_s \mathcal{P}_0) \tag{3.24}$$

It acts on elements of $\mathbb{V}$, but it doesn't take values in $\mathbb{V}$, because the function θ doesn't[5] belong to $\mathbb{V}$. Nevertheless we can formally compute

$$\Gamma f = \{\mathcal{G} f\} = \{\mathcal{G}_s f\} - \rho^{-1}\mathcal{A} f \partial_x - \{x\mathcal{G}_s Q(\mathcal{A} + \partial_\theta \mathcal{G}_s \mathcal{P}_0) f\} \tag{3.25}$$

[5] θ is not periodic in θ!

$$\oint\colon f(x,\theta,t) \mapsto \int_0^{2\pi} d\theta \int_0^{2\pi} dt f(x,\theta,t) \equiv f_{0,0}(x)$$

$$\chi \stackrel{\text{def}}{=} 1 - \oint\colon f \mapsto \sum_{l,m\in\mathbb{Z}_0} f_{l,m}(x) e^{im\theta+ilt}, \mathbb{Z}_0 \stackrel{\text{def}}{=} \mathbb{Z}\backslash\{0\}$$

$$\mathcal{P}_k\colon \sum_{n\geq 0} a_n x^n \mapsto a_k x^k \;\; k\in\mathbb{N},\; a_n \in \mathbb{R}\; \forall n \in \mathbb{N}$$

$$\mathcal{P}_{\geq k}\colon \sum_{n\geq 0} a_n x^n \mapsto \sum_{n\geq k} a_n x^n$$

$$\mathcal{R}_s \stackrel{\text{def}}{=} \oint \mathcal{P}_0 + \mathcal{P}_{\geq 2}$$

$$\mathcal{N}_s \stackrel{\text{def}}{=} \chi\mathcal{P}_0 + \mathcal{P}_1 \equiv 1 - \mathcal{R}_s$$

$$\mathcal{G}_s\colon f \mapsto \sum_{l,m\in\mathbb{Z}_0} \frac{-i\mathcal{P}_{\leq 1} f_{l,m}(x)}{\rho x_0 \Delta m + l} e^{im\theta+ilt}$$

$$\mathcal{A} \stackrel{\text{def}}{=} \left(\oint Q\right)^{-1} \oint \mathcal{P}_0(\partial_x - Q\partial_\theta \mathcal{G}_s)$$

$$\mathcal{K} \stackrel{\text{def}}{=} \rho^2 x_0 \Delta\mathcal{A} + \{\tfrac{1}{2}Qx^2\}\, x\mathcal{G}_s\left(\mathcal{P}_0\partial_x - Q\mathcal{A} - Q\partial_\theta\mathcal{G}_s\mathcal{P}_0\right)$$

Table 3.1: *Here we group some of the operators defined on $\mathbb{V}$ and needed for the KAM algorithm. We are using the Fourier representation* (3.15)*, the* 1*s are to be intended as identity operators, H is the Hamiltonian, and Q has been defined in equation* (3.21)*.*

where the "illegal" θ got cancelled by the bracket, so Γ goes from $\mathbb{V}$ to $\operatorname{der}\mathbb{V}$. This construction is usefull because we could apply Remark 2 and prove hypothesis 2.4 in the form (2.9). But before doing so, there is another complication to overcome, because $\mathcal{H} \neq \{H\}$; we have to verify also that

$$[\partial_t]\{P\} = \{\partial_t P\}, \quad \forall P \in \mathbb{V} \tag{3.26}$$

i.e.

$$\partial_t\{P\}R - \{P\}\partial_t R = \{\partial_t P\}R, \quad \forall R \in \mathbb{V}$$

For the l.h.s. above we have

$$\begin{aligned}
&\partial_t\{P\}R - \{P\}\partial_t R = \\
&\quad = \partial_t\{\sum_{l'\in\mathbb{Z}} P_{l'}(x,\theta)e^{il't}\}\sum_{l\in\mathbb{Z}} R_l(x,\theta)e^{ilt} - \{\sum_{l'\in\mathbb{Z}} P_{l'}(x,\theta)e^{il't}\}\partial_t\sum_{l\in F} R_l(x,\theta)e^{ilt} = \\
&\quad = \sum_{L,l\in\mathbb{Z}} e^{iLt} iL\{P_{L-l}(x,\theta)\}R_l(x,\theta) - \sum_{L,l\in\mathbb{Z}} e^{iLt}\{P_{L-l}(x,\theta)\}ilR_l(x,\theta) = \\
&\qquad\qquad = \sum_{L,l\in\mathbb{Z}} e^{iLt} i(L-l)\{P_{L-l}(x,\theta)\}R_l(x,\theta)
\end{aligned}$$

while for the r.h.s

$$\{\partial_t\sum_{l'\in\mathbb{Z}} P_{l'}(x,\theta)e^{il't}\}\sum_{l\in\mathbb{Z}} R_l(x,\theta)e^{ilt} = \sum_{L\in\mathbb{Z}} e^{iLt} i(L-l)\{P_{L-l}(x,\theta)\}R_l(x,\theta)$$

so formula (3.26) is established.

Now it's really sufficient to check (2.9), which we rewrite explicitly:

$$(\rho x_0\Delta\partial_\theta + \partial_t + \{\tfrac{1}{2}Qx^2\})\,(\mathcal{G}_s + \rho\theta\mathcal{A} - x\mathcal{G}_s Q(\mathcal{A} + \partial_\theta\mathcal{G}_s\mathcal{P}_0))f = \mathcal{N}f$$

We use

$$\{\tfrac{1}{2}Qx^2\} = xQ\partial_\theta - \tfrac{1}{2}x^2(\partial_\theta Q)\partial_x$$

and we get

$$\begin{aligned}
&\chi\mathcal{P}_{\leq 1}f + \{\tfrac{1}{2}Qx^2\}\mathcal{G}_s f - x\chi\mathcal{P}_{\leq 1}Q(\mathcal{A} + \partial_\theta\mathcal{G}_s\mathcal{P}_0)f + \underset{\bullet}{\underline{xQ\mathcal{A}f}} + \underset{\sim}{\underline{\rho x_0\Delta\mathcal{A}f}} \\
&+\underset{\square}{\underline{\{\tfrac{1}{2}Qx^2\}x\mathcal{G}_s Q(\mathcal{A} + \partial_\theta\mathcal{G}_s\mathcal{P}_0)f}} = (\chi\mathcal{P}_0 + \mathcal{P}_1)f + \underset{\sim}{\underline{\rho x_0\Delta\mathcal{A}f}} + \\
&-\underset{\square}{\underline{\{\tfrac{1}{2}Qx^2\}x\mathcal{G}_s Q\mathcal{A}f}} + \underset{\bullet}{\underline{xQ\mathcal{A}f}} + \{\tfrac{1}{2}Qx^2\}x\mathcal{G}_s\mathcal{P}_0\partial_x f - \underset{\square}{\underline{\{\tfrac{1}{2}Qx^2\}x\mathcal{G}_s\mathcal{P}_0 Q\partial_\theta\mathcal{G}_s f}}
\end{aligned}$$

All the terms underlined in the same way cancel among themselves, and we are left with

$$\{\tfrac{1}{2}Qx^2\}\mathcal{G}_s f + xQ\mathcal{A}f - x(\chi Q)\mathcal{A}f - x\chi Q\partial_\theta\mathcal{G}_s\mathcal{P}_0 f \;=\; \oint \mathcal{P}_1 f + \{\tfrac{1}{2}Qx^2\}x\mathcal{G}_s\mathcal{P}_0\partial_x f$$

Now we observe that $x\mathcal{P}_0\partial_x f = \mathcal{P}_1 f$ so that there is a partial cancellation among the first and the latter term in the above equation, and there remains

$$\{\tfrac{1}{2}Qx^2\}\mathcal{G}_s\mathcal{P}_0 f \;+\; x\big(\oint Q\big)\mathcal{A}f \;-\; x\chi Q\partial_\theta\mathcal{G}_s\mathcal{P}_0 f = \oint \mathcal{P}_1 f$$

Then we insert the explicit expressions of $\{\frac{1}{2}Qx^2\}$ and that of $\mathcal{A}$ as it can be found in table 3.1,

$$\underset{\diamond}{\underline{xQ\partial_\theta\mathcal{G}_s\mathcal{P}_0 f}} \;+\; \underset{\triangle}{\underline{x\oint \mathcal{P}_0\partial_x f}} \;-\; \underset{\diamond}{\underline{x\oint Q\partial_\theta\mathcal{G}_s\mathcal{P}_0 f}} \;-\; \underset{\diamond}{\underline{x\chi Q\partial_\theta\mathcal{G}_s\mathcal{P}_0 f}} \;=\; \underset{\triangle}{\underline{\oint \mathcal{P}_1 f}}$$

Again we underlined in the same way all the terms that cancels out. We conclude that equation (2.9) is satisfied.

4. Here we show that $\mathcal{GR} = 0$, so that $\mathcal{HGR} = \mathcal{NR} = 0$. So,

$$\begin{aligned}
\mathcal{GR} \;=\; &[\mathcal{G}_s \;+\; \theta\mathcal{A} \;-\; x\mathcal{G}_s\, Q\,(\mathcal{A} \;+\; \partial_\theta\mathcal{G}_s\mathcal{P}_0)\,]\times \\
&\times\;[\mathcal{R}_s \;-\; \rho\Delta x_0\mathcal{A} \;-\; \{\tfrac{1}{2}Qx^2\}\,x\,\mathcal{G}_s\,(\mathcal{P}_0\partial_x \;-\; Q\mathcal{A} \;-\; Q\partial_\theta\mathcal{G}_s\mathcal{P}_0)]
\end{aligned}$$

Next we observe the following equalities

$$\begin{aligned}
\mathcal{G}_s\mathcal{R}_s &\propto \chi\mathcal{P}_{\leq 1}(\oint \mathcal{P}_0 + \mathcal{P}_{\geq 2})) = 0 \\
\mathcal{AR}_s f &= \Big(\oint Q\Big)^{-1}\Big(\oint \mathcal{P}_0\partial_x\mathcal{P}_{\geq 2} - \oint Q\partial_\theta\mathcal{G}_s\mathcal{R}_s\Big) = 0 \\
\mathcal{AA} &\propto \mathcal{A}\oint \mathcal{P}_0 = 0 \\
\mathcal{G}_s\mathcal{A} &\propto \mathcal{G}_s\oint \mathcal{P}_0 = 0
\end{aligned}$$

so we are left with

$$\mathcal{GR} = \big(-\mathcal{G}_s - \theta\mathcal{A} + x\mathcal{G}_s Q(\mathcal{A} + \partial_\theta\mathcal{G}_s\mathcal{P}_0)\big)\underbrace{\{\tfrac{1}{2}Qx^2\}x\mathcal{G}_s(\mathcal{P}_\prime\partial_x - Q\mathcal{A} - Q\partial_\theta\mathcal{G}_s\mathcal{P}_0)}_{\in\mathrm{ran}(\mathcal{R}_s)} = 0$$

5. In next proposition 3 we prove the inequalities (3.28) and (3.29), which are respectively of type 2.5 and 2.3, with

$$\Lambda(d,\delta) \;=\; \frac{C(\tau,\gamma,r,\delta,d,q)}{q^3\, d\,(d+\delta)^{2\tau+2}}\,, \qquad \Xi(\delta) \;=\; \frac{\tilde{C}}{q^3\,\delta^{2\tau+3}}$$

We see that some extra hypothesis on Q and on the product $\rho\Delta x_0$ are required. In particular condition (1) of Proposition 3 is usually called the "diophantine condition"; in section 3.2.5 we show that, once ρ and Δ have been fixed, the set of values of x_0 which satisfy this hypothesis is large in a measure theoretic sense. We can also compute, according to equation (2.6),

$$\epsilon_\mu \;=\; \frac{1}{2}\sup_{n\in\mathbb{N}}\left(\frac{1}{n!}\prod_{j=1}^{n}\Lambda\Big(\tfrac{\mu}{n},\tfrac{(j-1)\mu}{n}\Big)\right)^{-1/n} \;=\; \frac{1}{2}\sup_{n\in\mathbb{N}}\left(\frac{1}{n!}\prod_{j=1}^{n}\frac{C}{q^3\,\frac{\mu}{n}\left(\frac{j\mu}{n}\right)^{2\tau+2}}\right)^{-1/n} \;=$$

$$=\; \frac{q^3\mu^{2\tau+3}}{2\,C}\sup_{n\in\mathbb{N}}\left(\frac{n^n}{n!}\right)^{2\tau+3} \;=\; \frac{q^3\mu^{2\tau+3}}{2\,C}\sup_{n\in\mathbb{N}}\left(e^{1-1/n}\right)^{2\tau+3} \;=\; \frac{q^3\mu^{2\tau+3}}{2\,C} \tag{3.27}$$

Proposition 3. *Consider the Lie algebra $\mathbb{V}_{symm}$ with the scale of Banach norms* (3.16)*. Let $V, Q \in \mathbb{V}_r$ for some $r \in \mathbb{R}_+$.*

Define two operators $\mathcal{R}, \mathcal{N}$ by (3.23) *and an operator Γ by* (3.25)*.*

Assume there exist real numbers $\gamma, \rho, \Delta > 0$, $\tau > 1$, $0 < q < 1$ and $-1 < x_0 < 1$ such that:

(1) $x_0, \rho, \Delta, \gamma$ and τ satisfy $\big|\,\rho\Delta x_0 m \,+\, l\,\big| \;\geq\; \gamma\,\big(|l| + |m|\big)^{-\tau}$, $\forall l, m \in \mathbb{Z}_0$;

(2) $|Q_{00}| \;\geq\; q$;

(3) $\|Q\|_r \;\leq\; q^{-1}$;

Then $\forall d, \delta \in \mathbb{R}_+, d+\delta < r$ and $\forall W \in \mathbb{V}_r$, $\forall Z \in \mathbb{V}_{r-\delta}$, the following inequalities hold

$$\|(\Gamma W)Z\|_{r-\delta-d} \;\leq\; \frac{C\,\|W\|_r\,\|Z\|_{r-\delta}}{q^3 d(d+\delta)^{2\tau+2}} \tag{3.28}$$

$$\|\mathcal{N}W\|_{r-\delta} \leq \frac{\tilde{C}\,\|W\|_r}{q^3\,\delta^{2\tau+3}} \tag{3.29}$$

$$\|\mathcal{R}W\|_{r-\delta} \leq \frac{\tilde{C}\,\|W\|_r}{q^3\,\delta^{2\tau+3}} \tag{3.30}$$

where C and $\tilde{C}$ are constants depending on $\tau, \gamma, e, q, \rho, \delta, d$.

Proof. By definition,

$$(\Gamma W)Z \;=\; \{\mathcal{G}_s W\}Z - (\mathcal{A}W)\,\partial_x Z - \{x\mathcal{G}_s Q\mathcal{A}W\}Z - \{x\mathcal{G}_s Q\partial_\theta\mathcal{G}_s\mathcal{P}_0 W\}Z$$

We will study each of the four terms on the r.h.s. separately.

$$\|\{\mathcal{G}_s W\}Z\|_{r-\delta-d} = \sum_{L,M\in\mathbb{Z}} \frac{e^{(r-\delta-d)(|L|+|M|)}}{\rho}\times$$

$$\times\left|\sum_{l,m\in\mathbb{Z}_0} \frac{W_{l,m,1}\,(M-m)\,Z_{L-l,M-m} - m\,(\mathcal{P}_{\leq 1}W_{l,m})\,\partial_x Z_{L-l,M-m}}{x_0\Delta\rho m + l}\right|_{r-d} \leq$$

$$\leq \sum_{L,M\in\mathbb{Z};l,m\in\mathbb{Z}_0} e^{(r-\delta-d)(|L-l|+|M-m|)}\, e^{(r-\delta-d)(|l|+|m|)}\,\frac{(|m|+|l|)^\tau}{\gamma\rho}\times$$

$$\times\left(|W_{l,m,1}|_{r-\delta-d}\,|M-m|\,|Z_{L-l,M-m}|_{r-\delta-d} - |m|\,\left|\mathcal{P}_{\leq 1}W_{l,m})\right|_{r-\delta-d}\left|\partial_x Z_{L-l,M-m}\right|\right)\Bigg|_{r-\delta-d} \leq$$

$$\leq \frac{1}{\rho\gamma}\sum_{L,M,l,m\in\mathbb{Z}} e^{r(|l|+|m|)}\left(\frac{1}{ed}\left(\frac{\tau}{e(d+\delta)}\right)^\tau e^{(r-\delta)(|L-l|+|M-m|)}\,\frac{|W_{l,m}|_r}{d+\delta}\,|Z_{L-l,M-m}|_{r-\delta-d} +\right.$$

$$\left. + \left(\frac{\tau+1}{e(d+\delta)}\right)^{\tau+1} e^{(r-\delta-d)(|L-l|+|M-m|)}\,|W_{l,m}|_{r-\delta-d}\,\frac{|Z_{L-l,M-m}|_{r-\delta}}{d}\right) \leq$$

$$\leq \frac{\tau^\tau + (\tau+1)^{(\tau+1)}}{\gamma\rho e^{\tau+1}(d+\delta)^{\tau+1}d}\,\|W\|_r\,\|Z\|_{r-\delta} \equiv \frac{C_1}{(d+\delta)^{\tau+1}d}\,\|W\|_r\,\|Z\|_{r-\delta}$$

In going from the 4th to the 5th line we used the condition (1); and in the last passage we introduced a constant C_1 for conciseness. Next we consider

$$\mathcal{A}W = \frac{1}{Q_{0,0}}\left(\mathcal{P}_0\sum_{n\geq 0}(n+1)\,W_{0,0,n+1}\,x^n - \oint\sum_{L,M\in\mathbb{Z};l,m\in\mathbb{Z}_0}\frac{me^{iLt+iM\theta}}{x_0\rho\Delta m + l}\,Q_{L-l,M-m}\,W_{l,m,0}\right) =$$

$$= \frac{W_{0,0,1}}{Q_{0,0}} - \frac{1}{Q_{0,0}}\sum_{l,m\in\mathbb{Z}_0}\frac{m}{x_0\rho\Delta m + l}\,Q_{-l,-m}\,W_{l,m,0}$$

So that

$$|\mathcal{A}W| \leq \left|\frac{W_{0,0,1}}{Q_{0,0}}\right| + \sum_{l,m\in\mathbb{Z}_0}\left|\frac{mQ_{-l,-m}W_{l,m,0}}{Q_{0,0}(x_0\rho\Delta m + l)}\right| \leq$$

$$\leq \frac{1}{q}\,\|\partial_x W_{0,0}\|_{r-\delta-d} + \frac{1}{q\gamma}\sum_{l,m\in\mathbb{Z}_0}(|m|+|l|)^{(\tau+1)}\,\|\mathcal{P}_0 W\|_r\,e^{-r(|m|+|l|)}\,|Q_{-l,-m}|$$

where I used $|\mathcal{P}_0 W_{l,m}|_r \leq \|\mathcal{P}_0 W\|_r e^{-r(|l|+|m|)}$. Continuing:

$$|\mathcal{A}W| \leq \frac{\|W\|_r}{q(d+\delta)} + \frac{(\tau+1)^{\tau+1}}{q\gamma(er)^{\tau+1}}\|W\|_r\sum_{l,m\in\mathbb{Z}_0}|Q_{l,m}| \leq$$

$$\leq \frac{\|W\|_r}{q(d+\delta)} + \frac{(\tau+1)^{\tau+1}}{q^2\gamma(er)^{\tau+1}}\|W\|_r \leq \frac{C_2\|W\|_r}{q^2(d+\delta)^{\tau+1}} \qquad (3.31)$$

where in the last passage we used $\delta + d \leq r$ so that $r^{-1} \leq (d+\delta)^{-1}$, and C_2 is a constant. Then for the second term of equation (3.2.2) we have

$$\|(\mathcal{A}W)\partial_x Z\|_{r-\delta-d} \leq \sum_{l,m\in\mathbb{Z}} e^{r(|l|+|m|)}\,|aW|\,\frac{1}{d}\,|Z_{l,m}|_{r-\delta} \leq \frac{C_2\|W\|_r\|Z\|_{r-\delta}}{q^2\,d\,(d+\delta)^{\tau+1}}$$

The third term of equation (3.2.2) reads

$$\{x\mathcal{G}_s Q\mathcal{A}W\}Z \;=\; \frac{1}{\rho}\sum_{L,M\in\mathbb{Z}} e^{iLt+iM\theta}(\mathcal{A}W)\times$$
$$\times\sum_{l,m\in\mathbb{Z}_0}\left(\frac{Q_{l,m}\,(M-m)\,Z_{L-l,M-m}\;-\;x\,m\,Q_{l,m}\,\partial_x Z_{L-l,M-m}}{x_0\rho\Delta m+l}\right)$$

so that $\left\|\{x\mathcal{G}_s Q\mathcal{A}W\}Z\right\|_{r-\delta-d}\leq\frac{|\mathcal{A}W|}{\rho}\sum_{L,M\in\mathbb{Z}} e^{(r-\delta-d)(|L|+|M|)}\times$

$$\times\left|\sum_{l,m\in\mathbb{Z}_0}\frac{Q_{l,m}\,(M-m)\,Z_{L-l,M-m}\;-\;x\,m\,Q_{l,m}\,\partial_x Z_{L-l,M-m}}{x_0\rho\Delta m+l}\right|_{r-\delta-d}\leq$$
$$\leq\frac{|\mathcal{A}W|}{\rho\gamma}\sum_{L,M,l,m\in\mathbb{Z}} e^{(r-\delta-d)(|L-l|+|M-m|)}\;e^{(r-\delta-d)(|l|+|m|)}|Q_{l,m}|\times$$
$$\times\left(|M-m|\,\frac{(|m|+|l|)^{\tau}}{\gamma}\;|Z_{L-l,M-m}|_{r-\delta-d}+|x|_{r-\delta-d}\;\left|\partial_x Z_{L-l,M-m}\right|_{r-\delta-d}\frac{(|m|+|l|)^{\tau+1}}{\gamma}\right)\leq$$
$$\leq\frac{|\mathcal{A}W|}{\rho\gamma}\left(\frac{1}{ed}\,\|Z\|_{r-\delta}\left(\frac{\tau}{e(d+\delta)}\right)^{\tau}\|Q\|_r\;+\;\frac{|r|}{d}\left(\frac{\tau+1}{e(d+\delta)}\right)^{\tau+1}\|Z\|_{r-\delta}\|Q\|_r\right)\;\leq$$
$$\leq\frac{C_3\|W\|_r\|Z\|_{r-\delta}}{q^3\,d\,(d+\delta)^{2\tau+2}}$$

where C_3 is another constant.

Finally, the fourth term of equation (3.2.2) is

$$\{x\mathcal{G}_s Q\partial_\theta\mathcal{G}_s\mathcal{P}_0 W\}Z\;=$$
$$=\left\{x\mathcal{G}_s\sum_{L,M,l,m\in\mathbb{Z}_0} e^{iLt+iM\theta}\frac{mW_{l,m,0}Q_{L-l,M-m}}{x_0\rho\Delta m+l}\,,\;\sum_{l_1,m_1\in\mathbb{Z}} Z_{l_1,m_1}e^{il_1t+im_1\theta}\right\}=$$
$$=\sum_{L_1,M_1\in\mathbb{Z};L,M,l,m\in\mathbb{Z}_0} e^{iL_1t+iM_1\theta}\Big(m\,W_{l,m}\;Q_{L-l,M-m}\,(M_1-M)\,Z_{L_1-L,M_1-M}+$$
$$-\;x\,M\,m\,W_{l,m}\,Q_{L-l,M-m}\,\partial_x Z_{L_1-L,M_1-M}\Big)/\Big(\rho(x_0\rho\Delta M+L)(x_0\rho\Delta m+l)\Big)$$

so that

$$\|\{x\mathcal{G}_s Q\partial_\theta \mathcal{G}_s \mathcal{P}_0 W\}Z\|_{r-\delta-d} \le \sum_{L_1,M_1,L,M,l,m\in\mathbb{Z}} \Bigg(e^{(r-\delta-d)(|L_1-L|+|M_1-M|)} \times$$

$$\times\, e^{(r-\delta-d)(|L|+|M|)} |W_{l,m,0}|\, |Q_{L-l,M-m}|\, \frac{(|m|+|l|)^{\tau+1}}{\rho\gamma^2} \Big(|M_1-M| \times$$

$$\times\, |Z_{L_1-L,M_1-M}|\, (|M|+|L|)^\tau \,+\, |x|(|M|+|L|)^{\tau+1} |\partial_x Z_{L_1-L,M_1-M}|_{r-\delta-d} \Bigg) \le$$

$$\le \frac{\|Z\|_{r-\delta}}{q\rho\gamma^2 d e^{\tau+1}} \|W\|_r \left(\left(\frac{2\tau}{d+\delta}\right)^\tau \left(\frac{2(\tau+1)}{e(d+\delta)}\right)^{\tau+1} + |r| \left(\frac{2(\tau+1)}{d+\delta}\right)^{\tau+1} \left(\frac{2(\tau+1)}{e(d+\delta)}\right)^{\tau+1} \right) \equiv$$

$$\equiv \frac{C_4 \|W\|_r \|Z\|_{r-\delta}}{q\, d\, (d+\delta)^{2\tau+1}}$$

with a fourth constant C_4. By defining

$$C \,=\, (C_1\, q^3 + C_2\, q)(d+\delta)^{\tau+1} + C_3 + (d+\delta) q^2 C_4$$

we end up with the thesis.

To prove the second and third inequalities, we start by observing that

$$\|\mathcal{N}_s V\|_{r-\mu} = \|\chi\mathcal{P}_0 V + \mathcal{P}_1 V\|_{r-\mu} \le \|V\|_{r-\mu}$$
$$\|\mathcal{R}_s V\|_{r-\mu} = \| \oint \mathcal{P}_0 V + \mathcal{P}_2 V\|_{r-\mu} \le \|V\|_{r-\mu}$$

Then we consider

$$\|\mathcal{K}V\|_{r-\mu} \,\le\, |x_0 \Delta\rho\mathcal{A}V| \,+\, \left\|\{\tfrac{1}{2}Qx^2\} x\mathcal{G}_s(\mathcal{P}_0\partial_x - Q\mathcal{A} - Q\partial_\theta\mathcal{G}_s\mathcal{P}_0)V\right\|_{r-\mu}$$

By (3.31),

$$|x_0\Delta\rho\mathcal{A}V| \,\le\, |x_0|\, \rho\Delta \left|\mathcal{A}V\right| \,\le\, \rho\Delta\, \frac{C_2}{q^2\, \mu^{\tau+1}}\, \|V\|_r$$

To the next term we apply equation (3.19) of Proposition 2 with $\delta = d = \mu/2$,

$$\left\|\{\tfrac{1}{2}Qx^2\} x\mathcal{G}_s(\mathcal{P}_0\partial_x - Q\mathcal{A} - Q\partial_\theta\mathcal{G}_s\mathcal{P}_0)V\right\|_{r-\mu} \le$$
$$\le \frac{4}{\rho e \mu^2} \left\|\tfrac{1}{2}Qx^2\right\|_r \left\|x\mathcal{G}_s(\mathcal{P}_0\partial_x - Q\mathcal{A} - Q\partial_\theta\mathcal{G}_s\mathcal{P}_0)V\right\|_{r-\mu/2}$$

We have

$$\left\|\tfrac{1}{2}Qx^2\right\|_r \,\le\, \tfrac{1}{2}\|Q\|_r |x^2|_r \le (2q)^{-1}$$

and

$$\left\|x\mathcal{G}_s\mathcal{P}_0\partial_x V\right\|_{r-\mu/2} \,=\, \sum_{l,m\in\mathbb{Z}_0} \left|\frac{V_{lm1}x}{x_0\rho\Delta m + l}\right|_{r-\frac{\mu}{2}} e^{(r-\mu/2)(|l|+|m|)} \le$$

$$\le \sum_{l,m\in\mathbb{Z}_0} |V_{lm1}x|_{r-\mu/2} \left(|m|+|l|\right)^\tau \gamma^{-1}\, e^{(r-\mu/2)(|l|+|m|)} \,\le\, \frac{1}{\gamma}\left(\frac{2\tau}{e\mu}\right)^{\tau+1} \|V\|_r$$

and

$$\left\|x\mathcal{G}_sQ\mathcal{A}V\right\|_{r-\mu/2} = \left|\mathcal{A}V\right| |x|_{r-\mu/2} \left\|\mathcal{G}_sQ\right\|_{r-\mu/2} \leq$$

$$\leq \left|\mathcal{A}V\right| \sum_{l,m\in\mathbb{Z}_0} \gamma^{-1}(|l|+|m|)^\tau e^{(r-\mu/2)(|l|+|m|)}|Q_{l,m}| \leq$$

$$\leq \frac{C_2}{\gamma q^2\mu^{\tau+1}} \|V\|_r \frac{1}{} \left(\frac{2\tau}{e\mu}\right)^\tau \|Q\|_r \leq \frac{C_2}{\gamma q^3\mu^{\tau+1}} \left(\frac{2\tau}{e\mu}\right)^\tau \|V\|_r$$

And finally

$$\left\|x\mathcal{G}_sQ\partial_\theta\mathcal{G}_s\mathcal{P}_0V\right\|_{r-\frac{\mu}{2}} = \sum_{L,M\in\mathbb{Z}_0} \frac{e^{(r-\frac{\mu}{2})(|L|+|M|)}}{|x_0\Delta\rho M+L|} \left|\sum_{l,m\in\mathbb{Z}_0} \frac{mQ_{L-l,M-m}v_{l,m,0}}{x_0\Delta\rho m+l}\right| \leq$$

$$\leq \sum_{L,M,l,m\in\mathbb{Z}} e^{(r-\frac{\mu}{4})(|L|+|M|)}e^{-\frac{\mu}{4}(|L|+|M|)} \frac{(|M|+|L|)^\tau}{\gamma} \frac{(|m|+|l|)^\tau}{\gamma} |Q_{L-l,M-m}||V_{l,m,0}| \leq$$

$$\leq \left(\frac{4\tau}{e\mu}\right)^\tau \frac{1}{\gamma^2} \sum_{L,M,l,m\in\mathbb{Z}} e^{r(|L|+|M|)}e^{-\frac{\mu}{4}(|L-l|+|M-m|)}e^{-\frac{\mu}{4}(|l|+|m|)} \frac{(|m|+|l|)^\tau}{\gamma} |Q_{L-l,M-m}||V_{l,m,0}| \leq$$

$$\leq \left(\frac{4\tau}{e\mu}\right)^\tau \left(\frac{4(\tau+1)}{e\mu}\right)^{\tau+1} \frac{\|V\|_r}{\gamma^2 q}$$

We can conclude that

$$\|\mathcal{K}V\|_{r-\mu} \leq \frac{\tilde{C}\|V\|_r}{q^3\mu^{2\tau+3}}$$

and so

$$\|\mathcal{N}V\|_{r-\mu} \leq \|\mathcal{N}_sV\|_{r-\mu} + \|\mathcal{K}V\|_{r-\mu} \leq \|V\|_r + \frac{\tilde{C}_1\|V\|_r}{q^3\mu^{2\tau+3}} \equiv \frac{\tilde{C}\|V\|_r}{q^3\,\mu^{2\tau+3}}$$

and analogously

$$\|\mathcal{R}V\|_{r-\mu} \leq \|\mathcal{R}_sV\|_{r-\mu} + \|\mathcal{K}V\|_{r-\mu} \leq \|V\|_r + \frac{\tilde{C}_1\|V\|_r}{q^3\mu^{2\tau+3}} \equiv \frac{\tilde{C}\|V\|_r}{q^3\,\mu^{2\tau+3}}$$

where $\tilde{C}_1, \tilde{C}$ are constants depending on $\mu, \tau, \gamma, q, \rho, \Delta, e$.

□

3.2.3 A KAM theorem for the Symmetric Throbbing Top

Now we prove a KAM theorem for the symmetric Throbbing Top by iteratively applying Proposition 1.

Theorem 1. *Consider the dynamical system* (3.11) *on the algebra* $\mathbb{V}_{symm}$ *and with* $E = E_{symm}$. *Define* Q *as in equation* (3.21) *and* $\rho, \Delta, \gamma, \tau, q \in \mathbb{R}_+$ *as in Proposition 3. Then there exist* $\epsilon_0, r \in \mathbb{R}_+$ *such that if* $\|V\|_r \leq \epsilon$, *the for a large class of initial data, the trajectories of this dynamical system can be mapped to trajectories of a static Top.*

Proof. We have shown in the previous section that Proposition 1 can be applied to the dynamical system we are considering, by choosing $\mathbb{B}$ as in (3.22), $\mathcal{R}$ as in (3.23) and Γ as in (3.25). Thus by formula (2.7) we can map $\mathcal{H} + \{V\}$ into $\mathcal{H}_* + \{V_*\}$. This is possible, in particular, if $X(t = 0) = x_0$ satisfy the Diophantine condition (1) of Proposition 3; as we discussed, the set of these values is large in a measure theoretic sense. We have to choose a loss $\mu_0 \in \mathbb{R}_+$ such that $\mu_0 < r/3$, so that $\|V_*\|_{r-3\mu} < \kappa\, \epsilon_0^2$.

Now we want to show that Proposition 1 can be applied to $\mathcal{H}_* + \{V_*\} \equiv \mathcal{H} + \{RV\} + \{V_*\}$, so with the same values of ρ, Δ and x_0 but

$$Q \mapsto Q^* \stackrel{\text{def}}{=} Q + \partial^2_{xx}(\mathcal{R}V) \qquad r \mapsto r_* \equiv r - \mu$$

and of course $V \mapsto V_*$. We need to verify that there exist a new constant q_* for which hypothesis (2) and (3) of Proposition 3 are again satisfied. By using inequality (3.30) of Proposition 3,

$$\|Q^* - Q\|_{r-\mu} \;=\; \|\partial^2_{xx}\mathcal{R}V\|_{r-\mu} \;\le\; \frac{2}{r^2}\|\mathcal{R}V\|_r \;\le\; \frac{i2\,\epsilon\,\tilde{C}}{r^2 q^3 \mu^{2\tau+3}}$$

where we used equations (3.17), (3.30), and the Cauchy inequality. In the same way

$$|Q^*_{0,0} - Q_{0,0}| \;=\; \Big|\oint \partial^2_{xx}\mathcal{R}V\Big| \;\le\; \frac{2}{r^2}\Big|\oint \mathcal{R}V\Big|_r \;\le\; \frac{2\,\epsilon\,\tilde{C}}{r^2 q^3 \mu^{2\tau+3}}$$

So we have

$$\|Q^*\|_{r-\mu} \;\le\; \|Q\|_{r_0-\mu} \;+\; \|Q^* - Q\|_{r_0-\mu} \;\le\; \frac{1}{q} \;+\; \frac{\epsilon\tilde{C}}{r^2 q^3 \mu^{2\tau+3}} \;\le\; \frac{1}{q - \frac{\epsilon\tilde{C}}{r^2 q^3 \mu^{2\tau+3}}} \;\equiv\; \frac{1}{q_*}$$

where the last inequality holds as long as

$$0 \le \frac{\epsilon\tilde{C}}{r^2 q^3 \mu^{2\tau+}} \le q \le 1$$

This condition is to be confronted with formula (3.27); they are compatible if

$$q \;\ge\; \tilde{C}/(C\,r^2) \tag{3.32}$$

At the same time, $|Q_{0,0}| \le |Q_{0,0} - Q^*_{0,0}| + |Q^*_{0,0}|$ so

$$|Q^*_{0,0}| \;\ge\; |Q_{0,0}| - |Q^*_{0,0} - Q_{0,0}| \;\ge\; q \;-\; \frac{\epsilon\tilde{C}}{q^3 \mu^{2\tau+5}} \;\equiv\; q_*$$

So Proposition 1 can be applied to $\mathcal{H}_* + \{V_*\}$. We may build a sequence of dynamical systems by

$$e^{[\Gamma_i V_i]}(\mathcal{H}_i + \{V_i\}) = \mathcal{H}_{i+1} + \{V_{i+1}\}$$

where $V_0 \equiv V$, and

$$\mathcal{H}_i = x_0\rho\Delta\partial_\theta + \partial_t + \{\tfrac{1}{2}Q^i x^2\}, \qquad Q^0 \equiv Q$$

$$\Gamma_i f = \{\mathcal{G}f\} = \{\mathcal{G}_s f\} - \rho^{-1}\mathcal{A}_i f\partial_x - \{x\mathcal{G}_s Q^i(\mathcal{A}_i + \partial_\theta\mathcal{G}_s\mathcal{P}_0)f\}$$

$$\mathcal{A}_i = \Big(\oint Q^i\Big)^{-1}\oint \mathcal{P}_0(\partial_x - Q^i\partial_\theta\mathcal{G}_s\mathcal{P}_0).$$

The sequence would converge to the static Top

$$\mathcal{H}_\infty \;=\; \mathcal{H} \;+\; \sum_{i=0}^{\infty} \mathcal{R}_i V_i$$

To show that the sequence exists, we need three sequences $\{\epsilon_i, \mu_i, q_i\}_{i\in\mathbb{N}}$ such that

$$\Gamma_i V_i \colon \mathbb{V}_{r_i} \to \mathbb{V}_{r_{i+1}}\,, \quad \|Q^i\|_{r_i} < (q_i)^{-1}\,, \quad |Q^i{}_{0,0}| > q_i\,, \quad \|V_i\|_{r_i} < \epsilon_{\mu_i}^2$$

where $r_i \equiv r - \sum_{j=1}^{i-1} \mu_j$. Moreover they must satisfy:

(a) $\epsilon_i \;=\; q_i^3 \mu_i^{2\tau+3}/(2\,C)$ as we computed at point **5** of section 3.2.2;

(b) $q_i \;\geq \tilde{C}/(C r_i^2)$, coherently with equation (3.32);

(c) $0 < \mu_i < \frac{r_i}{3}$, as required by Proposition 1;

(d) $\sum_{i=1}^{\infty} \mu_i < r$, to ensure that the operator $\mathcal{H}_\infty$ is well defined on $\mathbb{V}_{r_\infty}$, being

$$r_\infty = r - \sum_{j=1}^{\infty} \mu_j > 0$$

(e) $\lim_{i\to\infty} \epsilon_i = 0$, so we can conclude that $V_\infty \stackrel{\text{def}}{=} \lim_{i\to\infty} V_i = 0$;

(f) $0 < q_\infty < q_i < 1, \quad \forall i \in \mathbb{N}$ as required by Proposition 3;

We choose

$$\epsilon_i \;=\; \frac{\epsilon_0}{(i+1)^{2(2\tau+3)}}, \qquad \mu_i \;=\; \frac{1}{(1+i)^2}\left(\frac{2\,C\,\epsilon_0}{q_i^3}\right)^{\frac{1}{2\tau+3}}$$

so that condition (a) is satisfied. Also condition (e) is evidently satisfied. Now we compute

$$\sum_{i=0}^{\infty} \mu_i \;=\; \sum_{i=0}^{\infty} \frac{1}{(1+i)^2}\left(\frac{2\,C\,\epsilon_0}{q_i^3}\right)^{\frac{1}{2\tau+3}} \;\leq\; \left(\frac{2\,C\,\epsilon_0}{q_\infty^3}\right)^{\frac{1}{2\tau+3}} \sum_{i=1}^{\infty} \frac{1}{i^2} \;\leq\; \frac{\pi^2}{6}\left(\frac{2\,C\,\epsilon_0}{q_\infty^3}\right)^{\frac{1}{2\tau+3}}$$

Both conditions (c) and (d) are satisfied by imposing $\sum_{i=0}^{\infty} \mu_i < r/3$ and, by the result above, we get

$$\left(\frac{2\,C\,\epsilon_0}{q_\infty^3}\right)^{\frac{1}{2\tau+3}} \;\leq\; \frac{2\,r}{\pi^2} \tag{3.33}$$

Then for the sequence $\{q_i\}_{i\in\mathbb{N}}$ we propose

$$q_i = q_{i-1}\Big(1 - \frac{1}{(i+1)^2}\Big) = q_0 \prod_{j=2}^{i+1}\Big(1 - \frac{1}{j^2}\Big)$$

By taking the logarithm of both sides, and using that $\log(1-x) \geq -\log(4)x$ for $x \in [\frac{1}{2}, 1]$, we get

$$\log(q_i) = \log(q_0) + \sum_{j=2}^{i+1} \log\Big(1 - \frac{1}{j^2}\Big) \geq$$

$$\geq \log(q_0) - \sum_{j=2}^{i+1} \frac{\log(4)}{j^2} \geq \log(q_0) - \sum_{j=1}^{\infty} \frac{\log(4)}{j^2} = \log(q_0\, 2^{-\pi^2/3})$$

We set $q_\infty = q_0\, 2^{-\pi^2/3}$ and $q_0 < 1$ so condition (f) is satisfied. If we plug this value for q_∞ into equation (3.33) we get

$$\epsilon_0 \leq \frac{q_0^3}{C}\left(\frac{r}{\pi^2}\right)^{2\tau+3} 2^{(4+2\tau-\pi^2)(2\tau+3)/(1-\pi^2)} \tag{3.34}$$

Finally, we rewrite condition (b) as $r^2 \geq \tilde{C}/(qC)$ and, being $1/q \geq 1$, we get a lower bound on r,

$$r \geq \sqrt{\tilde{C}/C}$$

□

3.2.4 Numerical experiments

In this section we show some numerical simulations of dynamical system (3.11) with the Hamiltonian given by (3.14) and a perturbation V like in equation (3.10) with the matrix A given by (3.12), so that

$$V = \frac{\epsilon}{2} M_2^2 \cos(t) = \frac{\rho^2 \epsilon}{2}(1 - X^2)\sin^2(\theta)\cos(t) \tag{3.35}$$

That is, we solved the Euler-Poinsot equations (3.1) with $I_2^{-1} \mapsto (I_2^{-1} + \epsilon\cos(t))^{-1}$. We employed a Runge-Kutta numerical scheme of 4th order, with a time step equal to 10^{-4}. We checked for conservation of the Casimir ρ, and, in the unperturbed (thus static) case, also of energy.

The system is described in the four coordinates ρ, X, θ, t, but the trajectories can be visualized in 2-dimensional plots. Indeed, ρ is constant being a Casimir, while, to get rid

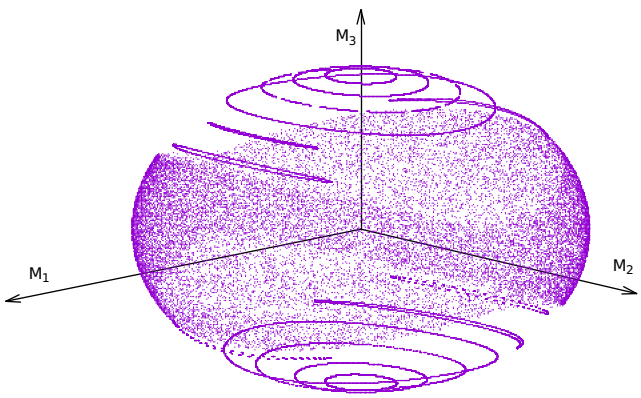
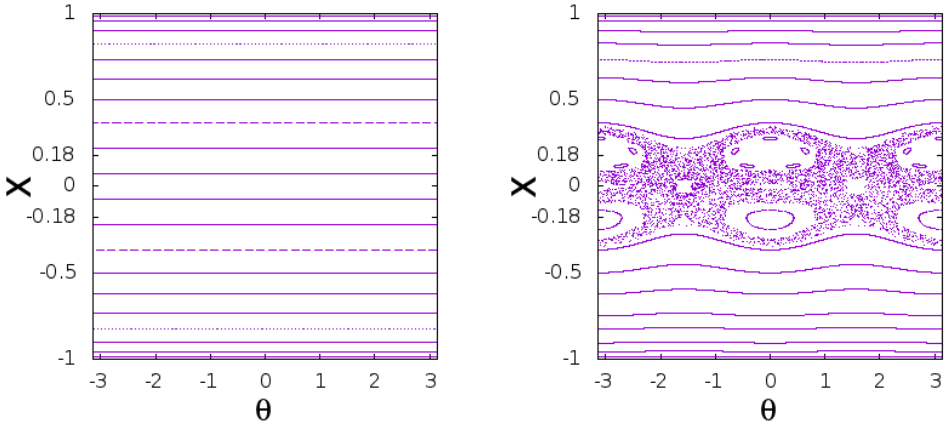
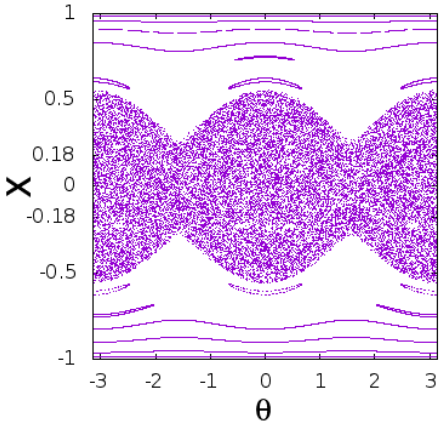


Figure 3.2: *A phase portrait for the model described in this subsection 3.2.4. So, the perturbation is given by Eq.* (3.35)*. Numerical values of the parameters are* $I_1 = I_2 = 1, I_3 = \frac{1}{3}$ *and* $\epsilon = 0.4$*. As one can see, for these intermediate values of the perturbation, some of the tori (those near to the separatrix) are broken, others persist.*

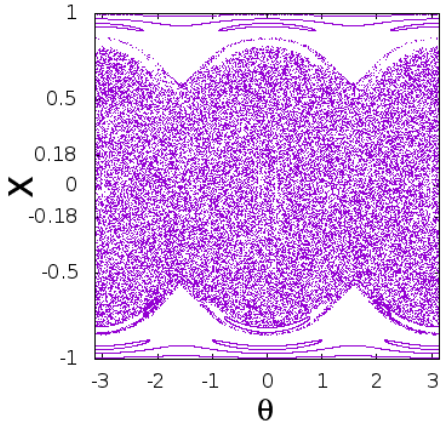


(a) $\epsilon = 0$

(b) $\epsilon = 0.1$



(c) $\epsilon = 0.4$. These points are same that were used to plot figure 3.2



(d) $\epsilon = 1$

Figure 3.3: *Symmetric case, with perturbation given by Eq.* (3.35). *Here* $I_1 = I_2 = 1$ *and* $I_3 = \frac{1}{3}$.

of time, we used the "time-2π-mapping": it is the map which results from the solution of a periodic dynamical system, by taking the evolved points only at instants which are integer multiples of the period of the perturbation, in this case 2π. In practice, the code solved the equations in the M_1, M_2, M_3 coordinates, we took their values at each time step for which $t = 2\pi n,\ n \in \mathbb{N}$, and then compute X and θ.

In figure 3.3 we can see the 4 plots relative to different values of ϵ. For $\epsilon = 0$, the trajectories are straight lines: the system is integrable and the trajectories are level surfaces of the Hamiltonian (which is the same as being surface levels of x). When $\epsilon > 0$, this is no more true and chaos appears. We see that it starts to appear around the *resonances*, which are the values $X = x_0$, where the diophantine condition (and thus theorem 1) is broken. They are solutions of

$$\rho \Delta x_0 m + l = 0$$

for the unknown x_0. In our simulations we fixed (arbitrarily) $\rho = \sqrt{2}$ and $\Delta = 2$. Because

of equation (3.35), $m \in \{0, \pm 2\}$ and $l \in \{\pm 1\}$. Thus the resonant values of x_0 are

$$x_0 = -\frac{l}{2\sqrt{2}m}, m \neq 0 \implies x_0 = \pm\frac{1}{4\sqrt{2}} \approx \pm 0.18$$

for any value of ϵ. Indeed, it is around these points that chaos appears; this is particularly evident in the upper right plot of figure 3.3.

In figure 3.4 we report the results of simulations for a slightly different perturbation,

$$V = \frac{\epsilon}{2} M_2^2 \cos^2(t) = \frac{\rho^2 \epsilon}{2}(1 - x^2)\sin^2(\theta)\cos^2(t) \tag{3.36}$$

so that in this case $l, m \in \{0, \pm 2\}$, and the resonances are located around

$$x_0 = \pm\frac{1}{2\sqrt{2}} \approx \pm 0.35 \quad \text{and} \quad x_0 = 0$$

Again, this are precisely the points around which we see that the tori are broken, even for very low values of ϵ.

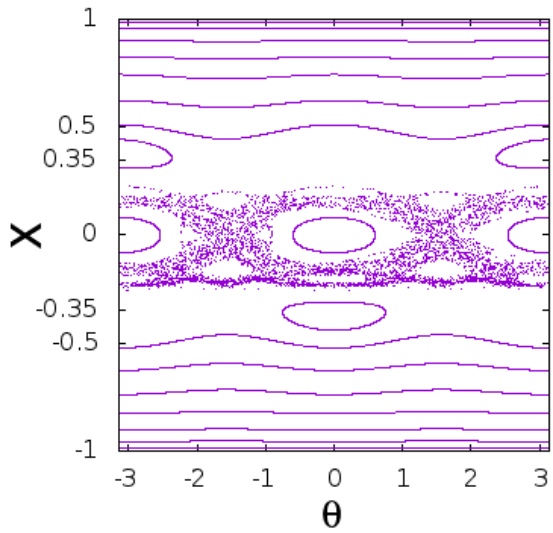


Figure 3.4: *Symmetric case, with the perturbation given by Eq.* (3.36)*. Here* $I_1 = I_2 = 1$, $I_3 = \frac{1}{3}$, $\epsilon = 0.1$.

3.2.5 About the Diophantine condition

The diophantine condition, that we required in hypothesis (1) of Proposition 3 was introduced by Kolmogorov in his paper of 1954 [47]. Here we want to show that the set of values of x_0 which satisfy the Diophantine condition have a great relative measure in $(-1, 1)$.

Let $\rho, \Delta \in \mathbb{R}$ be positive constant. The equality

$$\rho\,\Delta\,x_0\,m + l = 0 \tag{3.37}$$

denotes a straight line in the m, l plane. If m, l were continuous variables, any value x_0 would denote a straight line of angular coefficient $k \equiv x_0\rho\Delta$, with $k \in (-\rho\Delta, \rho\Delta)$. But

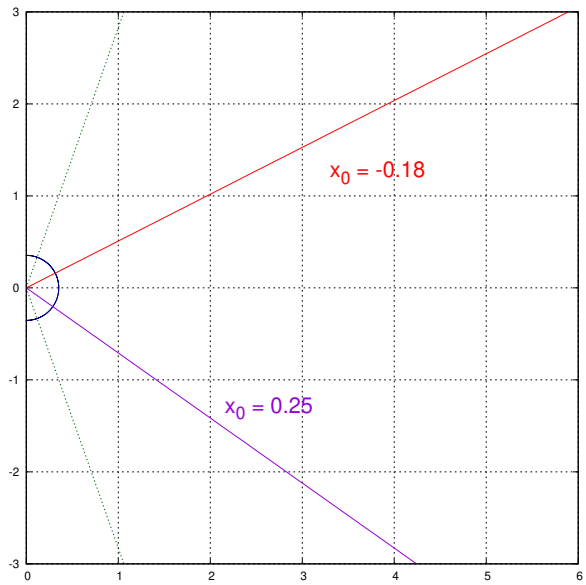


Figure 3.5: *In this picture we plot the portion of the m, l plane with positive m. We show a grid to underline the integer values, because we are mostly interested by them. We also plot a straight line given by equation* (3.37)*, with $\rho = \sqrt{2}$ and $\Delta = 2$ (these values are chosen to make some comparison with section 3.2.4. Two green lines correspond to $x_0 = \pm 1$, which are the limit values for x_0. One (purple) line corresponds to a value chosen randomly of $x_0 = 0.25$ and finally a red line corresponds to $x_0 = -0.18$, a value chosen again to confront with 3.2.4. In fact we see that this line passes very near to the point* $(2, 1)$*, a fact that we underlined in the previous section.*

m, l are descrete, so there are only certain values of x_0 which can satisfy the equality: those for which k is rational.

Now consider the inequality

$$|\rho\Delta\, x_0 m + l| \;<\; \Psi(|l|, |m|) \tag{3.38}$$

with $\Psi : \mathbb{Z}_0^2 \to \mathbb{R}$ positive definite. Geometrically, the inequality above defines a "strip" of width 2Ψ around each of the straight lines defined by equation (3.37) which has a rational angular coefficient $x_0\rho\Delta$. Moreover the points x_0 which satisfy the inequality (3.38) evidently are those which do not satisfy the diophantine condition; we will show now that they constitute a "small" (in a measure theoretic sense) subset of $(-1, 1)$.

The segment $(-1, 1) \ni x_0$ can be identified with the arc on the circle of radius $r \equiv 1/(2\,\mathrm{atan}(\rho\Delta))$ centered at the origin and bounded by the intersections with the two straight lines of angular coefficiens $\pm\rho\Delta$ (see figure 3.5). Its total length is 1. We cut away from this arc all the small arcs which are the intersections with the strips defined by (3.38), and we assume $\Psi(l, m) \equiv \varphi(|l| + |m|)$. In the end we cut away a set of measure

$$\mu \;=\; \sum_{l,m\in\mathbb{Z}_0} \Psi(l, m) \;=\; \sum_{s>0} \varphi(s) \sum_{l,m\in\mathbb{Z}_0 \mathrm{s.t.} |l|+|m|=s} \;=\; \sum_{s>0} 2(s-1)\varphi(s)$$

Let $\mathbb{R} \ni \gamma, \tau > 0$, and define $\varphi(s) = \gamma s^{-\tau}$. Now

$$\mu = 2\gamma\Big(\sum_{s>0} \frac{1}{s^{\tau-1}} - \sum_{s>0} \frac{1}{s^{\tau}}\Big)$$

so that the term in parenthesis is bounded for $\tau > 2$. We see that $\mu \propto \gamma$, and γ can be taken arbitrarily small.

3.3 The case of the Non-Symmetric Throbbing Top

In this section we tackle the problem of the Throbbing Top, with $I_1 > I_2 > I_3$. The dynamical system is again (3.11), set on the algebra (3.9) with $t \in \mathbb{R}$. We recall that the bracket is again given by (3.8) in accordance to the Lie-Poisson theorem.

We will consider as unperturbed the flow of the autonomous system, so $\mathbb{B} \equiv \ker\big(\partial_t\big)$. This is a subalgebra because the bracket is time-independent. Moreover $\mathbb{B}$ is invariant by $\mathcal{H}$ since $H \in \mathbb{B}$ while ∂_t gives 0 while acting on $\mathbb{B}$. By using Proposition 1 we will prove that under suitable conditions on the perturbation, the trajectories of the Throbbing Top stick to those of a static top. But first we need to build a new scale of Banach norms.

3.3.1 A new scale of norms

The Poisson structure for the Top is built by applying the Lie-Poisson theorem to the algebra $\mathbb{R}^3 \ni \boldsymbol{M}$, (we recall that the bracket is given by the cross product). Now we apply to $\mathbb{R}^3$ by the general complexification method described in the appendix A, section A.2; the result is

$$\mathbb{C}^3 \stackrel{\text{def}}{=} \mathbb{R}^3 \oplus \mathbb{R}^3 \tag{3.39}$$

with dual $\mathbb{C}^{3^*}$. Moreover, as the space $\mathbb{C}^3$ is metric, we can identify $\mathbb{C}^{3^*} = \overline{\mathbb{C}^3}$.

The Lie-Poisson theorem can be applied to (3.39), to get the new Lie-Poisson algebra

$$\mathbb{V}_C \stackrel{\text{def}}{=} \mathcal{C}^\infty(\overline{\mathbb{C}^3} \to \mathbb{C})$$

with the Lie-Poisson bracket

$$\big(\{f\}g\big)(\bar{Z}) \;=\; \bar{Z}\big[\!\big[\partial_{\bar{Z}} f\big]\!\big]\partial_{\bar{Z}} g$$

and the Casimir $\rho_C^2(\bar{Z}) \;=\; \bar{Z}Z$. We can also define the non-autonomous problem on the algebra

$$\tilde{\mathbb{V}_C} \stackrel{\text{def}}{=} \mathcal{C}^\infty(\mathbb{R} \to \mathbb{V}_C)$$

again with the same (time-independent) bracket of $\mathbb{V}_C$. A few remarks:

- the transposition on $\mathbb{C}^3$ obeys (A.18) and it shouldn't be confused with complex conjugation, denoted by $Z^\dagger$;
- the quantity ρ_C is not positive definite (but this is not a problem);
- however, when the imaginary part of $\bar{Z}$ is zero, the bracket on $\mathbb{V}_C$ reduces to that of $\mathbb{V}$, and the Casimir ρ_C becomes the real Casimir ρ;
- as a consequence of last point, we have that $\mathbb{V}_{TT} \subset \tilde{\mathbb{V}}_C$.

Let us consider the set

$$\mathbb{E} \stackrel{\text{def}}{=} \big\{\, \bar{Z} \in \bar{\mathbb{C}}^3\,, \quad \rho_C(\bar{Z}) \;=\; \rho_C^{\mathbb{E}} \,\big\} \quad \supset \quad \mathbb{S}_2$$

Any dynamics of the complexified bracket takes place on $\mathbb{E}$ because it preserves the Casimir ρ_C; again, for zero imaginary part of $\bar{Z}$ the set $\mathbb{E}$ reduces to the sphere $\mathbb{S}_2$.

Finally we introduce the domain

$$\mathbb{E}_\beta \stackrel{\text{def}}{=} \bigcup_{\bar{Z}\in\mathbb{E}} \mathbb{B}_\beta(\bar{Z})$$

where $\mathbb{B}_\beta(\bar{Z})$ is a ball in $\mathbb{C}^3$ centered in $\bar{Z}$ and of radius β. Now we can define a scale of Banach spaces, endowing, for all $r \in \mathbb{R}_+$. $\mathbb{E}_\beta$ with the Banach norm

$$\big\|f\big\|_\beta \stackrel{\text{def}}{=} \sup_{\bar{Z}\in\mathbb{E}_\beta,\, t\in\mathbb{R}} \big|f(\bar{Z},t)\big| \tag{3.40}$$

which bound from above, for any r, the supremum norm on $\mathbb{V}$

$$\big\|f\big\|_\infty \stackrel{\text{def}}{=} \sup_{\overline{\boldsymbol{M}}\in\mathbb{S}^2,\, t\in\mathbb{K}} \big|f(\overline{\boldsymbol{M}},t)\big|$$

The derivative of a function $f \in \mathbb{V}_C$ is defined as

$$\bar{\zeta}\partial_{\bar{Z}} f(\bar{Z}) \;=\; \lim_{\eta\to 0} \frac{f(\bar{Z}+\bar{\zeta}\eta) \;-\; f(\bar{Z})}{\eta}$$

so it belongs to $\mathbb{C}^3 \otimes \tilde{\mathbb{V}}_C$. A natural extension of the norm (3.40) to this set is

$$\|\partial_{\bar{Z}} f(\bar{Z})\|_\beta \stackrel{\text{def}}{=} \sup_{\bar{W}\in\bar{\mathbb{C}}^3} \sup_{\overline{\boldsymbol{Z}}\in\mathbb{E}_\beta,\, t\in\mathbb{R}} \frac{\big|\bar{W}\partial_{\bar{Z}} f(\bar{Z},t)\big|}{|\bar{W}|_3}$$

where $|\quad|_3$ is the norm on $\mathbb{C}^3$, defined by $|Z| \stackrel{\text{def}}{=} \sqrt{\bar{Z}^\dagger Z}$, and extended by duality to $\mathbb{C}^{3*}$, i.e. $|\bar{Z}| \equiv |Z|$. We assume that on $\mathbb{E}$ the following variant of the Cauchy equality holds,

$$\bar{W}\partial_{\bar{Z}} H(\bar{Z}) \;=\; \oint d\theta e^{-i\theta} H(\bar{Z} \,+\, \bar{W}e^{i\theta})$$

by which we can bound (with loss) the derivation operator on $\mathbb{V}_C$. Indeed, for any $\bar{\delta} \in \bar{\mathbb{C}}^3$, with $\delta \;=\; |\bar{\delta}| \in \mathbb{R}_+$ we have

$$\big\|\partial_{\bar{Z}} f\big\|_\beta \;\le\; \frac{\big\|\bar{\delta}\partial_{\bar{Z}} f\big\|_\beta}{|\bar{\delta}|} \;\le\; \sup_{\bar{Z}\in\mathbb{E}_\beta} \frac{\big|\oint d\theta e^{-i\theta} f(\bar{Z} \,+\, \bar{\delta}e^{i\theta})\big|}{|\bar{\delta}|} \;\le\; \sup_{\bar{Z}\in\mathbb{E}_\beta,\bar{\zeta}\in\mathbb{E}_\delta(\bar{Z})} \frac{|f(\bar{\zeta})|}{\delta} \;\le\; \frac{\big\|f\big\|_{\beta+\delta}}{\delta}$$

We can also bound the bracket. Let δ, d be positive reals such that $d < \delta, d+\delta < \beta$; we have

$$\big\|\{f\}g\big\|_{\beta-\delta-d} \;=\; \sup_{\bar{Z}\in\mathbb{E}_{\beta-\delta-d},\, t\in\mathbb{R}} \Big|\bar{Z}\big[\big[\partial_{\bar{Z}} f\big]\big]\partial_{\bar{Z}} g\Big| \;\le$$

$$\le \sup_{\bar{Z}\in\mathbb{E}_{\beta-\delta-d},\, t\in\mathbb{R}} \big|\bar{Z}\big|\,\big|\partial_{\bar{Z}} f\big|\,\big|\partial_{\bar{Z}} g\big| \;\le\; \frac{\rho_C^{\mathbb{E}}}{d(d+\delta)}\big\|f\big\|_\beta\,\big\|g\big\|_{\beta-\delta} \tag{3.41}$$

Now we have all the tools to formulate the following:

Proposition 4. *Consider the dynamical system* (3.11) *on the algebra* (3.9) *with* $t \in \mathbb{R}$. *Assume that the perturbation* $V(\overline{\boldsymbol{M}},t) \in \mathbb{V}_T T$ *has compact time support* $[0,T]$. *If there exists two positive real numbers* β *and* ε *such that* $\|V\|_\beta \le \epsilon_0$ *and* $2\pi^2\sqrt{T\,\epsilon_0\,\rho_C} < 3r$, *then there exists a static Top with Hamiltonian* $\tilde{H}$ *whose flow is conjugated to that of* H_{TT}.

Proof. Having chosen the invariant subalgebra $\mathbb{B}$ as the set of time independent functions, a choice for the operators $\mathcal{R}$ and Γ is

$$\mathcal{R}f(\overline{\boldsymbol{M}},t) \;=\; f(\overline{\boldsymbol{M}},0) \tag{3.42}$$

$$\Gamma f \stackrel{\text{def}}{=} \{\mathcal{G}f\}\,, \quad (\mathcal{G}f)(t) \stackrel{\text{def}}{=} \int_0^t d\sigma e^{(\sigma-t)\{H\}}(\mathcal{N}f)(\sigma)\,,\, \mathcal{N} \stackrel{\text{def}}{=} 1-\mathcal{R} \tag{3.43}$$

It's evident that $\mathcal{R}$ takes values in $\mathbb{B}$ and that $\mathcal{R}^2 = \mathcal{R}$. To check hypotheses 2.4, after recalling the remark 2, we consider

$$\big(\{H\} + \partial_t\big)(\mathcal{G}f)(t) = \{H\}\int_0^t d\sigma e^{(\sigma-t)\{H\}}(\mathcal{N}f)(\sigma) +$$

$$+ e^{(\sigma-t)\{H\}}(\mathcal{N}f)(\sigma)|_{\sigma=t} - \int_0^t d\sigma e^{(\sigma-t)\{H\}}\{H\}(\mathcal{N}f)(\sigma) = (\mathcal{N}f)(t) \quad (3.44)$$

Hypothesis 2.3 is proven without much effort,

$$\big\|\mathcal{N}f\big\|_{\beta-\delta} = \big\|f(\bar{Z},t) - f(\bar{Z},0)\big\|_{\beta-\delta} \leq 2\big\|f\big\|_{\beta}$$

so that $\Xi = 2$. To find a function Λ which satify hypothesis 2.5, we start by using equation (3.41),

$$\big\|\{\mathcal{G}V\}g\big\|_{\beta-\delta-d} \leq \frac{1}{d(d+\delta)}\big\|(\mathcal{G}V)\big\|_{\beta}\big\|g\big\|_{\beta-\delta}, \forall g \in \mathbb{V}_{TT}$$

Then we consider

$$\big\|\mathcal{G}f\big\|_{\beta} \leq \sup_{\bar{Z}\in\mathbb{E}_{\beta},t}|t|\sup_{0\leq\sigma\leq t}\Big|e^{(\sigma-t)\{H\}}(\mathcal{N}f)(\overline{\boldsymbol{M}},\sigma)\Big| \leq \sup_{\bar{Z}\in\mathbb{E}_{\beta},t}2|t|\Big|V\Big|$$

where in the last passage we used the fact that the set $\mathbb{D}_\beta$ is invariant by $e^{(\sigma-t)\{H\}}$. If now we use the hypothesis that V has compact time support $[0,T]$, we find

$$\big\|(\Gamma V)g\big\|_{\beta-\delta-d} \leq \frac{2\,T\,\rho_C}{d(d+\delta)}\big\|V\big\|_{\beta}\big\|g\big\|_{\beta-\delta}, \forall g \in \mathbb{V}_{TT} \quad (3.45)$$

which is a inequality of type 2.5, with $\Lambda(d,\delta)) = (2\,\rho_C\,T)/(d(d+\delta))$. And we can compute

$$\begin{aligned}\epsilon_\mu &= \frac{1}{2}\sup_{n\in\mathbb{N}}\left(\frac{1}{n!}\prod_{j=1}^{n}\Lambda\Big(\frac{\mu}{n},\frac{(j-1)\mu}{n}\Big)\right)^{-1/n}\\ &= \frac{1}{2}\sup_{n\in\mathbb{N}}\left(\frac{1}{n!}\prod_{j=1}^{n}\frac{2\,\rho_C\,T\,n^2}{\mu^2\,j}\right)^{-1/n} \leq \frac{\mu^2}{4\,\rho_C\,T}\sup_{n\in\mathbb{N}}e^{(}1/n-1) \equiv \frac{\mu^2}{4\,\rho_C\,T}\end{aligned}$$

We can also prove a simple iteration mechanism. Indeed, by the perturbation formula we get a new Hamiltonian $\mathcal{H}_* = \partial_t + \{H + \mathcal{R}V\}$, but we can keep the same operator $\mathcal{R}$ (3.42) and define the new operator $\mathcal{G}_*$ as

$$\mathcal{G}_*f(t) \stackrel{\text{def}}{=} \int_0^t d\sigma e^{(\sigma-t)\{H+\mathcal{R}V\}}(\mathcal{N}f)(\sigma)$$

and it's evident now that it will satisfy the same bound (3.45). In general at ith iteration we will have

$$\mathcal{G}_if(t) \stackrel{\text{def}}{=} \int_0^t d\sigma e^{(\sigma-t)\{H+\sum_{k=1}^{i}\mathcal{R}V_k\}}(\mathcal{N}f)(\sigma), \qquad V_0 = V\,,\,\mathcal{G}_0 = \mathcal{G}\,,\,\mathcal{G}_1 = \mathcal{G}_*$$

To perform the iteration we need then two sequences $\{\epsilon_i,\mu_i\}_{i\in\mathbb{N}}$ representing, respectively, the perturbation strength and the loss of domain at i-th step. They are related by $\epsilon_i = \mu_i^2/(4T\rho_C)$. We require

$$\lim_{i\to\infty}\epsilon_i = 0 \qquad \sum_{i=0}^{\infty}\mu_i < \beta/3$$

The limit implies that the final perturbation will be equal to 0; the inequality implies that after the iteration the new static dynamical system will be defined on $\mathbb{V}_{\beta_\infty}$ with $\beta_\infty > 2\beta/3$. We choose (ansatz)

$$\epsilon_i = \frac{\epsilon_0}{(i+1)^4} \qquad \Longrightarrow \mu_i = \frac{2\sqrt{T\,\rho_C\,\epsilon_0}}{(i+1)^2}$$

Evidently $\epsilon_i \to 0$ for $i \to \infty$ while

$$\sum_{i=0}^{\infty} \mu_i = 2\sqrt{\epsilon_0\,T\,\rho_C}\sum_{i=1}^{\infty}\frac{1}{i^2} = \frac{2\pi^2}{6}\sqrt{T\epsilon_0\rho_C} \Longrightarrow \frac{\sqrt{T\epsilon_0\rho_C}}{\beta} < \frac{3}{2\pi^2}$$

We can conclude that the flow of H_{TT} is conjugated to that of $\tilde{H} = H + \sum_{i=0}^{\infty}\mathcal{R}V_i$. $\square$

3.4 Conclusion

In this chapter we applied a general perturbation theory for Poisson systems, presented in chapter 2, to study the dynamics of a non-autonomous Top.

We first considered a symmetric periodic Top. We found many common elements with classical mechanics, as we have a set of action-angle-like coordinates: two angles, which are the azimuthal angle and time, plus the cosine of the polar angle, which behaves like the action associated to the azimuthal angle[6]. However, we also see some novelties: one angle does not have an associated action[7] and it does not appear even in the bracket; also, the phase space has not the structure of a cotangent bundle. We introduced a scale of Banach norms on the algebra of functions, and proved that our algebraic perturbation formula can be applied. Moreover, we have shown that the process can be iterated, by following the standard KAM procedure, to conjugate the Throbbing Top to a Static Top.

We also considered a non-symmetric Top, with a non-periodic prescribed time-dependence. In this case even more novelties appear: we don't have action-angle variables nor a canonical structure. The concept of invariant "tori" was replaced by a properly chosen invariant subalgebra of the phase space: the set of time-independent observables. It is "invariant" as a consequence of our choice about the splitting of the full evolution operator into "main part" and "perturbation". We see more analogies with Floquet theory [35], [25], [90] rather than with KAM theory; this makes the result even more interesting, as a generalization of Floquet theory for non-periodic dynamical systems has not yet been proposed. But we still borrowed some ideas of KAM theory: to introduce a scale of Banach norms we used a complexification of the dominion, and a 3-dimensional generalization of the Cauchy inequality. The algebra of the Top is built by the Lie-Poisson theorem applied to the Lie algebra (phase space) $(\mathbb{R}^3, [\![\ \ ]\!])$. To get the complexified Top we don't complexify directly the Top algebra; we complexified instead the phase space, and then applied again the Lie-Poisson theorem. In this way we naturally get the new bracket (with the complexified Casimir). We were able again to apply iteratively our perturbation theory to conjugate also in this case, the non-autonomous Top to an autonomous one, for a perturbation with compact support in time. Note that this second analysis can be

[6]Note that the polar angle has the units of an angle, in the sense that it is measured in radiants; however, it takes values on the compact $[0, \pi]$, and not on the torus.

[7]In classical mechanics, when the Hamiltonian doesn't depend on all of the actions, the system is called *degenerate*. For a KAM theorem for degenerate systems see for instance [4] and [77].

applied, as a special case, to a symmetric top in the case of a non-periodic perturbation, which wasn't considered in the first part of the chapter.

The investigation the non-symmetric top with periodic perturbation is left for future work. Another improvement to look for is related to convergence with respect to the time variable, in the case of a non-periodic system: we assumed finite time support for the perturbation, $t < T$, to get an upper bound on the perturbation strength of order $\epsilon \approx T^{-1}$, as in the theory of adiabatic invariants. It would be interesting to go beyond this type of estimate, by introducing some scaling of the norm also with respect to time. Finally, a difference emerged between the two examples we considered: in the second one (non-periodic non-symmetric Top) the operator Γ was defined for any function of the algebra $\mathbb{V}_{TT}$; however, it was bounded only for the perturbation V. Instead, in the other example (symmetric and periodic Top) Γ was defined by fixing Q, so it's related to a particular system; by suitable hypothesis on Q, it is bounded when acting on any function of $\mathbb{V}_{TT}$. We feel not at ease with this behaviour, as if there is a more elegant definition of Γ waiting to be unveiled.

4

The Charged Particle

In this chapter we consider the following problem: to characterize the dynamics of a relativistic charged particle under the influence of external electric and magnetic fields (or simply, an electromagnetic field). The fields are fixed: we assume that retroaction of the particle on the field is negligible. The dynamical system, as well as its geometric and Poisson strutures, are discussed in section 4.1.

This problem is of interest for many reasons. On a fundamental level, the Maxwell-Vlasov equations provide a self-consistent description of a set of charged particles interacting with an electromagnetic field. These equations may be solved perturbatively: the fluctuations of the fields are perturbations of the trajectories in a fixed field. For the laboratory and industrial applications, knowledge of the particles trajectory is a key element to design magnetic fields for an efficient particle confinement, in view, for instance, of magnetized fusion. For the numerical applications, a partial knowledge of the trajectories may lead to a semplification of the equations of motion, allowing to speed up the computations.

While finding an analytical solution for the trajectory of a particle in an arbitrary field is an impossible task, we aim instead at developing some methods to reduce the number of degrees of freedom of the problem. Indeed, in this chapter we will consider two approaches of this type. We call the first one, described in section 4.2, an "Eulerian reduction": we fix half of the degrees of freedom by computing the velocity vector field associated to the Lorentz force. This vector field is the result of a partial differential equation, so that a formal analogy with the Eulerian description of a fluid arises, hence the name. We also derive the Poisson structure in this picture, as well as some simple solutions. The second reduction theory that we consider, described in section 4.3, is a new approach to the well known Guiding Centre Theory, proposed by Di Troia in [23]. The key is the geometric (not perturbative!) definition of the Guiding Centre as the reference frame in which a charged particle is seen moving on a closed and periodic trajectory. Interestingly enough, the Guiding Centre velocity happens to be a field again. Here we compute for the first time the Guiding Centre equations of motion, and their Poisson structure, in the relativistic regime. We find also a curious geometrical consequence, that the Guiding Centre is not a particle.

4.1 The dynamical system

The equations of motion for a particle of mass m and charge e, subject to an electric field $\boldsymbol{E}$ and a magnetic field $\boldsymbol{B}$ are

$$\begin{cases} \dfrac{d\boldsymbol{x}}{dt} = \dfrac{\boldsymbol{w}}{m\gamma} \\ \dfrac{d\boldsymbol{w}}{dt} = e\boldsymbol{E} - \dfrac{e}{m\gamma}\left[\!\left[\boldsymbol{B}\right]\!\right]\boldsymbol{w} \qquad \gamma = \sqrt{1 + \dfrac{w^2}{m^2}} \end{cases} \tag{4.1}$$

Here t denotes time and $\boldsymbol{x}$ the position of the particle in some ("laboratory") reference frame. The velocity of the particle is $\boldsymbol{v} \stackrel{\text{def}}{=} d\boldsymbol{x}/dt$ while $\boldsymbol{w} \stackrel{\text{def}}{=} m\gamma\boldsymbol{v}$ is the kinetic momentum. The particle kinetic energy is $m\gamma = \sqrt{m^2 + \overline{\boldsymbol{w}}\boldsymbol{w}}$.

We considered the equations of motion in cgs units and fixed the physical dimensions as follows. The 3 fundamental dimensions are a length L, a time T, a mass M. c, e, m are fundamental constants. We set $c = 1$, so L = T and $v = dx/dt$ is dimensionless. Instead $[m] = \mathrm{M}$ as one would expect. Then we have that $[w] = \mathrm{M}, [\varepsilon] = \mathrm{M}$. Finally, e appears always next to a field so it's sufficient to set $[eE] = \mathrm{M/L}$ and $[eB] = \mathrm{M/L}$.

7-dimensional dynamics

The phase space of system (4.1) is a 6-dimensional space of coordinates $(\boldsymbol{x}, \boldsymbol{w})$, however, usually in special relativity one considers a 7-dimensional space of coordinates $\{t, \boldsymbol{x}, \boldsymbol{w}\}$, where the trajectory is parametrized by a scalar variable called *proper time* and denoted here by s. A 7th equation is then added to the system,

$$\frac{dt}{ds} = \gamma$$

while in the other equations we replace $d/dt = \gamma^{-1} d/ds$. The extended system is chosen to underline the geometric structure underlying special relativity: t and $\boldsymbol{x}$ are the coordinates of a four-dimensional manifold, called the spacetime, which is endowed with a metric structure:

$$\eta = dt \otimes dt - d\overline{\boldsymbol{x}} \otimes d\boldsymbol{x}$$

The proper time s has the physical meaning of time coordinate of a reference frame in which the particle is seen at rest (hence the name: "proper time").

Electromagnetic Potentials and Gauge Invariance

It is known [34] that an electromagnetic field can be charaterized by a vector potential $\boldsymbol{A}(t, \boldsymbol{x})$ and a scalar potential $\Phi(t, \boldsymbol{x})$ such that

$$\boldsymbol{B} = \left[\!\left[\partial_{\overline{\boldsymbol{x}}}\right]\!\right]\boldsymbol{A}, \qquad \boldsymbol{E} = -\partial_t \boldsymbol{A} - \partial_{\overline{\boldsymbol{x}}}\Phi \tag{4.2}$$

Their definition is not unique, as we have the freedom to choose a scalar function $\boldsymbol{g}$ and to redefine

$$\hat{\boldsymbol{A}} = \boldsymbol{A} + \partial_{\overline{\boldsymbol{x}}} g, \qquad \hat{\Phi} = \Phi - \partial_t g \tag{4.3}$$

If we compute the fields from the hatted potentials, we get the same result as from the original ones. This is called the "gauge freedom", and the function g is called the *gauge function*. The gauge freedom is considered somehow "unphysical" because it is not possible

to detect the gauge in a physical experiment[1] (in physical experiments we measure fields and not potentials), so there's not a way to define it.

Poisson structure, static case

It is possible to introduce a Poisson structure for the system (4.1). We start by considering a static situation, with no electric field. In this case the phase space is

$$\mathbb{V} = \mathcal{C}^\infty(\{\overline{\boldsymbol{x}}, \overline{\boldsymbol{w}}\}) \tag{4.4}$$

We observe that in this case energy is conserved.

As the phase space has a differentiable structure, the Lie product can be defined by the aid of the hodge differential d and of a bivector π, in the following way

$$\{v\}\, w \stackrel{\text{def}}{=} dw(dv\,\pi), \qquad v, w \in \mathcal{C}^\infty(\{\overline{\boldsymbol{x}}, \overline{\boldsymbol{w}}\})$$

We call π the *Poisson bivector* (see also appendix C.1). This bracket is anticommutative by definition. The Jacobi identity is implied either by

$$\{\pi, \pi\}_M = 0$$

($\{,\}_M$ being the Schouten-Nijenhuis bracket [48]), or if π is the inverse of a symplectic form (a closed non-degenerate 2-form, see again appendix C.1).

For the static charged particle we consider the bivector

$$\pi \stackrel{\text{def}}{=} \partial_{\boldsymbol{w}} \wedge \partial_{\overline{\boldsymbol{x}}} + e\,\partial_{\boldsymbol{w}} [\![\boldsymbol{B}]\!]\, \partial_{\overline{\boldsymbol{w}}} \tag{4.5}$$

which is the inverse of the symplectic form

$$\sigma_B = d\overline{\boldsymbol{w}} \wedge d\boldsymbol{x} - e\, d\overline{\boldsymbol{x}}\, [\![\boldsymbol{B}]\!]\, d\boldsymbol{x} \tag{4.6}$$

in the sense that

$$\sigma_B\,(d\overline{\boldsymbol{w}}\,\pi) = d\overline{\boldsymbol{w}}\,, \quad \sigma_B\, d\overline{\boldsymbol{x}}\,\pi = d\overline{\boldsymbol{x}}$$

In our notation, the rules for annihilating forms and vectors read

$$d\boldsymbol{w}\,\partial_{\boldsymbol{w}} = \mathbb{I} = d\boldsymbol{x}\,\partial_{\boldsymbol{x}}$$

The first is readily checked

$$d\boldsymbol{x}\pi = -\partial_{\overline{\boldsymbol{w}}} \implies d\overline{\boldsymbol{x}}\pi = -\partial_{\boldsymbol{w}} \implies \sigma_B(d\overline{\boldsymbol{x}}\pi) = d\overline{\boldsymbol{x}}$$

Then the second

$$d\boldsymbol{w}\,\pi = \partial_{\boldsymbol{x}} + e\,[\![\boldsymbol{B}]\!]\partial_{\overline{\boldsymbol{w}}} \implies d\overline{\boldsymbol{w}}\pi = \partial_{\boldsymbol{x}} - e\,\partial_{\boldsymbol{w}}\,[\![\boldsymbol{B}]\!]$$

so

$$\sigma_B\,(d\overline{\boldsymbol{w}}\,\pi) = d\overline{\boldsymbol{w}} - e\, d\overline{\boldsymbol{x}}\,[\![\boldsymbol{B}]\!] + b\, d\overline{\boldsymbol{x}}\,[\![\boldsymbol{B}]\!] = d\overline{\boldsymbol{w}}$$

Finally, the closedness of σ is equivalent to the homogeneous Maxwell equation $\operatorname{div}\boldsymbol{B} = 0$,

$$d\sigma_B = ed\Big(\overline{\boldsymbol{B}}\,[\![d\boldsymbol{x}]\!]\,d\boldsymbol{x}\Big) = e\,\underbrace{\partial_{\boldsymbol{x}}\boldsymbol{B}}_{\equiv \operatorname{div}\boldsymbol{B}}\,\underbrace{d\overline{\boldsymbol{x}}\,[\![d\boldsymbol{x}]\!]\,d\boldsymbol{x}}_{\text{volume form on}\mathbb{R}^3}$$

[1] Not in classical physics

The other homogeneous Maxwell equation reads $[\![\partial_{\overline{\boldsymbol{x}}}]\!]\boldsymbol{E} = 0$ and it implies $\boldsymbol{E} = -\partial_{\overline{\boldsymbol{x}}}\Phi(\overline{\boldsymbol{x}})$ (this is of course a particular case of (4.2)). Now we consider as Hamiltonian the *canonical energy*

$$\varepsilon = m\gamma + e\Phi \tag{4.7}$$

so the Hamiltonian vector field reads

$$\mathcal{H} = d\varepsilon\pi = \frac{\overline{\boldsymbol{w}}d\boldsymbol{w}}{m\gamma_v}\pi = \frac{\overline{\boldsymbol{w}}}{m\,\gamma_v}\partial_{\overline{\boldsymbol{x}}} + \frac{e\,\overline{\boldsymbol{w}}[\![\boldsymbol{B}]\!]}{m\,\gamma_v}\partial_{\overline{\boldsymbol{w}}} + e\,\overline{\boldsymbol{E}}\,\partial\overline{\boldsymbol{w}}$$

Then the system

$$\begin{cases} \overline{\dot{\boldsymbol{w}}} = \mathcal{H}\overline{\boldsymbol{w}} \\ \overline{\dot{\boldsymbol{x}}} = \mathcal{H}\overline{\boldsymbol{x}} \end{cases}$$

is equivalent to (4.1) with $\partial_t\boldsymbol{E} = 0$ and $\partial_t\boldsymbol{B} = 0$.

Poisson structure, general case

If the fields are no more static, the inhomogeneous Maxwell equations read,

$$\partial_{\boldsymbol{x}}\boldsymbol{B} = 0\,, \qquad [\![\partial_{\overline{\boldsymbol{x}}}]\!]\boldsymbol{E} + \partial_t\boldsymbol{B} = 0 \tag{4.8}$$

but only the first of them is satisfied as a consequence of the Jacobi identity of π; the second equation does not hold in this picture. The correct dynamical system may be recovered by the Hamiltonian vector field

$$\mathcal{H} = \big(d(m\gamma) - e\overline{\boldsymbol{E}}d\boldsymbol{x}\big)\pi$$

but the inhomogeneous Maxwell equations wouldn't be implied.

Another possibility, that was considered for instance in [72], is to autonomize the system (4.1), by extending the phase space with two dimensions, $\{t, P_t\}$ where P_t is the momentum conjugate to t. The symplectic form is also extended to

$$\sigma_{EB} = d\overline{\boldsymbol{w}} \wedge d\boldsymbol{x} + dP_t \wedge dt - e\,d\overline{\boldsymbol{x}}[\![\boldsymbol{B}]\!]d\boldsymbol{x} + e\,d\overline{\boldsymbol{x}}\boldsymbol{E} \wedge dt$$

So time has become a coordinate and evolution is parametrized by a new parameter τ. Now the closedness of σ_{EB} implies both Maxwell equations:

$$d\sigma_{EB} = -ed\overline{\boldsymbol{x}}[\![\partial_t\boldsymbol{B}]\!]d\boldsymbol{x} \wedge dt - e\partial_{\boldsymbol{x}}\boldsymbol{B}\,d\overline{\boldsymbol{x}}[\![d\boldsymbol{x}]\!]d\boldsymbol{x} + ed\overline{\boldsymbol{x}}[\![[\![\partial_{\overline{\boldsymbol{x}}}]\!]\boldsymbol{E}]\!]d\boldsymbol{x} \wedge dt \tag{4.9}$$

$$d\sigma_{EB} \overset{!}{=} 0 \implies \big\{ -\partial_t\boldsymbol{B} + [\![\partial_{\overline{\boldsymbol{x}}}]\!]\boldsymbol{E} = 0\partial_{myVecx}\boldsymbol{B} = 0 \tag{4.10}$$

By a calculation similar to point **E**, but a bit more tedious, the inverse bivector is found,

$$\pi_e = \partial_{\boldsymbol{w}} \wedge \partial_{\overline{\boldsymbol{x}}} + \partial_{P_t} \wedge dt + e\,\partial_{\boldsymbol{w}}[\![\boldsymbol{B}]\!]\partial_{\overline{\boldsymbol{w}}} + e\,\partial_{P_t} \wedge \overline{\boldsymbol{E}}\partial_{\overline{\boldsymbol{w}}}$$

The new Hamiltonian is

$$H_e = P_t + \sqrt{m^2 + \overline{\boldsymbol{w}}\boldsymbol{w}}$$

We compute the Hamiltonian vector field,

$$\mathcal{H}_e = \frac{\overline{\boldsymbol{w}}d\boldsymbol{w}}{m\gamma} + \partial_t + \frac{1}{m\gamma}\overline{\boldsymbol{w}}[\![\boldsymbol{B}]\!]\partial_{\overline{\boldsymbol{w}}} + e\overline{\boldsymbol{E}}\partial_{\overline{\boldsymbol{w}}} - \frac{e\overline{\boldsymbol{E}}\boldsymbol{w}}{m\gamma}\partial_{P_t}$$

which is equivalent to the equations of motion

$$\begin{aligned}
\frac{dt}{d\tau} &= 1 \\
\frac{d\boldsymbol{x}}{d\tau} &= \frac{\boldsymbol{w}}{m\gamma} \\
\frac{d\boldsymbol{w}}{d\tau} &= e\boldsymbol{E} + \frac{e}{m\gamma}[\![\boldsymbol{B}]\!]\boldsymbol{w} \\
\frac{dP_t}{d\tau} &= -\frac{e\overline{\boldsymbol{E}}\boldsymbol{w}}{m\gamma}
\end{aligned}$$

which is the autonomization of (4.1) (see [60]). We have two new equations: the first one says that the new coordinate we have introduced has the meaning of time, the fourth says instead that P_t, the momentum conjugate to time, has the meaning of energy. Indeed, the fourth equation states the (non-)conservation of energy, and in the autonomous description, where P_t is the Hamiltonian and equals $m\gamma$, is a consequence of the third equation.

4.2 The Eulerian reduction

In this section we try to develop a method to drastically reduce the number of degrees of freedom of system (4.1), from 6 to 3. In fact we search for a vector field $\boldsymbol{W}(\boldsymbol{x},t)$ such that the dynamical system

$$\frac{d\boldsymbol{x}}{dt} = \frac{\boldsymbol{w}}{m\gamma}\bigg|_{\boldsymbol{w}=\boldsymbol{W}(\boldsymbol{x},t)} \tag{4.11}$$

describe the same dynamics of the first equation of system (4.1), while the second equation is identically satisfied by setting $\boldsymbol{w} = \boldsymbol{W}$. It's here that we see an analogy with the Euler description of fluid dynamics.

We can accomplish this reduction by a change of coordinates, from $\boldsymbol{w}$ to a new set of three variables, $\boldsymbol{\alpha}$, without changing the position $\boldsymbol{x}$. The new variables will have a very simple dynamics,

$$\frac{d\boldsymbol{\alpha}}{dt} = 0 \tag{4.12}$$

The vector field $\boldsymbol{W}$ of equation (4.11) has the meaning of diffeomorphism from the old coordinates $\{\overline{\boldsymbol{x}}, \overline{\boldsymbol{w}}\}$ to a new set of coordinates $\{\overline{\boldsymbol{x}}, \overline{\boldsymbol{\alpha}}\}$.

To derive an equation for $\boldsymbol{W}$ we start by computing its time derivative

$$\frac{d}{dt}\overline{\boldsymbol{W}} = \partial_t\overline{\boldsymbol{W}} + \frac{d\overline{\boldsymbol{x}}}{dt}\partial_{\overline{\boldsymbol{x}}}\overline{\boldsymbol{W}} + \frac{d\overline{\boldsymbol{\alpha}}}{dt}\partial_{\overline{\boldsymbol{\alpha}}}\overline{\boldsymbol{W}} \overset{!}{=} \frac{d\overline{\boldsymbol{w}}}{dt}$$

By injecting in the above equations (4.1) and (4.12), we find

$$\partial_t\boldsymbol{W} + \Big(\frac{1}{m\gamma}\overline{\boldsymbol{W}}\partial_{\overline{\boldsymbol{x}}}\Big)\boldsymbol{W} = e\boldsymbol{E} - \frac{e}{m\gamma}[\![\boldsymbol{B}]\!]\,\boldsymbol{W} \tag{4.13}$$

A very simple (but also common) case is $\boldsymbol{E} = 0$ and static $\boldsymbol{B}$; in this case γ is constant, and equation (4.13) becomes simply

$$(\overline{\boldsymbol{W}}\nabla)\boldsymbol{W} = -e\,[\![\boldsymbol{B}]\!]\,\boldsymbol{W} \tag{4.14}$$

A question remains about the physical meaning of the variables $\boldsymbol{\alpha}$. They should be related to the initial data of the problem. In fact, in the original problem (4.1) 6 initial data are given; if the Eulerian Reduced equations are assumed to be equivalent to it, we have to use all the 6 initial data, and the three of them referred to the velocity can enter the problem only as values of the coordinates $\boldsymbol{\alpha}$. For instance, by imposing some constraint of the following type,

$$\boldsymbol{w}(\boldsymbol{x}_0, \boldsymbol{\alpha}) \,=\, \boldsymbol{\alpha} \tag{4.15}$$

In this case the new coordinates are strictly identified with the initial data for the velocity (actually, for the momentum). We may more generally choose them as combinations of the initial velocities, depending also on the problem at hand. For instance, in the absence of an electric field the modulus of $\boldsymbol{W}$ is constant and can be identified with one of them. Indeed we will make this choice in section 4.2.5.

An alternative look for equation (4.13)

We use the vectorial identity

$$[\![\boldsymbol{W}]\!]\,[\![\partial_{\overline{\boldsymbol{x}}}]\!]\,\boldsymbol{W} \,=\, \underbrace{\partial_{\overline{\boldsymbol{x}}}\overline{\boldsymbol{W}}\boldsymbol{W}}_{\equiv\frac{1}{2}\partial_{\overline{\boldsymbol{x}}}(h^2-1)} \,-(\overline{\boldsymbol{W}}\partial_{\overline{\boldsymbol{x}}})\boldsymbol{W} \implies$$

$$\Big(\frac{1}{m\gamma}\overline{\boldsymbol{W}}\partial_{\overline{\boldsymbol{x}}}\Big)\boldsymbol{W} \,=\, \frac{1}{2h}\partial_{\overline{\boldsymbol{x}}}(h^2-1) \,+\, \frac{1}{m\gamma}[\![\boldsymbol{W}]\!]\,[\![\partial_{\overline{\boldsymbol{x}}}]\!]\,\boldsymbol{W}$$

to rewrite equation (4.13) as

$$-\partial_t\boldsymbol{W} \,-\partial_{\overline{\boldsymbol{x}}}h + e\boldsymbol{E} \,+\, \frac{1}{m\gamma}\,[\![\boldsymbol{W}]\!]\Big(e\boldsymbol{B} \,+\, [\![\boldsymbol{W}]\!]\Big)$$

Upon defining

$$e\boldsymbol{E}_c \,=\, e\boldsymbol{E} \,-\, \partial_{\overline{\boldsymbol{x}}}h \,-\, \partial_t\boldsymbol{W}\,, \qquad e\boldsymbol{B}_c \,=\, e\boldsymbol{B} \,+\, [\![\partial_{\overline{\boldsymbol{x}}}]\!]\,\boldsymbol{W} \tag{4.16}$$

we find a new form for equation (4.13),

$$\boldsymbol{E}_c \,+\, [\![\,\frac{\boldsymbol{W}}{m\gamma}\,]\!]\,\boldsymbol{B}_c \,=\, 0 \tag{4.17}$$

The fields defined in (4.16) are called "canonical" because they are defined by using the energy and momentum of the particle as scalar and vector potential for the electromagnetic field. Equation (4.17), which was written for the first time in [22], will be useful to trace a parallelism with Guiding Centre theory, as we will do in section 4.3.2.

The Jacobian

Finally, we recall that the Jacobian J of a change of variable is defined as the proportionality factor between the old and new volume form on the phase space:

$$dw_1 \wedge dw_2 \wedge dw_3 \wedge dx_1 \wedge dx_2 \wedge dx_3 \,=\, J d\alpha_1 \wedge d\alpha_2 \wedge d\alpha_3 \wedge dx_1 \wedge dx_2 \wedge dx_3$$

and for our transformation $\{\overline{\boldsymbol{x}}, \overline{\boldsymbol{w}}\} \mapsto \{\overline{\boldsymbol{x}}, \overline{\boldsymbol{\alpha}}\}$ we find

$$J \,=\, \partial_{\alpha_1}\overline{\boldsymbol{W}}\,[\![\partial_{\alpha_2}\boldsymbol{W}]\!]\,\partial_{\alpha_3}\boldsymbol{W} \tag{4.18}$$

4.2.1 The Guiding Particle Solution

If we consider equation (4.17) and set $\boldsymbol{E_c} = 0$, we are left with

$$[\![\boldsymbol{W}]\!]\,(e\boldsymbol{B_c}) = [\![\boldsymbol{W}]\!]\,(e\boldsymbol{B} \,+\, [\![\partial_{\overline{\boldsymbol{x}}}]\!]\,\boldsymbol{W}) \,=\, 0$$

which is equivalent to

$$\frac{1}{\lambda}\,\boldsymbol{W} \,=\, e\boldsymbol{B} \,+\, [\![\partial_{\overline{\boldsymbol{x}}}]\!]\,\boldsymbol{W} \tag{4.19}$$

Here λ is some scalar function. The solution of this equation in the case $\left|\,[\![\partial_{\overline{\boldsymbol{x}}}]\!]\,\boldsymbol{W}\right| \ll |\boldsymbol{W}|$ was called the "Guiding Particle" solution in [22]. To have an idea of the physical meaning of the equation above, we search for a solution iteratively. First we assume that the curl of $\boldsymbol{W}$ is completely negligible. We get an approximate solution as

$$\boldsymbol{W}^{(0)} \,=\, \lambda e B\boldsymbol{b}$$

where we put $\boldsymbol{B} = B\boldsymbol{b}$ and $\boldsymbol{b}$ has unit norm. The approximate solution is then used to find a better approximation:

$$\boldsymbol{W}^{(1)} \,=\, \lambda e B\boldsymbol{b} \,+\, \lambda[\![\partial_{\overline{\boldsymbol{x}}}]\!]\,\boldsymbol{W}^{(0)} \tag{4.20}$$

If we set $w_{\|} \stackrel{\text{def}}{=} \overline{\boldsymbol{b}}\boldsymbol{W}$ and then $\lambda \,=\, w_{\|}/(eB_0)$, equation (4.20) turns into (4.21),

$$\boldsymbol{W}^{(1)} \,=\, w_{\|}\boldsymbol{b} \,+\, \frac{w_{\|}}{eB_0}\,[\![\partial_{\overline{\boldsymbol{x}}}]\!]\,(w_{\|}\boldsymbol{b}) \tag{4.21}$$

This formula is interesting in that it resembles the standard guiding centre velocity in the case of null magnetic moment [15].

4.2.2 The Poisson structure after the Eulerian Reduction

Now we look for the Poisson structure in the coordinates $\{\overline{\boldsymbol{x}}, \overline{\boldsymbol{\alpha}}\}$. We will consider as a starting point the static structure (4.5), as it is not clear to us if all of the three variables $\overline{\boldsymbol{\alpha}}$ can be constant in the non-static case.

A very efficient method to compute the new Poisson bivector, is to perform the change of variables into σ_B, and then invert the resulting new symplectic 2-form. In general we are proposing the following logical scheme to invert a bivector:

$$\pi_{\text{old}} \stackrel{\text{inversion}}{\longrightarrow} \sigma_{\text{old}} \stackrel{\text{change of variables}}{\longrightarrow} \sigma_{\text{new}} \stackrel{\text{inversion}}{\longrightarrow} \pi_{\text{new}}$$

The reason to choose this path is that a change of variables for forms is well defined, while for derivations it is not the case[2].

To compute the new symplectic 2-form, we start by computing

$$\boldsymbol{w} \,\mapsto\, \boldsymbol{W}(\boldsymbol{x}, \boldsymbol{\alpha}) \;\Longrightarrow\; d\overline{\boldsymbol{w}} \,\mapsto\, d\overline{\boldsymbol{W}} \,=\, d\overline{\boldsymbol{\alpha}}\,\mathsf{D} \,+\, d\overline{\boldsymbol{x}}\,\partial_{\overline{\boldsymbol{x}}}\,\overline{\boldsymbol{W}}$$

where we called $\mathsf{D} = \partial_{\overline{\boldsymbol{\alpha}}}\,\overline{\boldsymbol{W}}$. By plugging the above into σ_B we get

$$\begin{aligned}\sigma_* \,&=\, d\overline{\boldsymbol{\alpha}}\,\mathsf{D} \,\wedge\, d\boldsymbol{x} \,+\, d\overline{\boldsymbol{x}}\,\partial_{\boldsymbol{x}}\overline{\boldsymbol{W}} \,\wedge\, d\boldsymbol{x} \,-\, e\,d\overline{\boldsymbol{x}}\,[\![\boldsymbol{B}]\!]\,d\boldsymbol{x}\\ &=\, d\overline{\boldsymbol{\alpha}}\,\mathsf{D} \,\wedge\, d\boldsymbol{x} \,-\, d\overline{\boldsymbol{x}}\,[\![\,[\![\partial_{\overline{\boldsymbol{x}}}]\!]\,\boldsymbol{W} + e\boldsymbol{B}]\!]\,d\boldsymbol{x}\end{aligned}$$

[2] A choice of connection would be needed.

To find the inverse bivector, we pose an ansatz:

$$\pi_* \,=\, \partial_{\boldsymbol{\alpha}} \,\wedge\, \mathsf{M}\,\partial_{\overline{\boldsymbol{x}}} \,+\, e\,\partial_{\boldsymbol{\alpha}}\,\mathcal{B}_*\,\partial_{\overline{\boldsymbol{\alpha}}}$$

with unknown matrices M and $\mathcal{B}_*$, the latter being skew-symmetric by hypothesis. Then we impose the relations

$$\sigma_*\, d\boldsymbol{\alpha}\,\pi_* \,=\, d\boldsymbol{\alpha} \quad \text{and} \quad \sigma_*\, d\boldsymbol{x}\,\pi_* \,=\, d\boldsymbol{x} \tag{4.22}$$

So we compute

$$\begin{aligned} d\boldsymbol{\alpha}\,\pi_* \,&=\, \mathsf{M}\,\partial_{\overline{\boldsymbol{x}}} \,+\, e\,\mathcal{B}_*\,\partial_{\overline{\boldsymbol{\alpha}}} \;\Longrightarrow\; d\overline{\boldsymbol{\alpha}}\,\pi_* \,=\, \partial_{\boldsymbol{x}}\,\mathsf{M} \,-\, e\,\partial_{\boldsymbol{\alpha}}\,\mathcal{B}_* \;\Longrightarrow \\ \sigma_*\, d\overline{\boldsymbol{\alpha}}\,\pi_* \,&=\, d\overline{\boldsymbol{x}}\,\overline{\mathsf{D}}\,e\,\mathcal{B}_* \,-\, d\overline{\boldsymbol{x}}\,\big[\!\big[\big[\!\big[\partial_{\overline{\boldsymbol{x}}}\big]\!\big]\boldsymbol{W} + e\boldsymbol{B}\big]\!\big]\,\overline{M} \,+\, d\overline{\boldsymbol{\alpha}}\,\mathsf{D}\,\overline{M} \end{aligned}$$

and $\sigma_*\, d\overline{\boldsymbol{x}}\,\pi_* \,=\, d\overline{\boldsymbol{x}}\overline{\mathcal{A}}\,\mathsf{M}$. We see that the equations (4.22) are satisfied iff

$$\begin{cases} \mathsf{D}\,\overline{\mathsf{M}} \,=\, 1 \\ e\,\overline{\mathsf{D}}\,\mathcal{B}_* \,=\, \big[\!\big[\big[\!\big[\partial_{\overline{\boldsymbol{x}}}\big]\!\big]\boldsymbol{W} + e\boldsymbol{B}\big]\!\big]\,\overline{\mathsf{M}} \\ \overline{\mathsf{M}}\,\mathsf{D} \,=\, 1 \end{cases}$$

If $\det \mathsf{D} \neq 0$, so that D is invertible, the system is solved by $\overline{\mathsf{M}} \,=\, \mathsf{D}^{-1}$ and

$$e\,\mathcal{B}_* \,=\, \overline{\mathsf{D}}^{-1}\big[\!\big[\big[\!\big[\partial_{\overline{\boldsymbol{x}}}\big]\!\big]\boldsymbol{W} + e\boldsymbol{B}\big]\!\big]\,\mathsf{D}^{-1} \tag{4.23}$$

We see that imposing $\mathcal{B}_* = 0$ is equivalent to ask

$$\big[\!\big[\partial_{\overline{\boldsymbol{x}}}\big]\!\big]\boldsymbol{W} + e\boldsymbol{B} = 0 \tag{4.24}$$

If such condition is realized, the new coordinates $\boldsymbol{\alpha}$ resemble Hamiltonian actions, as they all commute among themselves (but they are not yet conjugated to to an angle). We observe that (4.24) corresponds to the limit $\lambda \to \infty$ of equation (4.19), so it gives one very particular solution for the Eulerian velocity.

If we introduce a vector potential $\boldsymbol{A}$, such that $\boldsymbol{B} \,=\, \big[\!\big[\partial_{\overline{\boldsymbol{x}}}\big]\!\big]\boldsymbol{A}$, then equation (4.24) becomes

$$\boldsymbol{W} \,=\, -e\,\boldsymbol{A} \,+\, \partial_{\overline{\boldsymbol{x}}} g \tag{4.25}$$

where g is some gauge function. This kind of solution was already considered in [23] where it is also noted the following: in classical mechanics the canonical momentum is defined by

$$\boldsymbol{P} \stackrel{\text{def}}{=} \boldsymbol{W} + e\boldsymbol{A} \equiv \partial_{\overline{\boldsymbol{x}}} g$$

In the Hamilton-Jacobi solution of the dynamics, the canonical momentum is equal to the spatial gradient of Hamilton's principal function (usually denoted by S). So, in this context the magnetic gauge coincides with Hamilton's principal function[3]. And the solution of equation (4.25) was called the gyrating particle, or simply *gyroparticle*, if the trajectory is closed and periodic. In that case, the origin of this coordinate system is called the Guiding Centre, according to definition 1.

[3]There exists in fact a general relation between gauge trasforms and canonical transforms, which is evident in the lagrangian description. This is discussed in appendix C.

4.2.3 Example: a constant uniform magnetic field

We consider a costant uniform magnetic field in cartesian coordinates, $\boldsymbol{B} = B_0\boldsymbol{e}_z$. The solution of equations (4.1) can be computed explicitly:

$$\begin{cases} \boldsymbol{w}(t) \;=\; \exp\Big(-\dfrac{\omega t}{\gamma}[\![\boldsymbol{e}_z]\!]\Big)(-[\![\boldsymbol{e}_z]\!]^2)\boldsymbol{w_0} \;+\; w_0^z\boldsymbol{e}_z \\ \boldsymbol{x}(t) \;=\; \dfrac{1}{eB_0}[\![\boldsymbol{e}_z]\!]\exp\Big(-\dfrac{\omega t}{\gamma}[\![\boldsymbol{e}_z]\!]\Big)\boldsymbol{w_0} \;+\; \dfrac{w_z^0\,t}{m\gamma}\boldsymbol{e}_z \;+\; X\boldsymbol{e}_x \;+\; Y\boldsymbol{e}_y \end{cases} \tag{4.26}$$

(see appendix A.3 for the exponential of a skew-symmetric matrix). The second equation is the parametrization of a helix. The particle is turning with constant frequency $\omega/\gamma = eB_0/(m\gamma)$ (ω is called the *Larmor frequency*) around a magnetic field line of coordinates $x = X, y = Y$. The particle is also traslating along the field line with constant velocity $w_z^0/m\gamma \equiv v_z^0$. By evaluating at $t = 0$ the second equation (4.26) we compute

$$X \;=\; x^0 \;-\; \frac{w_y^0}{e\,B_0}\,, \quad Y \;=\; y^0 \;+\; \frac{w_x^0}{e\,B_0}$$

The angle between the velocity vector of the particle and the magnetic axis, defined by $\tan\phi \;=\; (v_z^0)^2/\sqrt{(v_x^0)^2 + (v_y^0)^2}$, is called the *pitch angle*, while the angle $\zeta \;=\; \omega t/\gamma$ is called the *gyrophase*. The second equation of (4.26) can be inverted by multiplying on the left by $[\![\boldsymbol{e}_z]\!]$,

$$-\exp\Big(-\frac{\omega t}{\gamma}[\![\boldsymbol{e}_z]\!]\Big)[\![\boldsymbol{e}_z]\!]^2\boldsymbol{w}^0 \;=\; eB_0[\![\boldsymbol{e}_z]\!](\boldsymbol{x} - \boldsymbol{X})$$

where we defined $\boldsymbol{X} \stackrel{\text{def}}{=} X\boldsymbol{e}_x + Y\boldsymbol{e}_y$. and it can be plugged into the former equation to give

$$\boldsymbol{w} \;=\; eB_0[\![\boldsymbol{e}_z]\!](\boldsymbol{x} - \boldsymbol{X}) \;+\; w_z^0\boldsymbol{e}_z \tag{4.27}$$

The quantity $\boldsymbol{\rho} \stackrel{\text{def}}{=} -[\![\boldsymbol{e}_z]\!]^2(\boldsymbol{x}(t) \;-\; \boldsymbol{X})$ is constant and is called the *Larmor radius*. We have $\overline{\boldsymbol{w}}\boldsymbol{w} \;=\; (eB_0)^2\rho^2 \;+\; (w_z^0)^2$ by which the single particle energy can be rewritten as

$$h \;=\; m\sqrt{1 \;+\; \left(\tfrac{w_z^0}{m}\right)^2 \;+\; \omega^2\rho^2}$$

Formula (4.27) has the meaning of moment as a function the position, or moment field, and indeed is a solution of (4.14). If $\boldsymbol{x^0}$ is the initial position, we define

$$\boldsymbol{\alpha} \;=\; eB_0\,[\![\boldsymbol{e}_z]\!](\boldsymbol{x^0} - \boldsymbol{X}) + w_z^0\boldsymbol{e}_z$$

to rewrite equation (4.27) as

$$\boldsymbol{W} \;=\; \boldsymbol{\alpha} \;-\; eB_0\,[\![\boldsymbol{e}_z]\!](\boldsymbol{x} - \boldsymbol{x^0}) \tag{4.28}$$

so that $\boldsymbol{\alpha}$ is the initial velocity: $\boldsymbol{\alpha} \;=\; \boldsymbol{W}(\boldsymbol{x^0}, \boldsymbol{\alpha})$.

In figure 4.2 we plot the result of simple numerical simulations of the trajectory given by direct integration of (4.1), and that given by the GC solution (4.28). As one can see, there's perfect agreement between the two, i.e., between the full dynamics and its "Eulerian reduction".

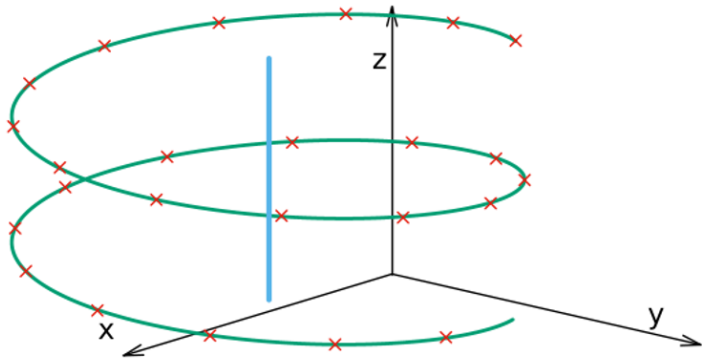


Figure 4.1: *The trajectory of a charged particle in a constant uniform magnetic field field in cartesian coordinates* (x, y, z)*. We solved numerically* (4.1) *with* $B = B_0 \boldsymbol{e}_z$*, for a total of 6 degrees of fredom. We also solved* $\frac{d}{dt}\boldsymbol{x} = \boldsymbol{W}(\boldsymbol{x}, \boldsymbol{\alpha})/(m\gamma)$*, with* $\boldsymbol{W}$ *given by* (4.28)*, and the result are the red crosses. As one can see, there's perfect agreeement with the two solutions (but in the second case we had to solve only 3 equations instead of 6). We used a Runge-Kutta 4th order scheme, with a step of* $h = 0.001$ *for both simulations. The initial data are* $\boldsymbol{x}^0 = (1, 0, 0)$*,* $\boldsymbol{\alpha} = (0.6, 1.1, 0.3) \equiv \boldsymbol{w}^0$ *and* $\omega = 1$*. The blue line is the magnetic field line of coordinates* $x = X, y = Y$ *around which the particle is turning, and it coincides also with the GC trajectory (see section 4.3).*

4.2.4 Another example: the toroidal field

We call "toroidal" a magnetic field of type $\boldsymbol{B} = B_0 \boldsymbol{e}_\tau L/R$, where R, τ, z are the cylindrical coordinates (τ is called *toroidal angle* in fusion context) and L is a constant with dimentions of a length. Here B_0 is a constant measuring the field strength. This magnetic field is produced, for instance, by a thin wire along the z-axis[4]

One can check that a solution of (4.14) for the toroidal magnetic field is

$$\boldsymbol{W} = B_0 \left(\frac{\alpha_1 \, \alpha_2}{R} \boldsymbol{e}_\tau + \log\left(\frac{R}{\alpha_2}\right) \boldsymbol{e}_z \pm \sqrt{\alpha_3^2 + \alpha_1^2\left(1 - \frac{\alpha_2}{R}\right) - \log\left(\frac{R}{\alpha_2}\right)} \, \boldsymbol{e}_R \right) \tag{4.29}$$

A particle following with this velocity is turning around a point which is itself turning around the z-axis, and drifting along it. We may also describe this trajectory as an helix, winding around another helix, winding around the z-axis. The meaning of the new coordinates $\overline{\boldsymbol{\alpha}}$ in this second example is as follows: $R = \alpha_2$ is the radial coordinate of the centre of gyration, while α_1 and α_3 are related to the extrema of this circle. In fact, the $R-$component of the field $\boldsymbol{W}$ is equal to zero in the two points $r_\pm$ (we may call them

[4]By a little perturbation of the "toroidal field", we get a common model for magnetic fields in tokamaks:

$$\boldsymbol{B} = B_0 \boldsymbol{e}_\tau / R + [\![\nabla f(R, \tau.z)]\!] \boldsymbol{e}_\tau$$

A field of this type is called axysimmetric, because it is invariant for rotation around the z-axis.

"turning points") defined by

$$\log(r_\pm) = \pm\sqrt{\alpha_3^2 + \alpha_1^2(1 - r_\pm^{-2})}$$

The factor $\pm$ in front of the $R-$component of $\boldsymbol{W}$ arises because the field is singular at the inversion points. In fact the Jacobian is

$$J \;=\; \frac{M\alpha_3}{R\sqrt{\alpha_3^2 + \alpha_1^2\Big(1 - \frac{\alpha_2}{R}\Big) - \log\Big(\frac{R}{\alpha_2}\Big)}}$$

so it is singular at the turning points: the solution (4.29) is not well defined over the whole space.

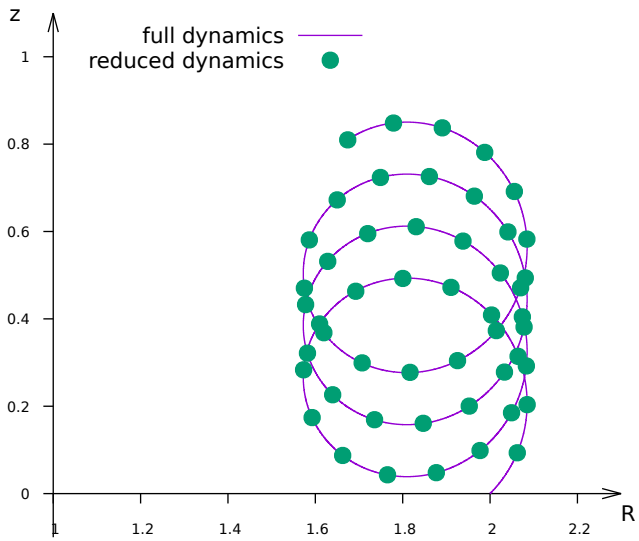


Figure 4.2: *Here we plot the trajectory of a charged particle in a "toroidal field"* $\boldsymbol{B} = B_0\boldsymbol{e}_\tau/R$ *in cylindrical coordinates* (R, τ, z)*. More precisely we are visualizing the* (R, z) *projection of motion. The continuous line is the numerical solution of* (4.1) *for a total of 6 degrees of fredom. The points instead come from the solution of* $\frac{d}{dt}\boldsymbol{x} = \boldsymbol{W}(\boldsymbol{x}, \boldsymbol{\alpha})/(m\gamma)$*, with* $\boldsymbol{W}$ *given by* (4.28)*, (we are plotting just one point every 100). The sign of the* $R-$*component of the field was changed as an external input each time there was a change in the sign We see perfect agreeement among the two solutions. We used a Runge-Kutta 4th order scheme, with a step of* $h = 0.001$ *for both simulations. The initial data are* $\boldsymbol{x}^0 = (1, 0, 0)$*,* $\boldsymbol{\alpha} = (0.6, 1.1, 0.3) \equiv \boldsymbol{w}^0$ *and* $eB_0 = 1$*.*

4.2.5 Euler angles decomposition

Here we discuss one method by which we tried to solve (4.14). Spoiler alert: this approach is still work in progress, and we have not come to a point yet, so the reader may wish to skip this section.

To get a natural generalization of the constant uniform field case (section 4.2.3), the pitch angle φ and gyrophase ζ can be interpreted as Euler angles; this leads to the following ansatz,

$$\boldsymbol{W} \;=\; \alpha_3\Big(\boldsymbol{b}\cos\varphi \;+\; \sin\varphi\big(\cos\zeta + \sin\zeta\,[\![\boldsymbol{b}]\!]\big)\boldsymbol{u}\Big) \tag{4.30}$$

Here $\boldsymbol{b}$ is again the unit norm along the magnetic field, while $\boldsymbol{u}$ is a unit vector in the plane orthogoanal to $\boldsymbol{b}$. As the modulus of $\boldsymbol{W}$ is constant, we identify it with with α_3 (we recall that the $\overline{\boldsymbol{\alpha}}$ are constants, which values are determined by any combination of the initial velocities).

In this description the pitch angle and the gyrophase cannot be independent variables (the only independent variables are $\overline{\boldsymbol{x}}$); rather, they are fields:

$$\varphi \,=\, \varphi(\boldsymbol{x},\boldsymbol{\alpha}) \qquad \zeta \,=\, \zeta(\boldsymbol{x},\boldsymbol{\alpha})$$

We choose to define

$$\alpha_1 \,=\, \varphi(\boldsymbol{x}|_{t=0},\boldsymbol{\alpha}) \qquad \alpha_2 \,=\, \zeta(\boldsymbol{x}|_{t=0},\boldsymbol{\alpha})$$

The Jacobian

One can also compute (here $k \in \{1,2,3\}$)

$$\begin{aligned}\partial_{\alpha_k}\boldsymbol{W} \,=\, &\big(\boldsymbol{b}\cos\varphi \,-\, \sin\varphi(\cos\zeta \,+\, \sin\zeta\,[\![\boldsymbol{b}]\!])\boldsymbol{u}\big)\delta_{k,3} + \\ &-\, \alpha_3\big(\boldsymbol{b}\sin\varphi \,+\, \cos\varphi(\cos\zeta \,+\, \sin\zeta\,[\![\boldsymbol{b}]\!])\boldsymbol{u}\big)\,\partial_{\alpha_k}\varphi + \\ &+\, \alpha_3\big(\boldsymbol{b}\cos\varphi \,+\, \sin\varphi(\sin\zeta \,-\, \cos\zeta\,[\![\boldsymbol{b}]\!])\boldsymbol{u}\big)\,\partial_{\alpha_k}\zeta\end{aligned}$$

where $\delta_{i,j}$ is the kronecker delta. Then

$$J(\boldsymbol{x},\boldsymbol{\alpha}) \,=\, \alpha_3^2\sin\varphi\big(\partial_{\alpha_1}\varphi\,\partial_{\alpha_2}\zeta \,-\, \partial_{\alpha_2}\varphi\,\partial_{\alpha_1}\zeta\big)$$

and in particular

$$J(\boldsymbol{x}|_{t=0},\boldsymbol{\alpha}) \,=\, \alpha_3^2\sin\varphi$$

So we need to ask that $\sin\varphi \neq 0$, for the map (4.30) to be well defined.

The new equations of motion

We define $\boldsymbol{w} \stackrel{\text{def}}{=} \boldsymbol{W}/\alpha_3$ and we compute

$$\begin{aligned}\alpha_3^{-2}(\overline{\boldsymbol{W}}\nabla)\boldsymbol{W} \,\equiv\, (\overline{\boldsymbol{w}}\nabla)\boldsymbol{w} \,=\, &(\cos\varphi \,-\, \sin\varphi\sin\zeta\,[\![\boldsymbol{u}]\!])(\overline{\boldsymbol{w}}\nabla)\boldsymbol{b} + \\ &+\, \sin\varphi\,(\cos\zeta \,+\, \sin\zeta\,[\![\boldsymbol{b}]\!])\,(\overline{\boldsymbol{w}}\nabla)\boldsymbol{u} \,+\, \sin\varphi\,(-\sin\zeta \,+\, \cos\zeta\,[\![\boldsymbol{b}]\!])\,\boldsymbol{u}\,(\overline{\boldsymbol{w}}\nabla)\zeta + \\ &+\, \big(-\boldsymbol{b}\sin\varphi \,+\, \cos\varphi(\cos\zeta \,+\, \sin\zeta\,[\![\boldsymbol{b}]\!])\boldsymbol{u}\big)\,(\overline{\boldsymbol{w}}\nabla)\varphi\end{aligned}$$

Now we observe that equation (4.14) lives on the plane $\boldsymbol{W}^{\perp}$, infact

$$\overline{\boldsymbol{W}}(\overline{\boldsymbol{W}}\nabla)\boldsymbol{W} \,=\, 0$$

This is the same as saying that equation (4.14), although being 3-dimensional, is in fact equivalent to two equations; they are found for instance by multiplying equation (4.14) on the left respectively by $\overline{\boldsymbol{b}}$ and by $\overline{\boldsymbol{W}}[\![\boldsymbol{b}]\!]$. The results are

$$\begin{cases}\overline{\boldsymbol{b}}(\overline{\boldsymbol{w}}\nabla)\boldsymbol{w} \,=\, 0 \\ \alpha_3^2\,\overline{\boldsymbol{b}}[\![\boldsymbol{w}]\!]\,(\overline{\boldsymbol{w}}\nabla)\boldsymbol{w} \,=\, \frac{e}{m}B\alpha_3\sin^2\varphi\end{cases}$$

By some vectorial algebra, the upper two equations lead to

$$\begin{cases}\overline{\boldsymbol{w}}\nabla\varphi \,=\, -\cos\zeta\,\overline{\boldsymbol{u}}(\overline{\boldsymbol{w}}\nabla)\boldsymbol{b} \,-\, \sin\zeta\,\overline{\boldsymbol{b}}[\![\boldsymbol{u}]\!]\,(\overline{\boldsymbol{w}}\nabla)\boldsymbol{b} \\ eB_0 \,=\, (\overline{\boldsymbol{w}}\nabla)\zeta \,+\, \overline{\boldsymbol{b}}[\![\boldsymbol{u}]\!]\,(\overline{\boldsymbol{w}}\nabla)\boldsymbol{u} \,+\, \Phi\,(\overline{\boldsymbol{w}}\nabla)\boldsymbol{b}\,(\cos\zeta\,[\![\boldsymbol{b}]\!] \,-\, \sin\zeta)\boldsymbol{u}\end{cases} \tag{4.31}$$

where we defined $\Phi \stackrel{\text{def}}{=} \cot\varphi$ which is well defined for $0 < \varphi < \pi$ as we assumed $\sin\varphi \neq 0$. If we plug $\boldsymbol{w} = \alpha_3^{-1}\boldsymbol{W}$ with $\boldsymbol{W}$ given by (4.30) in system (4.31) we end up with

$$\begin{cases} (\overline{\boldsymbol{w}}\nabla)\varphi = \frac{1}{2}\sin\varphi\Big(\overline{\nabla}\boldsymbol{b} + \sin 2\zeta\,\overline{\boldsymbol{u}}\big([\![\boldsymbol{b}]\!](\overline{\boldsymbol{u}}\nabla)\boldsymbol{b} - (\overline{\boldsymbol{u}}[\![\boldsymbol{b}]\!]\nabla)\boldsymbol{b}\big) + \\ \qquad - \cos 2\zeta\,\overline{\boldsymbol{u}}\big((\overline{\boldsymbol{u}}\nabla)\boldsymbol{b} + (\overline{\boldsymbol{u}}[\![\boldsymbol{b}]\!]\nabla)[\![\boldsymbol{b}]\!]\boldsymbol{b}\big)\Big) - \cos\varphi\,\overline{(\overline{\boldsymbol{b}}\nabla)\boldsymbol{b}}\,e^{\zeta[\![\boldsymbol{b}]\!]}\boldsymbol{u} \\ (\overline{\boldsymbol{w}}\nabla)\zeta = eB_0 + \cos\varphi\,\overline{\boldsymbol{u}}\Big([\![\boldsymbol{b}]\!](\overline{\boldsymbol{b}}\nabla)\boldsymbol{u} + \frac{1}{2}\big([\![\boldsymbol{b}]\!](\overline{\boldsymbol{u}}\nabla) + (\overline{\boldsymbol{u}}[\![\boldsymbol{b}]\!]\nabla)\big)\boldsymbol{b}\Big) + \\ \qquad + \sin\varphi\,\overline{\boldsymbol{u}}[\![\boldsymbol{b}]\!](\overline{\boldsymbol{u}}e^{\zeta[\![\boldsymbol{b}]\!]}\nabla)\boldsymbol{u} + \Phi\cos\varphi\,\overline{(\overline{\boldsymbol{b}}\nabla)\boldsymbol{b}}(\sin\zeta - \cos\zeta[\![\boldsymbol{b}]\!])\boldsymbol{u} + \\ \qquad + \frac{1}{2}\sin 2\zeta\cos\varphi\,\overline{\boldsymbol{u}}\big((\overline{\boldsymbol{u}}\nabla) + [\![\boldsymbol{b}]\!](\overline{\boldsymbol{u}}[\![\boldsymbol{b}]\!]\nabla)\big)\boldsymbol{b} + \\ \qquad + \frac{1}{2}\cos\varphi\cos 2\zeta\,\overline{\boldsymbol{u}}\big([\![\boldsymbol{b}]\!](\overline{\boldsymbol{u}}\nabla) - (\overline{\boldsymbol{u}}[\![\boldsymbol{b}]\!]\nabla)\big)\boldsymbol{b} \end{cases}$$

The new equations of motion are partial differential equations for the functions φ and ζ. Their solution is a formidable task, for which we have not figured out a general method. But we see that we have a great ambition: the vector field $\boldsymbol{W}$ is defined over a 3-dimensional space, but we want to "squeeze" all of the information on the 3 coordinates in only two scalar variables, that is, we are looking for a bijection between a 3 dimensional and a 2 dimensional space.

4.3 Non-perturbative Guiding Centre Theory

As it was seen in section 4.2.3, under the influence of a constant uniform magnetic field, the trajectory of a charged particle is a helix, given by the superposition of two motions: a gyration in the plane orthogonal to the magnetic field, with constant (Larmor) frequency ω; and a translation of the center of gyration along the magnetic field line, with constant velocity. Qualitatively, we observe a kind of helix even when the fields are no more constant and uniform, if the particle is subject to a sufficiently strong magnetic field; and the gyration period is always much smaller than any other time-scale involved.

Guiding Centre Theory is a mathematical theory to make this description quantitative: indeed the name "Guiding Centre" refers to the point drifting along the magnetic field lines, and around which the helix wraps. And it is a reduction theory, because the gyratory motion is decoupled from the drift.

Historically, the two cornerstones of Guiding Centre theory are represented by the works of Northrop [74], [73], [75] of the 1960ies, and those of Littlejohn [53], [54], [55], [56] of the 1980ies; a recent comprehensive survey has been provided by Cary and Brizard [15]. Northrop applied, to the Newton equations with Lorentz force, a method by Kruskal [49] to asymptotically remove fast oscillations from a generical dynamical system. By this method it is possible to compute an adiabatic invariant, also expressed as an asymptotic series. In principle the algorithm is defined up to arbitrary order, however, it requires a fastly increasing amount of algebra[5]. Littlejohn switched to the Lagrangian description of electrodynamics, and used a combination of Lie and gauge trasforms to define the GC coordinates perturbatively. The method is strongly model dependent, as the choice of the orders of magnitude of the various physical quantities has to be assumed a priori and has

[5]the method of Kruskal has been reconsidered by Burby in the recent paper [13] in the light of a variational principle for fast-slow systems.

an influence on the result. Moreover the method is not defined iteratively[6], and it has been pushed up to fourth order in the recent paper by Brizard and Tronko [11]. In this approach there is no control on the convergence of the transform, while some numerical evidencies on the presence of chaos in phase space have been found [14].

Another approach to Guiding Centre theory, based on canonical coordinates was recently proposed by Neishtadt and Artemyiev [72]. The change of coordinates is defined exactly, however, the magnetic moment is still only an adiabatic invariant, not an exactly conserved quantity. We also mention a recent effort by Scott [86] to build the Guiding Centre reduction only by gauge transforms.

In this section we will consider on a novel approach to GC theory, proposed by Di Troia in the two works [22] and [23]. The Guiding Centre coordinates are defined geometrically, not by their expansion in the particles coordinates, and thus closed equations to compute them are provided. To introduce the theory we start by revisiting the exact solution for the constant and uniform magnetic field.

The constant uniform magnetic field, again

The dynamics of a cherged particle in a constant and uniform magnetic field was described at length in section 4.2.3. From equation (4.26), or equivalently from equation (4.27), we deduce that there exists one reference frame in which the particle moves on a closed and periodic trajectory: the frame moving along z with constant velocity $\boldsymbol{V} = v_z^0 \boldsymbol{e}_z$. We call Guiding Centre, and denote by $\boldsymbol{X}$, the origin of this reference frame. It can be shown that in the presence of a uniform and constant electric field $\boldsymbol{E_0}$, the velocity of the center of gyration would be instead

$$\boldsymbol{V} \;=\; v_z^0 \boldsymbol{e}_z \;+\; \frac{[\![\boldsymbol{E_0}]\!]\,\boldsymbol{B_0}}{|\boldsymbol{B_0}|^2} \tag{4.32}$$

The trajectory of the particle is a circle parametrized by an angle ζ which evolves in time according to the law

$$\zeta \;=\; -\frac{\omega t}{\gamma} \equiv -\omega s \tag{4.33}$$

Hence we give the following definitions.

`Definition 1.` *The Guiding Centre (henceforth GC) is the reference frame in which a charged particle moves on a closed and periodic trajectory; its coordinates in the laboratory frame are denoted by $T, \boldsymbol{X}$. The gyrophase ζ is the angle which parametrizes the trajectory performed by a charged particle in the GC description.*

Here we recall that a trajectory is closed and periodic in space iff it is closed in phase space. So the gyrophase is an angle in *phase space*.

If we also introduce a constant $\mu \stackrel{\text{def}}{=} B_0\rho^2$ (where B_0 is the field strength and ρ was called the Larmor radius) and a function $k = (m/e)(\omega/\mu)$ the single particle energy can be written

$$h \;=\; m\sqrt{1 \;+\; \frac{\overline{\boldsymbol{W}}\boldsymbol{W}}{m^2} \;-\; \frac{e}{m}\,\omega\,\mu}$$

[6]We should however mention a work by de Guillebon and Vittot [43] in which an algorithm to push the trasformation to arbitrary order has been proposed.

where $\boldsymbol{W}$ is the guiding centre kinetic momentum. The vector potential for a constant uniform magnetic field is

$$\boldsymbol{A}(\boldsymbol{x}) \,=\, \tfrac{1}{2}\big[\!\big[\boldsymbol{B}\big]\!\big](\boldsymbol{x}-\boldsymbol{X}) \,+\, \partial_{\overline{\boldsymbol{x}}} g$$

with some gauge function g. If the potential is computed on the particle trajectory then it equals

$$\boldsymbol{A}|_{\text{traj}} \,=\, \tfrac{1}{2}\rho B_0 \boldsymbol{e}_\tau \,+\, \partial_{\overline{\boldsymbol{x}}} g$$

The velocity field was instead given by (4.28), the constant α_3 being 0 in the GC frame, so

$$\boldsymbol{w}|_{\text{traj}} \,=\, -e\rho B_0 \boldsymbol{e}_\tau$$

So there is one gauge function in which the integrability condition (4.25) can be achieved, and its value is

$$g \,=\, -\tfrac{1}{2} e B_0 \rho^2 \tau \equiv e\mu\zeta$$

where in the second passage we made the natural identification of the τ angle of cylindrical coordinates with the gyrophase ζ, and we used $\partial_{\overline{\boldsymbol{x}}}\tau \equiv \boldsymbol{e}_\tau/R$, R being equal to ρ on the trajectory. We are lead to this final definition,

`Definition` **2.** *The magnetic moment μ is propotional by (m/e) to the canonical momentum associated to the gyrophase. The gauge function of the Guiding Centre transform is m/e times the product of the magnetic moment and of the gyrophase.*

4.3.1 The Guiding Centre Transform

We start from the Lagrangian description of the charged particle dynamics, which is described in appendix C. The Lagrange form l of classical electrodynamics is exact, so there exists a Poincaré-Cartan 1-form

$$\varpi \,=\, -\overline{\boldsymbol{p}}\, d\boldsymbol{x} \,+\, \varepsilon\, dt \tag{4.34}$$

so that $l \,=\, d\varpi$. The form of Poincaré-Cartan involves the potentials $\boldsymbol{A}$ and Φ (see section 4.1): indeed $\boldsymbol{p}$ and ε are the canonical momentum and canonical energy,

$$\boldsymbol{p} \,\stackrel{\text{def}}{=}\, \boldsymbol{w} \,+\, e\boldsymbol{A}(t,\boldsymbol{x}) \qquad \varepsilon \,\stackrel{\text{def}}{=}\, m\gamma \,+\, e\Phi(t,\boldsymbol{x})$$

So we are shifting to a gauge-dependent description. However, as it can be seen by definition 2, the gauge freedom is a key element to introduce the magnetic moment.

We call *GC transform* $\mathcal{T}_{GC}$ the map conjugating the original phase space coordinates $\{t,\boldsymbol{x},\boldsymbol{w}\}$ to the new set of coordinates $\{T,\boldsymbol{X},\zeta,\mu,\varepsilon\}$, and is chosen in such a way that

$$\mathcal{T}_{GC}\colon \varpi \,\to\, \varpi_{GC}\,, \quad \varpi_{GC} \,=\, P^0\, dT \,-\, \overline{\boldsymbol{P}}\, d\boldsymbol{X} \,+\, \tfrac{m}{e}\,\mu\, d\zeta \tag{4.35}$$

where P^0 and $\boldsymbol{P}$ are functions of $\{T,\boldsymbol{X},\varepsilon,\mu\}$, defined by

$$\begin{aligned} \boldsymbol{P}(T,\boldsymbol{X},\varepsilon,\mu) \,&=\, \boldsymbol{W}(T,\boldsymbol{X},\varepsilon,\mu) + e\boldsymbol{A}(T,\boldsymbol{X}) \\ P^0(T,\boldsymbol{X},\varepsilon,\mu) \,&=\, W^0(T,\boldsymbol{X},\varepsilon,\mu) \,+\, e\Phi(T,\boldsymbol{X}) \end{aligned} \tag{4.36}$$

Here $\boldsymbol{W}$ and W^0, respectively the GC kinetic moment and kinetic energy, are also *fields*. According to definition 2, the magnetic moment μ appears in ϖ_{GC} as proportional to

the momentum canonically associated to the gyrophase, and the term $(m/e)\mu\zeta$ can be interpreted as the gauge function of the GC transform.

We compute

$$\begin{aligned} d\varpi_{GC} \;=\; & d\varepsilon\,\partial_\varepsilon P^0 \wedge dT \;+\; d\mu\,\partial_\mu P^0 \wedge dT \;-\; e\,d\overline{\boldsymbol{X}}\,\boldsymbol{E_c} \wedge dT \;-\; d\varepsilon\,\partial_\varepsilon \overline{\boldsymbol{P}} \wedge d\boldsymbol{X} \;+ \\ & -\; d\mu\,\partial_\mu \overline{\boldsymbol{P}} \wedge d\boldsymbol{X} \;+\; e\,d\overline{\boldsymbol{X}}\,[\![\boldsymbol{B_c}]\!]\,d\boldsymbol{X} \;+\; \tfrac{m}{e}\,d\mu \wedge d\zeta \end{aligned} \tag{4.37}$$

where we introduced again the "canonical magnetic field" $(e/m)\boldsymbol{B_c} \stackrel{\text{def}}{=} [\![\nabla]\!]\boldsymbol{P}$ and the "canonical electric field" $(e/m)\boldsymbol{E_c} \stackrel{\text{def}}{=} -\partial_{\overline{\boldsymbol{X}}} P^0 - \partial_T \boldsymbol{P}$, as in equations (4.16). However, the fields in equation (4.39b) are evaluated at the guiding centre coordinates $\{T, \overline{\boldsymbol{X}}\}$ while those in equation (4.17) at the particle coordinates $\{t, \overline{\boldsymbol{x}}\}$.

We look for a vector field $C = \dot{T}\partial_T + \overline{\dot{\boldsymbol{X}}}\nabla + \dot{\zeta}\partial_\zeta + \dot{\mu}\partial_\mu + \dot{\varepsilon}\partial_\varepsilon$ in the null space of $d\varpi_{GC}$[7],

$$\begin{aligned} d\varpi_{GC}C \;=\; & d\varepsilon\,\partial_\varepsilon P^0\,\dot{T} \;-\; dT\,\partial_\varepsilon P^0\,\dot{\varepsilon} \;+\; d\mu\,\partial_\mu P^0\,\dot{T} \;-\; dT\,\partial_\mu P^0\,\dot{\mu} \;-\; e\,d\overline{\boldsymbol{X}}\,\boldsymbol{E_c}\,\dot{T} \;+ \\ & +\; e\,dT\,\overline{\dot{\boldsymbol{X}}}\,\boldsymbol{E_c} \;+\; e\,d\overline{\boldsymbol{X}}\,[\![\boldsymbol{B_c}]\!]\,\dot{\boldsymbol{X}} \;-\; d\varepsilon\,\partial_\varepsilon \overline{\boldsymbol{P}}\,\dot{\boldsymbol{X}} \;+\; d\overline{\boldsymbol{X}}\,\partial_\varepsilon \boldsymbol{P}\,\dot{\varepsilon} \;+ \\ & -\; d\mu\,\partial_\mu \overline{\boldsymbol{P}}\,\dot{\boldsymbol{X}} \;+\; d\overline{\boldsymbol{X}}\,\partial_\mu \boldsymbol{P}\,\dot{\mu} \;+\; \tfrac{m}{e}\,d\mu\,\dot{\zeta} \;-\; \tfrac{m}{e}\,d\zeta\,\dot{\mu} \;=\; 0 \end{aligned}$$

The expression above is equivalent to the five equations

$$\dot{\mu} \;=\; 0 \tag{4.38a}$$

$$-\,\dot{T}\boldsymbol{E_c} \;+\; [\![\boldsymbol{B_c}]\!]\dot{\boldsymbol{X}} \;+\; \dot{\varepsilon}\partial_\varepsilon \boldsymbol{P} \;=\; 0 \tag{4.38b}$$

$$-\,\overline{\boldsymbol{E_c}}\dot{\boldsymbol{X}} \;+\; \dot{\varepsilon}\partial_\varepsilon P^0 \;=\; 0 \tag{4.38c}$$

$$\tfrac{m}{e}\dot{\zeta} \;=\; -\dot{T}\,\partial_\mu P^0 \;+\; \overline{\dot{\boldsymbol{X}}}\partial_\mu \boldsymbol{P} \tag{4.38d}$$

$$P^0\,\partial_\varepsilon P^0 \;-\; \overline{\boldsymbol{P}}\partial_\varepsilon \boldsymbol{P} \;=\; 0 \tag{4.38e}$$

where we have deleted a term in (4.38b) as a consequence of (4.39a). We have to check the compatibility of equations (4.38b) and (4.38c). We multiply the former scalarly by $\dot{\boldsymbol{X}}$ and the latter by $\dot{T}$,

$$-\dot{T}\,\overline{\boldsymbol{E_c}}\dot{\boldsymbol{X}} \;+\; \dot{\varepsilon}\,\overline{\dot{\boldsymbol{X}}}\partial_\varepsilon \boldsymbol{P} \;=\; 0, \qquad -\dot{T}\,\overline{\boldsymbol{E_c}}\dot{\boldsymbol{X}} \;+\; \dot{\varepsilon}\,\dot{T}\,\partial_\varepsilon P^0 \;=\; 0$$

So, they are compatible in virtue of equation (4.38e), or equivalently, the latter equation is redundant in system (4.38).

Finally, we have the freedom to set $\dot{T} = 1$ to get the equations of motion with respect

[7]Here the overdot denotes derivation with respect to some parameter, say τ. As we will see, we have the freedom to set $\dot{T} = 1$, so that this parameter has the meaning of (GC) time, but also $\dot{T} = \gamma$, so that the parameter may be identified with proper time. See also the appendix C.2.

to the GC time. By setting also $\boldsymbol{V} \equiv d\boldsymbol{X}/dT$, we find

$$\frac{d\mu}{dT} = 0 \tag{4.39a}$$

$$-\boldsymbol{E_c} + [\![\boldsymbol{B_c}]\!]\boldsymbol{V} + \frac{d\varepsilon}{dT}\partial_\varepsilon \boldsymbol{P} = 0 \tag{4.39b}$$

$$-\overline{\boldsymbol{E_c}}\boldsymbol{V} + \frac{d\varepsilon}{dT}\partial_\varepsilon P^0 = 0 \tag{4.39c}$$

$$\frac{m}{e}\frac{d\zeta}{dT} = -\partial_\mu P^0 + \overline{\dot{\boldsymbol{X}}}\partial_\mu \boldsymbol{P} \tag{4.39d}$$

$$P^0\partial_\varepsilon P^0 - \overline{\boldsymbol{P}}\partial_\varepsilon \boldsymbol{P} = 0 \tag{4.39e}$$

As a final note, we can confront equation (4.39b) with equation (4.17). Here we have a term proportional to the derivative in time of the single-particle energy, which doesn't appear in the Eulerian reduction. But in the Eulerian reduction we set $d\overline{\boldsymbol{\alpha}}/dt = 0$, and one of the $\overline{\boldsymbol{\alpha}}$ had the meaning of energy. If we consider a static system, where we would find $d\varepsilon/dT = 0$, the equations would be the same. So, not only in both reductions the velocity of the particle is replaced by a field, but we also find that this field obeys in both cases to the same equation.

4.3.2 From Lagrangian to Hamiltonian Formulation

In this section we build a Poisson structure for the system (4.39). on the phase space $\mathbb{X}$ of coordinates $\{\boldsymbol{X}, \zeta, \mu\, \varepsilon\}$. A generical overview of the method can be found in the appendix C, section C.1. The lagrange form $l = d\varpi_{GC}$ is split as

$$l = \sigma - \mathfrak{h} \wedge dT$$

where σ is the symplectic 2-form and does not contain terms proportional to dT, and $\mathfrak{h}$ the "Hamiltonian 1-form". From the Lagrange form (4.37) we get

$$\sigma = -d\varepsilon\, \partial_\varepsilon \overline{\boldsymbol{P}} \wedge d\boldsymbol{X} - d\mu\, \partial_\mu \overline{\boldsymbol{P}} \wedge d\boldsymbol{X} + e\, d\overline{\boldsymbol{X}}[\![\boldsymbol{B_c}]\!]d\boldsymbol{X} + (m/e)\, d\mu \wedge d\zeta \tag{4.40}$$

$$\mathfrak{h} = e\, d\overline{\boldsymbol{X}}\boldsymbol{E_c} - d\varepsilon\, \partial_\varepsilon P^0 - d\mu\, \partial_\mu P^0 \tag{4.41}$$

Then we search for a (Poisson) bivector π_{GC} which "inverts" σ in the sense that $\sigma Z \pi_{GC} = Z$ for any vector field Z on $\mathbb{X}$. Then the Hamiltonian vector field is given by $\mathfrak{h}\pi_{GC}$. The inverse of σ exists if σ is closed and with zero kernel. Having defined the Hodge differential on $\mathbb{X}$ by

$$d_{\mathbb{X}} = d\overline{\boldsymbol{X}}\partial_{\overline{\boldsymbol{X}}} + d\zeta\partial_\zeta + d\mu\partial_\mu + d\varepsilon\partial_\varepsilon$$

and then the closedness of σ on $\mathbb{X}$ follows from nothing that $\sigma = d_{\mathbb{X}}\big(\frac{m}{e}\mu\zeta - \overline{\boldsymbol{P}}d\boldsymbol{X}\big)$. The candidate π_{GC} has to solve the equations

$$(\sigma\partial_{\overline{\boldsymbol{X}}})\pi_{GC} = \partial_{\overline{\boldsymbol{X}}} \qquad (\sigma\partial_\zeta)\pi_{GC} = \partial_\zeta \qquad (\sigma\partial_\varepsilon)\pi_{GC} = \partial\varepsilon \qquad (\sigma\partial_\mu)\pi_{GC} = \partial_\mu \tag{4.42}$$

We consider an ansatz for π_{GC},

$$\pi_{GC} = \xi\partial_\varepsilon \wedge \partial_\zeta + \partial_{\boldsymbol{X}}\mathcal{A}\partial_{\overline{\boldsymbol{X}}} + \partial_\varepsilon \wedge \overline{\boldsymbol{z}}\partial_{\overline{\boldsymbol{X}}} + \tfrac{e}{m}\,\partial_\mu \wedge \partial_\zeta + \partial_\zeta \wedge \overline{\boldsymbol{s}}\partial_{\overline{\boldsymbol{X}}} \tag{4.43}$$

with vector fields $\boldsymbol{z}$ and $\boldsymbol{s}$, a scalar function ξ and a skew-symmetric matrix $\mathcal{A}$ to be determined. In principle we may also include a term coupling μ and $\boldsymbol{X}$, and another one coupling μ and ε, but they turn out to be necessarily null, so we omitted them. We compute

$$\sigma\partial_{\boldsymbol{X}} \;=\; e\, d\overline{\boldsymbol{X}}\big[\!\big[\boldsymbol{B_c}\big]\!\big] \;-\; d\mu\,\partial_\mu\overline{\boldsymbol{P}} \;-\; d\varepsilon\overline{\boldsymbol{P}}$$

$$\sigma\partial_\varepsilon \;=\; d\overline{\boldsymbol{X}}\partial_\varepsilon\boldsymbol{P} \qquad\qquad \sigma\partial_\zeta \;=\; \tfrac{m}{e}d\mu$$

$$\sigma\partial_\mu \;=\; d\overline{\boldsymbol{X}}\partial_\mu\boldsymbol{P} \;-\; \tfrac{m}{e}d\zeta \;=\; \partial_\mu\overline{\boldsymbol{P}}d\boldsymbol{X} \;-\; e^{-1}d\zeta$$

By plugging (4.43) and the above into equations (4.42), we get

$$\sigma\partial_{\overline{\boldsymbol{X}}}\pi_{GC} = -e\big[\!\big[\boldsymbol{B_c}\big]\!\big]\mathcal{A}\partial_{\overline{\boldsymbol{X}}} + e\big[\!\big[\boldsymbol{B_c}\big]\!\big]\boldsymbol{z}\partial_\varepsilon + e\big[\!\big[\boldsymbol{B_c}\big]\!\big]\boldsymbol{s}\partial_\zeta - \tfrac{e}{m}\partial_\mu\boldsymbol{P}\partial_\zeta - \xi\partial_\varepsilon\boldsymbol{P}\partial_\zeta - \partial_\varepsilon\boldsymbol{P}\overline{\boldsymbol{z}}\partial_{\overline{\boldsymbol{X}}}$$

$$\sigma\,\partial_\varepsilon\,\pi_{GC} \;=\; \partial_\varepsilon\overline{\boldsymbol{P}}\mathcal{A}\partial_{\overline{\boldsymbol{X}}} \;-\; \partial_\varepsilon\overline{\boldsymbol{P}}\boldsymbol{z}\partial_\varepsilon \;-\; \partial_\varepsilon\overline{\boldsymbol{P}}\boldsymbol{s}\partial_\zeta$$

$$\sigma\partial_\zeta\pi_{GC} \;=\; \tfrac{m}{e}\tfrac{e}{m}\partial_\zeta \;=\; \partial_\zeta$$

$$\sigma\partial_\mu\pi_{GC} \;=\; \partial_\mu\overline{\boldsymbol{P}}\mathcal{A}\,\partial_{\overline{\boldsymbol{X}}} \;-\; \partial_\mu\overline{\boldsymbol{P}}\boldsymbol{z}\partial_\varepsilon \;+\; \partial_\mu\overline{\boldsymbol{P}}\boldsymbol{s}\partial_\zeta \;+\; \partial_\mu$$

which are equivalent to the following

$$-\,e\,\big[\!\big[\boldsymbol{B_c}\big]\!\big]\mathcal{A} \;-\; \partial_\varepsilon\boldsymbol{P}\,\overline{\boldsymbol{z}} \;=\mathbb{I} \tag{4.44a}$$

$$\big[\!\big[\boldsymbol{B_c}\big]\!\big]\,\boldsymbol{z} \;=\; 0 \tag{4.44b}$$

$$\xi\,\partial_\varepsilon\boldsymbol{P} \;+\; (e/m)\partial_\mu\boldsymbol{P} \;-\; e\,\big[\!\big[\boldsymbol{B_c}\big]\!\big]\boldsymbol{s} \;=\; 0 \tag{4.44c}$$

$$\partial_\varepsilon\overline{\boldsymbol{P}}\,\boldsymbol{z} \;=\; -1 \tag{4.44d}$$

$$\partial_\varepsilon\overline{\boldsymbol{P}}\mathcal{A} \;=\; 0 \tag{4.44e}$$

$$\partial_\varepsilon\overline{\boldsymbol{P}}\boldsymbol{s} = 0 \tag{4.44f}$$

$$\partial_\mu\overline{\boldsymbol{P}}\mathcal{A} \;-\; (m/e)\overline{\boldsymbol{s}} \;=\; 0 \tag{4.44g}$$

$$\partial_\mu\overline{\boldsymbol{P}}\,\boldsymbol{z} \;-\; (m/e)\xi \;=\; 0 \tag{4.44h}$$

$$\partial_\mu\overline{\boldsymbol{P}}\boldsymbol{s} = 0 \tag{4.44i}$$

The last equation is automatically satisfied as a consequence of (4.44g). Consider the following ansatz:

$$\mathcal{A} \;=\; \frac{m}{e\,B_c}\big[\!\big[\boldsymbol{b_c}\big]\!\big] \qquad \boldsymbol{z} \;=\; -\eta\boldsymbol{b_c} \qquad \partial_\varepsilon\boldsymbol{P} \;=\; \frac{\boldsymbol{b_c}}{\eta} \tag{4.45}$$

for some unknown scalar function η, and having defined $\boldsymbol{b_c} \overset{\text{def}}{=} \boldsymbol{B_c}/B_c$. Now equations (4.44a), (4.44b), (4.44d) and (4.44e) are satisfied. Equations (4.44g) and (4.44h) give then the following definitions

$$\boldsymbol{s} \;=\; -(1/mB_c)\big[\!\big[\boldsymbol{b_c}\big]\!\big]\partial_\mu\boldsymbol{P} \tag{4.46a}$$

$$\xi \;=\; (e/m)\eta\overline{\boldsymbol{b_c}}\partial_\mu\boldsymbol{P} \tag{4.46b}$$

As a consequence (4.44f) is satisfied. Finally, equation (4.44c) gives

$$e\tfrac{-1}{mB_c}\left[\!\left[\boldsymbol{b_c}\right]\!\right]^2\partial_\mu P \,-\, \frac{e}{m}\,\partial_\mu P \,-\, \frac{e}{m}\partial_\varepsilon P\overline{\boldsymbol{z}}\partial_\mu P \,=\, 0$$

$$\big(-\left[\!\left[\boldsymbol{b_c}\right]\!\right]^2 \,-\, \mathbb{I} \,+\, \boldsymbol{b_c}\overline{\boldsymbol{b_c}}\big)\partial_\mu P \,=\, 0$$

which is an identity. We end up with the following bivector,

$$\pi_{GC} \,=\, \frac{e}{m}\,\eta\,\big(\overline{\boldsymbol{b_c}}\partial_\mu\boldsymbol{P}\big)\,\partial_\zeta\wedge\partial_\varepsilon \,+\, \frac{e}{m}\,\partial_\mu\wedge\partial_\zeta \,-\, \eta\,\partial_\varepsilon\wedge\overline{\boldsymbol{b_c}}\partial_{\overline{\boldsymbol{X}}} \,+$$

$$+\,\frac{1}{eB_c}\,\partial_{\boldsymbol{X}}\left[\!\left[\boldsymbol{b_c}\right]\!\right]\partial_{\overline{\boldsymbol{X}}} \,+\, \partial_\zeta\wedge\frac{\partial_\mu\overline{\boldsymbol{P}}\left[\!\left[\boldsymbol{b_c}\right]\!\right]}{mB_c}\,\partial_{\overline{\boldsymbol{X}}} \qquad (4.47)$$

As a final note, we underline that the condition for the existence of (4.47) is that $\boldsymbol{B_c}\neq 0$. This is the condition for σ to have empty kernel.

Hamilton's equations

With the Hamiltonian form (4.41), and the bivector (4.47) we find the Hamiltonian vector field

$$\mathfrak{h}\pi_{GC} \,=\, \eta\,\partial_\varepsilon P^0\,\overline{\boldsymbol{b_c}}\partial_{\overline{\boldsymbol{X}}} \,-\, \frac{e}{m}\,\partial_\mu P^0\,\partial_\zeta \,+\, \overline{\boldsymbol{E_c}}\frac{\left[\!\left[\boldsymbol{B_c}\right]\!\right]}{B_c^2}\,\partial_{\overline{\boldsymbol{X}}} \,+$$

$$+\,\frac{e}{m}\,\eta\,\overline{\boldsymbol{b_c}}\boldsymbol{E_c}\partial_\varepsilon \,-\, \frac{e}{mB_c^2}\,\big(\overline{\boldsymbol{E_c}}\left[\!\left[\boldsymbol{B_c}\right]\!\right]\partial_\mu\boldsymbol{P}\big)\partial_\zeta \,+\, \eta\,\frac{e}{m}\,\partial_\varepsilon P^0\,\big(\overline{\boldsymbol{b_c}}\partial_\mu P\big)\,\partial_\zeta$$

which is equivalent to the equations of motion

$$\frac{d\overline{\boldsymbol{X}}}{dT} \,=\, \eta\,\partial_\varepsilon P^0\,\overline{\boldsymbol{b_c}} \,+\, \frac{\overline{\boldsymbol{E_c}}\left[\!\left[\boldsymbol{B_c}\right]\!\right]}{B_c^2}$$

$$\frac{d\zeta}{dT} \,=\, \frac{e}{m}\left(\eta\partial_\varepsilon P^0\,\overline{\boldsymbol{b_c}}\partial_\mu\boldsymbol{P} \,+\, \frac{\overline{\boldsymbol{E_c}}\left[\!\left[\boldsymbol{B_c}\right]\!\right]\partial_\mu\boldsymbol{P}}{B_c^2} \,-\, \partial_\mu P^0\right)$$

$$\frac{d\varepsilon}{dT} \,=\, \eta\,\overline{\boldsymbol{b_c}}\boldsymbol{E_c}$$

$$\frac{d\mu}{dT} \,=\, 0$$

As η is free, we define

$$\eta\partial_\varepsilon P^0 \,=\, V_b \overset{\text{def}}{=} \overline{\boldsymbol{b_c}}\boldsymbol{V} \qquad (4.48)$$

and the system above is reduced to

$$\frac{d\boldsymbol{X}}{dT} \,=\, V_b\boldsymbol{b_c} \,+\, \frac{\left[\!\left[\boldsymbol{E_c}\right]\!\right]\boldsymbol{B_c}}{B_c^2} \qquad (4.49a)$$

$$\frac{d\zeta}{dT} \,=\, \frac{e}{m}\,\big(\overline{\boldsymbol{V}}\,\partial_\mu\boldsymbol{P} \,-\, \partial_\mu P^0\big) \qquad (4.49b)$$

$$\frac{d\varepsilon}{dT} \,=\, \frac{V_b\,\overline{\boldsymbol{b_c}}\boldsymbol{E_c}}{\partial_\varepsilon P^0} \qquad (4.49c)$$

$$\frac{d\mu}{dT} \,=\, 0 \qquad (4.49d)$$

The compatibility of the system (4.49) with (4.38) is clear, after one considers the following: from the ansatz (4.45) and the definition (4.48) we get

$$\partial_\varepsilon \boldsymbol{P} \;=\; \boldsymbol{b_c}\partial_\varepsilon P^0/V_b \tag{4.50}$$

then, if we multiply equation (4.38b) on the left by $[\![\boldsymbol{B_c}]\!]$ we get

$$\boldsymbol{V} \;=\; V_b\boldsymbol{b_c} \;+\; \frac{[\![\boldsymbol{E_c}]\!]\,\boldsymbol{B_c}}{B_c^2} \tag{4.51}$$

and if this result is plugged into $d\boldsymbol{X}/dT = \boldsymbol{V}$, it leads to equation (4.49a).

Finally, we stress the analogy between equation (4.51) and equation (4.32), which is an exact equation, but only for constant and uniform fields, and between equation (4.51) and equation (3.13) of Cary and Brizard [15]: in that work, instead of the canonical fields, the "starred" fields $\boldsymbol{E}^*, \boldsymbol{B}^*$ appear; in the light of this new approach they can be interpreted as a first order approximation to the canonical fields. Or, the canonical fields are the resummation of any perturbative expansion.

4.3.3 Conservation of the Lagrangian

The lagrangian function L is defined by $\varpi = Lds$, s being the proper time, and for the charged particle, by using (4.34),

$$L \;=\; -m \;-\; e\big(\Phi(t,\boldsymbol{x})u^0 \;-\; \overline{\boldsymbol{A}}(t,\boldsymbol{x})\boldsymbol{u}\big)$$

where $u^0 \equiv dt/ds = \gamma$ and $\boldsymbol{u} \stackrel{\text{def}}{=} d\boldsymbol{x}/ds = \gamma\boldsymbol{v}$. Instead by $\varpi_{GC} = L_{GC}ds$ we get

$$L_{GC} \;=\; -\,W^0U^0 \;+\; \overline{\boldsymbol{U}}\boldsymbol{W} \;-\; e\big(\Phi(T,\boldsymbol{X})U^0 \;-\; \overline{\boldsymbol{A}}(T,\boldsymbol{X})\boldsymbol{U}\big) \;+\; \mu\frac{m}{e}\frac{d\zeta}{ds}$$

A reasonable hypothesis is to ask invariance in form of the Lagrangian [23]: we want the new trajectories to describe again a charged particle. This is achieved by imposing

$$\frac{m}{e}\mu\frac{d\zeta}{ds} \;=\; -\,m \;-\; U^0W^0 \;+\; \overline{\boldsymbol{W}}\boldsymbol{U}$$

Now we recall equation (4.38d) and we inject it in the above, with the result

$$\begin{aligned}
-\mu\big(\overline{\boldsymbol{U}}\partial_\mu\boldsymbol{W} - U^0\partial_\mu U^0\big) \;&=\; -\,m \;+\, U^0W^0 \;-\, \overline{\boldsymbol{U}}\boldsymbol{W}\\
\log\big((W^0)^2 - \overline{\boldsymbol{W}}\boldsymbol{W} - m^2\big) \;&=\; \log\big(e\,k(T,\boldsymbol{X})\,\mu\big)^2\\
W^0 \;&=\; \sqrt{m^2 + \overline{\boldsymbol{W}}\boldsymbol{W} \;+\; e^2\,k^2(T,\boldsymbol{X})\,\mu^2}
\end{aligned} \tag{4.52}$$

for some real $k = k(T,\overline{\boldsymbol{X}})$ which is constant with respect to μ and also to ε (as a consequence of equation (4.38e)).

In special relativity "massive particles" are point particles obeying the *dispersion relation* $w^0 = \sqrt{m^2 + \overline{\boldsymbol{w}}\boldsymbol{w}}$. Equation (4.52) implies that in the limit $\mu \to 0$ the GC becomes a massive particle, otherwise it is not, and the magnetic moment measures this difference.

From equation (4.52) we can compute

$$\partial_\mu W^0 \;=\; \overline{\boldsymbol{V}}\partial_\mu\boldsymbol{W} \;+\; \frac{k^2\mu}{W^0}$$

By plugging the above into equation (4.39d) or into equation (4.38d) we respectively get

$$\frac{d\zeta}{dT} = -\frac{e}{m}\frac{k^2\mu}{W^0} \iff \frac{d\zeta}{ds} = -\frac{e}{m}k^2\mu \tag{4.53}$$

and the latter equation defines the gyrofrequency $\omega \stackrel{\text{def}}{=} -(e/m)k^2\mu$. We may finally rewrite

$$\frac{W^0}{m} = \sqrt{1 + \frac{\overline{\boldsymbol{W}}\boldsymbol{W}}{m^2} - \frac{e}{m}\omega\,\mu}$$

4.3.4 A hint for future developments

In the Hamiltonian description, we would involve $\boldsymbol{V}$ rather than $\boldsymbol{U}$. If we assume, by analogy with the particle description, that $\boldsymbol{W} = W^0\boldsymbol{V}$, we may rewrite equation (4.52) as

$$(W^0)^2 = m^2 + \overline{\boldsymbol{W}}\boldsymbol{W} + e^2k^2\mu^2 \implies W^0 = \sqrt{\frac{m^2 + e^2k^2\mu^2}{1 - \overline{\boldsymbol{V}}\boldsymbol{V}}} \equiv \gamma\big|_V\sqrt{m^2 + e^2\,k^2\,\mu^2}$$

or, upon setting $\kappa \equiv ek/m$,

$$U^0 = \gamma\big|_V\sqrt{1 + \kappa^2\,\mu^2}$$

In terms of metric, $U^0 = dT/ds$; if we also define $d\tau = \sqrt{(e/m)\mu}\,d\zeta$, then the last formula can be rewritten in the following suggestive way,

$$dT^2 = ds^2 + dX^2 + d\tau^2 \tag{4.54}$$

to be confronted with the usual formula of special relativity (see [50])

$$dt^2 = ds^2 + dx^2$$

The latter says that the (infinitesimal(time interval experienced by a particle in the lab frame is the sum of a (infinitesimal) displacement in space plus a (infinitesimal) increment of the proper-time. Equation (4.54) says that the infinitesimal time interval experienced by a particle in the GC frame is proportional to the sum of the (infinitesimal) GC displacement, plus the (infinitesimal) proper-time increment, plus the (infinitesimal) gyrorotation. Equation (4.54) may also be read as the definition of a metric over a 5-dimensional space. For more quantitative considerations on this last aspect, see [23].

4.4 Conclusions

In this chapter we studied two possible reduction theories for the relativistic charged particle dynamics. They are methods to reduce the number of degrees of freedom of this dynamical system.

One is Guiding Centre theory, which has been studied for decades in plasma physics. In this theory we aim at associating a pair of hamiltonian action-angle coordinates μ, ζ to the (fast) gyrating motion of a charged particle around a magnetic field line. The remaining degrees of freedom describe the (slow) motion of a point sticking to the field lines (the Guiding Centre). We focused on a novel non-perturbative and relativistic approach, introduced in [22], [23]. Emphasis is put not on the explicit form of the GC transform,

which we don't claim to know[8]; instead we impose some properties of the new Guiding Centre coordinates, and deduce a new Lagrange form. The Guiding Centre transform in then exact, but the drawback is that we don't know its explicit form, nor the relation between the particle coordinates and the GC coordinates. However, it has been shown [23] that, if the non-perturbative Lagrange form is expanded in the inverse of the magnetic field strength, at first order the result coincides with those of the perturbative approach. We have computed for the first time the new equations of motion, and their Poisson structure.

We also presented an entirely new theory, that we called the Eulerian reduction. In this second case we aim at finding the vector field associated to the Lorentz equation, so that half of the degrees of freedom of the problem is cut away. It was natural to build this theory as a change of variable from the particle coordinates $\{t, \overline{\boldsymbol{x}\boldsymbol{w}}\}$ to a new set $\{t, \overline{\boldsymbol{x}}, \overline{\boldsymbol{\alpha}}\}$ for which we have also computed the Poisson structure. The $\overline{\boldsymbol{\alpha}}$ are constants, with values determined by the initial data on the velocity. We mention (see section 4.2.1 that very similar equations were found in [22], in what was called the "Guiding Particle" solution of motion.

Both after the Eulerian and GC reductions, the kinetic momentum $\boldsymbol{w}$ is a field obeying

$$\boldsymbol{w} = \frac{(\overline{\boldsymbol{B}_c}\boldsymbol{w})\,\boldsymbol{B}_c + [\![\boldsymbol{E}_c]\!]\,\boldsymbol{B}_c}{|\boldsymbol{B}_c|^2}$$

Unfortunately, this is not an algebraic relation, but a very complicated Partial Differential Equation. We recall that in the two cases the fields are evaluated at different coordinates: at the particle coordinates for the Eulerian velocity, at the GC coordinates for the GC velocity. Nevertheless, the fact that the same equation appears suggests that some relation exists between the two theories; understanding this relation may be the key to fill the gaps still affecting both of them.

Aknowledgements

I would like to sincerely thank Claudio Di Troia for his help, patience, guidance and support while developing the work discussed in the second part of this chapter. And to beg his pardon for changing all of his notations in one night!

[8]In particular we don't claim that all magnetic fields are well approximated by a straight and uniform one.

5

Conclusions and Future Perspectives

So, the point of this work was to propose a perturbation theory for non-canonical Poisson systems. Many dynamical systems in physics are Poisson systems. Classical mechanics is a very particular example of Poisson system, so particular that it is called "canonical". The KAM theorem is a way to perform perturbation theory in classical mechanics, with nice mathematical properties. Even when it is possible to turn a Poisson system into a canonical one, this may require cumbersome calculations and/or restrictive hypothesis, so that being able to work in the original setting would be desirable. Extensions of the KAM theorem have already been provided for classical mechanics without action-angle variables [20], presymplectic mechanics [2], vector fields [8]; a conjecture for Lie algebroids has also been proposed [1]. We took inspiration from the KAM theorem to build a first order perturbation formula for Poisson systems (equivalent to what is commonly called "iterative lemma" in KAM theory); we applied it to study the dynamics of a non-autonomous Top, and we showed that our formula could be iterated to provide a full KAM theorem. Then, in an effort to apply our method also to classical electrodynamics, we studied the dynamics of a charged particle in a given electromagnetic field. This material was divide into three chapters, and we tried to make each of them self-contained, so here we will focus only on the links between the different chapters, and the most interesting hints for future developments.

In chapter 2 we showed how to build a Lie series to perform a first order perturbation reduction on a generical Poisson algebra $\mathbb{V}$. We start from a derivation of $\mathcal{H}$ and we split it into a "main part" $\mathcal{H}_0$, which preserves a given subalgebra $\mathbb{B} \subseteq \mathbb{V}$, and a "perturbation" $\{V\}$; then we conjugate the given system to a second system, which preserves $\mathbb{B}$ at first order in the magnitude of V. To build the conjugation, we need a kind of *right inverse* Γ of the Hamiltonian, that can be defined "outside" of $\mathbb{B}$. To bound this operator a scale of Banach norms was needed; we don't know any general method to build them outside the classical setting, so we required generically the existence of such a metric structure. An iteration mechanism to recover a full KAM theorem was described qualitatively, and it was made quantitative in the examples of chapter 3. The most important development of this part would be to identify an algorithm to build the "right inverse" Γ at each iteration of the "main step map"; in principle this would also provide a sequence of functions to bound their norm at each step. A second improvement would be to provide a generical scale of Banach norms for Poisson algebras. With these elements it should be possible to build quantitatively the iteration in the generical setting.

In the application to the symmetric Throbbing Top, we proposed an awkward expression for the right inverse Γ, which requires the unperturbed Hamiltonian. When compared

to the clean expression of the same operator in the context of the non-periodic Top, we feel that a more straighforward expression for Γ is waiting to be unveiled; it may also shed some light on the general iteration process and also on a general method to build the right inverse, adapted to $\mathbb{B}$ but not to the Hamiltonian.

In the study of the non-periodic Throbbing Top, we see that by our algebraic perturbation method we managed to conjugate a non-autonomous system to an autonomous one. This creates a parallel with Floquet theory, and also a starting point (hope?) for an extension of the latter to non-periodic systems. A major point needing improvement here, is the very rough estimates for the Banach norms; we couldn't apply the common tools of KAM theory: neither weighted Fourier norms, as we don't have a Fourier series in time, nor the Cauchy estimates, because they require analiticity in time, so they exclude the constant functions, which are exactly the set that we were targeting.

In chapter 4 we focused on the non-perturbative approach to Guiding Centre theory [23], which we consider a major novelty in the field; it sheds new light on the physical meaning of the Guiding Centre reduction, leaving aside the cumbersome calculations and heavy mathematical artillery that often come together with the more widespread approach of Littlejohn [15]. It allows to go beyond the perturbative expansion, which convergence is a major issue, often ignored [14]. And it is naturally set in a relativistic framework, which is not the case for the perturbative approach. The major drawback is that the explicit form of the Guiding Centre trasform has not been provided yet, so that the relation between the new coordinates and the old ones is not known; one interesting development would be to build it, at least in some simple example. At the same time, the Guiding Centre velocity is the solution of a partial differential equation involving the explicit expressions of the electromagnetic fields: it may be possible to work out the Guiding Centre coordinates directly from this solution, so that their relation with the particle coordinates could be deduced a posteriori. Then, again, the most interesting development would be to provide working examples.

" If you had really started this affair, you might be expected to finish it. But you know well enough now that *starting* is too great a claim for any, and that only a small part is played in great deeds by any hero. "

Gandalf, in *The Lord of the Rings* by J. R. R. Tolkien

A

Some mathematical tools

A.1 About Lie Algebras

For the contents of this section the main reference is [46].

A *Lie algebra* is a vector space $\mathbb{V}$ over a field $\mathbb{K}$ with a bilinear operation $\{\,,\}$ which is alternating (here $V, W, Z \in \mathbb{V}$)

$$\{V, W\} \,=\, -\{W, V\}$$

and satisfies the Jacobi identity

$$\{V, \{W, Z\}\} \,+\, \{W, \{Z, V\}\} \,+ \{Z, \{V, W\}\} \,=\, 0$$

This operation is called a *Lie bracket*.

If on $\mathbb{V}$ we define both a Lie bracket and an associative product

$$(A \cdot B) \cdot C \,=\, A \cdot (B \cdot C)$$

such that the Leibnitz identity holds,

$$\{A,\, B \cdot C\} \,=\, \{A, B\} \cdot C \,+\, B \cdot \{A, C\}$$

then $\mathbb{V}$ is a *Poisson algebra* (and the bracket is called *Poisson bracket*).

The *dual* of $\mathbb{V}$ is the set $\mathbb{V}^*\colon \mathbb{V} \to \mathbb{K}$. The space of *endomorphisms* of $\mathbb{V}$ is defined by

$$\operatorname{End}\mathbb{V} = \mathbb{V} \otimes \mathbb{V}^* \tag{A.1}$$

This space is a Lie algebra on its own with the bracket given by the commutator $[\,,]$,

$$[\mathcal{F}, \mathcal{G}] = \mathcal{F}\mathcal{G} - \mathcal{G}\mathcal{F}, \quad \mathcal{F}, \mathcal{G} \in \operatorname{End}\mathbb{V}$$

The space of *derivations* of $\mathbb{V}$ is defined by

$$\operatorname{der}\mathbb{V} \stackrel{\text{def}}{=} \{\, \mathcal{D} \in \operatorname{End}\mathbb{V} \text{ s.t. } \forall\, V, W \in \mathbb{V}, \qquad \mathcal{D}\{V, W\} \,=\, \{\mathcal{D}V, W\} \,+\, \{V, \mathcal{D}W\}\,\} \tag{A.2}$$

and it'a Lie subalgebra of $\operatorname{End}\mathbb{V}$.

For any element $F \in \mathbb{V}$, we can consider the application "bracket with F" which is an element of $\operatorname{der}\mathbb{V}$ which maps any $G \in \mathbb{V}$ into $\{F, G\} \in \mathbb{V}$. Analogously, for any $\mathcal{F} \in \operatorname{der}\mathbb{V}$ we can consider the application "bracket with $\mathcal{F}$" which, in this case, is an element of $\operatorname{der}\operatorname{der}\mathbb{V}$ and maps any $\mathcal{G} \in \operatorname{der}\mathbb{V}$ into $[\mathcal{F}, \mathcal{G}] \in \operatorname{der}\mathbb{V}$. We introduce the following notation for this applications: we write the bracket with a single argument. So $\{H\} \equiv \{H, \cdot\}$ and $[H] \equiv [\mathcal{H}, \cdot]$. In the manuscript we adopted directly this notation, and we nearly never wrote the bracket with two arguments. The image of "bracket with F" is always a derivation; any derivation built in this way is called an *inner* derivation. A derivation which is not inner is called *outer*.

For any derivation $\mathcal{A}$ the operator

$$e^{\mathcal{A}} \equiv \sum_{n=0}^{\infty} \frac{\mathcal{A}^n}{n!} \tag{A.3}$$

is called a *Lie series*.

A dynamical system on the algebra is given by

$$\begin{cases} \dot{F} = \mathcal{H}F \\ F(0) = F_0 \end{cases}$$

where $\mathcal{H}$ is a derivation, either inner or outer. Its formal solution is $F(t) = e^{t\mathcal{H}}F_0$, and $e^{t\mathcal{H}}$ is called the *flow* of $\mathcal{H}$.

There exists an application $\wedge$ on $\mathbb{V}$, called *exterior product* or *wedge product* which acts in the following way: if we call "vectors" the elements of $\mathbb{V}$, then the exterior product of two vectors is a "bivector",

$$\boldsymbol{v},\, \boldsymbol{w} \;\mapsto\; \boldsymbol{v} \wedge \boldsymbol{w}$$

Its name is due to the following: if σ is an element of $\mathbb{V}^*$, then it maps a byvector into a vector

$$\sigma(\boldsymbol{v} \wedge \boldsymbol{w}) = (\sigma\boldsymbol{v})\boldsymbol{w} - (\sigma\boldsymbol{w})\boldsymbol{v}$$

The exterior product is alternating by definition.

We say that the algebra $\mathbb{V}$ is *metric* if there exists an application

$$\bar{\cdot\cdot}\colon \mathbb{V} \to \mathbb{V}$$

called the transposition, and with the property

$$\bar{v}\, w \;=\; \bar{w}\, v \quad \forall v, w \in \mathbb{V} \tag{A.4}$$

If the transposition is injective, it is possible to identify

$$\bar{\mathbb{V}} \;=\; \mathbb{V}^*$$

Moreover, we say that the transposition has the Killing property if

$$\bar{v}\{v\} \;=\; 0 \tag{A.5}$$

A.1.1 Metric structure on the algebra

The main source for this section is the book by Reed and Simon on functional analysis [80].

An expression like (A.3) has only a formal meaning, unless we introduce a quantitative way to define a notion of convergence. This can be done by a introducing a *norm*. A (Banach) norm is a function $\|\cdot\|\colon \mathbb{V} \to \mathbb{R}_+$ (where $\mathbb{R}_+$ are the positive real numbers) with properties

$$\| A + B \| \leq \|A\| + \|B\| \tag{A.6}$$
$$\|\lambda A\| = |\lambda|\|A\| \tag{A.7}$$
$$\|A\| = 0 \implies A = 0 \tag{A.8}$$

A function from $\mathbb{V}$ to $\mathbb{R}_+$ which satisfy only properties A.6 and A.7 is called a *seminorm* instead.

Most of times it's impossible to bound from above a derivation by means of a norm. By "bounding a derivation" $\mathcal{D} \in \operatorname{der} \mathbb{V}$ we mean that $\forall A \in \mathbb{V}$,

$$\|\mathcal{D}A\| \leq \|A\|$$

This becomes possible if instead of a norm, we consider a *scale of norms* $\{\|\cdot\|_s\}_{s\in\mathbb{I}}$, where s is called an "index" and $\mathbb{I}$ is some set, usually the positive integers or the positive reals. For an algebra $\mathbb{V}$ with a scale of Banach norms indexed by $s \in \mathbb{I}$ we introduce the notation

$$\mathbb{V}_s = \{f \in \mathbb{V} \quad \text{s.t.} \quad \|f\|_s \leq \infty\}$$

and we assume by convention that

$$W \in \mathbb{V}_{s_1} \implies W \in \mathbb{V}_{s_2}, \forall s_2 < s_1$$

We call $\{\mathbb{V}_s\}_{s\in\mathbb{I}}$ a *scale of Banach spaces.*

We say that a derivation $\mathcal{D}$ is *bounded with loss* if

$$\|\mathcal{D}A\|_{s-\delta} \leq \alpha(\delta)\|A\|_s$$

for any $A \in \mathbb{V}_s$, $s, \delta \in \mathbb{I}$.

`Example 1.` *A paradigmatical example is the following [41]: on the complex plane $\mathbb{C}$ we define the sets*

$$\mathbb{B}_r(0) \stackrel{def}{=} \{z \in \mathbb{C}\, s.t.\, |z| < r\}$$

Then, on the space $\mathcal{C}^\infty(\mathbb{C})$ we consider the scale of norms (indexed by $r \in \mathbb{R}_+$)

$$|f(z)|_r \stackrel{def}{=} \sup_{z\in\mathbb{B}_r} |f(z)|$$

The Cauchy inequality states that $|\partial_z f(0)| \leq \frac{1}{r}|f(z)|_r$ from which we get the upper bound

$$|\partial_z f|_{r-\delta} \leq \frac{1}{\delta}|f|_r$$

So by loosing a "layer" of width δ of the original domain, it was possible to bound from above the derivation operator on $\mathbb{C}$.

A.1.2 Non-autonomous Hamiltonian Systems

Here we discuss a general method to add an explicit time-dependence to a dynamical system on a Lie algebra; it is an alternative to the commonly employed methods of autonomization [31] or contact structures [10].

Let $\mathbb{K}$ be an additive group, and consider the following space,

$$\tilde{\mathbb{V}} \stackrel{\text{def}}{=} C^\infty(\mathbb{K} \mapsto \mathbb{V}) \cong \mathbb{V} \otimes C^\infty(\mathbb{K} \to \mathbb{R})$$
$$\tilde{\mathbb{V}} \ni v(\cdot)\colon t \mapsto v(t) \in \mathbb{V}\,,\ \forall t \in \mathbb{K}$$

The variable $t \in \mathbb{K}$ will represent time, so $\mathbb{K} = \mathbb{T}$ (a torus) if we consider a periodic time-dependence, or $\mathbb{K} = \mathbb{R}$ for a non-periodic time-dependence.

We extend the bracket of $\mathbb{V}$ to $\tilde{\mathbb{V}}$ by the following rule,

$$\forall v, w \in \mathbb{V}, \quad [v, w](t) \equiv t \mapsto [v(t), w(t)]$$

and so $\tilde{\mathbb{V}}$ inherits the Lie-algebra structure of $\mathbb{V}$.

The operator $\partial_t\colon \tilde{\mathbb{V}} \to \tilde{\mathbb{V}}$, defined by

$$\partial_t\colon v(t) \mapsto \frac{dv(t)}{dt}$$

is a derivation of $\tilde{\mathbb{V}}$: infact, by the linearity of ∂_t and the bilinearity of $\{\,,\}$, we have

$$\partial_t\{v, w\}(t) \;=\; \left\{\frac{dv(t)}{dt}, w(t)\right\} + \left\{v(t), \frac{dw(t)}{dt}\right\},\ \forall v, w \in \mathbb{V}$$

Finally, we can define a non-autonomous dynamical system on $\tilde{\mathbb{V}}$,

$$\dot{F} = \mathcal{H}F + \partial_t F \tag{A.9}$$

This choice is made for coherence: if F is time independent, then we have the same dynamics of $\mathbb{V}$, while $\dot{t} = 1$ as one would naturally expect.

A.1.3 The Lie-Poisson theorem

The Lie-Poisson theorem states that, if $\mathbb{V}$ is a Lie algebra with some bracket $[\,,]$, then the space $\mathcal{C}^\infty(\mathbb{V}^*)$ is a Poisson algebra, called the "Lie-Poisson" algebra associated to $\mathbb{V}$, with an associative product $\circ$ and a Lie bracket $\{\,,\}$. We propose a constructive proof of this theorem. It is only a formal proof, as we won't introduce any metric structure to study the convergence of the series involved. This proof also enlightens an analogy with a bosonic Fock space in quantum mechanics.

We denote by $\mathbb{V}^*$ the dual space of $\mathbb{V}$, i.e. the space of linear functionals on $\mathbb{V}$,

$$\mathbb{V}^* \stackrel{\text{def}}{=} \{\sigma : \mathbb{V} \mapsto \mathbb{R},\ \sigma \text{ linear}\} \tag{A.10}$$

Then the double dual space of $\mathbb{V}$ is the dual space of $\mathbb{V}^*$,

$$\mathbb{V}^{**} \stackrel{\text{def}}{=} C^\infty(\mathbb{V}^*) \equiv \{f\colon \mathbb{V}^* \to \mathbb{R}\} \tag{A.11}$$

It is easy to show that $\mathbb{V} \subset \mathbb{V}^{**}$; indeed, one can build the map

$$\mathbb{V} \mapsto \mathbb{V}^{**}\colon v \mapsto \check{v} \tag{A.12}$$
$$\check{v}(\sigma) \stackrel{\text{def}}{=} \sigma(v), \forall \sigma \in \mathbb{V}^*$$

We call $\check{\mathbb{V}}$ the image of the map (A.12); moreover, for the rest of this section, we'll denote by small letters $(v, w, \dots)$ elements of $\mathbb{V}$ and by the same checked letters $(\check{v}, \check{w}, \dots)$ their images by the map (A.12).

On $\check{\mathbb{V}}$ we can define an associative product $\circ$ by

$$(\check{v} \circ \check{w})\, \sigma \stackrel{\text{def}}{=} \sigma(v)\sigma(w)$$

where, on the right, we have the product of two real numbers. The new product $\circ$ inherits linearity and commutativity from the product on $\mathbb{R}$; moreover it is associative by construction:

$$((\check{v} \circ \check{w}) \circ \check{z})\, \sigma = \sigma(v)\, \sigma(w)\, \sigma(z)$$

On $\check{\mathbb{V}}$ we can also define a Lie product $\{\cdot\}$ by

$$\big(\{\check{v}\}\check{w}\,\big)\, \sigma \stackrel{\text{def}}{=} \sigma([v]w) \tag{A.13}$$

The properties of linearity, antisymmetry and Jacobi identity of $\{\cdot\}$ follow from those of $[\,,]$ and from the linearity of σ. Finally, we have the freedom to impose the Leibnitz property of $\{\cdot\}$ with respect to $\circ$ by construction:

$$\{\check{v} \circ \check{w}\}\check{z} \stackrel{\text{def}}{=} \check{v} \circ \{\check{w}\}\check{z} + \big(\{\check{v}\}\check{z}\,\big) \circ \check{w}$$

Next, we define the space of "n-vectors" as the n-th power of $\check{\mathbb{V}}$ by itself,

$$\check{\mathbb{V}}^{\circ n} = \{\check{v}_1 \circ \check{v}_2 \circ \cdots \circ \check{v}_n;\ v_i \in \check{\mathbb{V}}\ i = 1, \dots, n\}$$

The space

$$\hat{\mathbb{V}} \equiv \bigoplus_{n=0}^{\infty} \check{\mathbb{V}}^{\circ n} \tag{A.14}$$

is the algebra of sums of n-vectors; we may call them "polynomials". This space, with the product $\circ$ and $\{\cdot\}$, is the Lie-Poisson algebra.

A.2 Complexification of a Lie Algebra

Let $\mathbb{V}$ be a Lie algebra with bracket $\{\,\}$. The complexification of $\mathbb{V}$ is defined as the set

$$\mathbb{V}^C \stackrel{\text{def}}{=} \mathbb{V} \oplus \mathbb{V}$$

with the rules

$$\{v \oplus w\}(x \oplus y) \stackrel{\text{def}}{=} (\{v\}x - \{w\}y) \oplus (\{w\}x + \{v\}y) \tag{A.15}$$

$$(v \oplus w)^* \stackrel{\text{def}}{=} v \oplus (-w) \tag{A.16}$$

$$\mathcal{I}(v \oplus w) \stackrel{\text{def}}{=} (-w) \oplus v \tag{A.17}$$

The second and third rule, in particular, agree with the usual definition of transposition and imaginary unit of the complex numbers. If $\mathbb{V}$ is metric with the Killing property, the complexification $\hat{\mathbb{V}}$ is endowed with the transposition

$$\overline{v \oplus w} = \bar{v} \oplus \bar{w} \tag{A.18}$$

We check that the property (A.4) is satisfied,

$$\begin{aligned}\overline{v \oplus w}\,(x \oplus y) &= (\bar{v} \oplus (-\bar{w}))(x \oplus y) = \bar{v}x - \bar{w}y \\ \overline{x \oplus y}\,(v \oplus w) &= (\bar{x} \oplus (-\bar{y}))(v \oplus w) = \bar{x}v - \bar{y}w\end{aligned}$$

We can check that the complexified algebra has also the Killing property (if the original algebra does),

$$\begin{aligned}\overline{v \oplus w}\,\{v \oplus w\}\,(x \oplus y) &= (\bar{v} \oplus (-\bar{w}))\{v \oplus w\}(x \oplus y) = \\ &= (\bar{v} \oplus (-\bar{w})(\{v\}x - \{w\}y) \oplus (\{w\}x + \{v\}y) = -(\bar{v}\{w\} + \bar{v}\{w\})y \qquad \text{(A.19)}\end{aligned}$$

but by the Killing property on $\mathbb{V}$ we have

$$\overline{v + w}\{v + w\} = 0 = \bar{v}\{w\} + \bar{w}\{v\}$$

so also equation (A.19) equals 0.

A.3 About $\mathbb{R}^n$

We denote by boldletters vectors on $\mathbb{R}^n$; their set is again $\mathbb{R}^n$. The space $\mathbb{R}^n$ has a canonical metric structure, that we will denote by an overbar. As a consequence, $\mathbb{R}^{n*} \cong \overline{\mathbb{R}^n}$; here $\mathbb{R}^{\ltimes *}$ is defined according to (A.10), and its elements are called *covectors*. For example, if $\boldsymbol{v} \in \mathbb{R}^n$, then

$$\boldsymbol{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad \overline{\boldsymbol{w}} = (w_1, \ldots, w_3), \quad \overline{\boldsymbol{w}}\boldsymbol{v} = w_1 v_1 + \ldots + w_3 v_3$$

A matrix M is an object mapping a vector into a vector, so it can be represented as the justaposition of a vector and a covector,

$$M = \boldsymbol{a_1}\overline{\boldsymbol{a_2}},\ \boldsymbol{a_1}, \boldsymbol{a_2} \in \mathbb{R}^n$$

The derivative operator $\partial_{\overline{\boldsymbol{x}}}$ is defined by

$$\overline{\boldsymbol{v}}\partial_{\overline{\boldsymbol{x}}} f = \lim_{t \to 0} \frac{f(\overline{\boldsymbol{x}} + t\overline{\boldsymbol{v}}) - f(\overline{\boldsymbol{x}})}{t}, \quad \forall f \colon \mathbb{R}^{n*} \to \mathbb{R}$$

The result is a vector: in fact we act on $\partial_{\overline{\boldsymbol{x}}} f$ with a covector $\overline{\boldsymbol{v}}$ to get a scalar.

A.3.1 Vector fields

Let $\mathbb{W}$ be an affine space, called the *phase space*. We call *vector field* an element of $\mathcal{C}^\infty(\mathbb{W} \to \mathbb{R}^n)$. All of the notation above evidently applies to vector fields as well. We can also derive a vector field:

$$\overline{\boldsymbol{v}}\partial_{\overline{\boldsymbol{x}}} \boldsymbol{w} = \lim_{t \to 0} \frac{\boldsymbol{w}(\overline{\boldsymbol{x}} + t\overline{\boldsymbol{v}}) - \boldsymbol{w}(\overline{\boldsymbol{x}})}{t}, \quad \forall\, \boldsymbol{w} \in \mathcal{C}^\infty(\mathbb{W} \to \mathbb{R}^n)$$

In this case, the result is a matrix

A.3.2 Coordinate systems on $\mathbb{R}^n$

A coordinate system on $\mathbb{R}^n$ is a set of n functions $\{q_i\}_{i=1}^n$. They define the basis vectors by ∇q_i. If these basis vectors are orthogonal to each other, we say that $\{q_i\}_{i=1}^n$ is a system of *orthogonal* coordinates. Usually (and in particular in this dissertation) orthogonal coordinate systems are preferred. The basis vectors have not unit norm,

$$\overline{\nabla} q_i \, \nabla q_j \;=\; \lambda_i^2 \delta_{i,j}$$

where δ_{ij} is the Kronecker delta. Unit vectors are defined by

$$\boldsymbol{e}_i \stackrel{\text{def}}{=} \lambda_i^{-1} \nabla q_i$$

The 1-forms coincides with covectors on $\mathbb{R}^n$; their set is denoted by $\Lambda^1(\mathbb{R}^n)$ and it has a canonical basis $\{dq_i\}_{i\in\mathbb{N}}$

$$dq_i \stackrel{\text{def}}{=} \lambda_i^{-1} \overline{\nabla} q_i$$

so that one has $dq_i \, \boldsymbol{e}_j \;=\; \delta_{ij}$. A directional derivative can be defined by

$$\partial_i \stackrel{\text{def}}{=} \lambda_i^a \, \overline{\nabla} q_i \, \nabla$$

with a to be determined by the condition

$$\delta_{ij} \stackrel{!}{=} \partial_i q_k \;=\; \lambda_i^a \overline{\nabla} q_i \, \nabla q_k \;=\; \lambda_i^{a+2} \delta_{ik} \;\Longrightarrow\; a = -2$$

so

$$\partial_i \;\equiv\; \lambda_i^{-2} \overline{\nabla} q_i \, \nabla \;\equiv\; \lambda_i^{-1} \overline{\boldsymbol{e}_i}$$

For a basis of unit vectors, the completeness relation holds,

$$\mathbb{I} \;=\; \sum_{i=1}^n \boldsymbol{e}_i \, \overline{\boldsymbol{e}_i}$$

Usually the *cartesian coordinates* are chosen as "fundamental", and different coordinate systems are defined starting from them. The cartesian coordinates, usually denoted by $\{x_i\}_{i=1}^n$, are defined by $\lambda_i = 1$, $\forall\, i \in \{1, \dots n\}$.

A.3.3 A special case: $n = 3$

$\mathbb{R}^3$ is a Lie algebra, the bracket being the cross product,

$$[\![\boldsymbol{v}]\!] \, \boldsymbol{z} = \begin{pmatrix} v^2 z^3 - v^3 z^2 \\ v^3 z^1 - z^3 v^1 \\ v^1 z^2 - v^2 z^2 \end{pmatrix}$$

The bracket is usually denoted by $\boldsymbol{v} \times \boldsymbol{w}$ or also by $\boldsymbol{v} \wedge \boldsymbol{w}$, but we denote it by a double square bracket $[\![\cdot]\!]$ to underline that it's a Lie bracket. And coherently with the notations adopted in the previous sections, we consider the application "bracket with $\boldsymbol{v}$" and denote it by the brackets with a single argument. It is known that the image of this map is explicitly given by a matrix,

$$\boldsymbol{v} \mapsto [\![\boldsymbol{v}]\!] \equiv \begin{pmatrix} 0 & -v^3 & v^2 \\ v^3 & 0 & -v^1 \\ -v^2 & v^1 & 0 \end{pmatrix}$$

The rule for the double cross product reads

$$[\![\boldsymbol{v}]\!]\,[\![\boldsymbol{w}]\!]\,\boldsymbol{z} \;=\; \boldsymbol{w}\overline{\boldsymbol{v}}\boldsymbol{z} \;-\; \overline{\boldsymbol{v}}\boldsymbol{w}\boldsymbol{z} \tag{A.20}$$

In particular, for a unit vector $\boldsymbol{e}$,

$$\boldsymbol{e}^2 \;=\; \boldsymbol{e}\overline{\boldsymbol{e}} \;-\; \mathbb{I} \tag{A.21}$$

With the latter result it is possible to evaluate

$$\exp\big(-t[\![\boldsymbol{e}]\!]\big) \;=\; \sum_{k_0}^{\infty}\frac{t^k}{k!}[\![\boldsymbol{e}]\!]^k \;=\; \sum_{n=0}^{\infty}\frac{t^{2n}}{(2n)!}(\boldsymbol{e}\overline{\boldsymbol{e}}-\mathbb{I})^n \;+\; \sum_{n=0}^{\infty}\frac{t^{2n+1}}{(2n+1)!}(\boldsymbol{e}\overline{\boldsymbol{e}}-\mathbb{I})\,[\![\boldsymbol{e}]\!]$$

It is possible to prove by induction,

$$(\boldsymbol{e}\overline{\boldsymbol{e}}-\mathbb{I})^n \;=\; (-1)^{n+1}\boldsymbol{e}\overline{\boldsymbol{e}} \;+\; (-1)^n\mathbb{I}$$

And so

$$\begin{aligned}\exp\big(-t[\![\boldsymbol{e}]\!]\big) \;&=\; \sum_{n=0}^{\infty}\frac{(-1)^n t^{2n}}{(2n)!}\mathbb{I}) \;+\; \sum_{n=1}^{\infty}\frac{(-1)^{n+1}t^{2n}}{(2n)!}\boldsymbol{e}\overline{\boldsymbol{e}} \;+\; \sum_{n=0}^{\infty}\frac{(-1)^n t^{2n+1}}{(2n+1)!}[\![\boldsymbol{e}]\!] \;=\\ &=\; \cos(t)(\boldsymbol{e}\overline{\boldsymbol{e}}-\mathbb{I}) \;+\; \sin(t)\,[\![\boldsymbol{e}]\!] \;+\; \boldsymbol{e}\overline{\boldsymbol{e}}\end{aligned}$$

In the case of a vector without unit norm, equation (A.21) is replaced by

$$\boldsymbol{e}^2 \;=\; \frac{\boldsymbol{e}\overline{\boldsymbol{e}}}{\overline{\boldsymbol{e}}\boldsymbol{e}} \;-\; \mathbb{I}$$

By inserting the nabla into (A.20), we get the followig rules

$$[\![\boldsymbol{v}]\!]\,[\![\nabla]\!]\,\boldsymbol{z} \;=\; \boldsymbol{z}'\boldsymbol{v} - \overline{\boldsymbol{z}}'\boldsymbol{v} \tag{A.22}$$

$$[\![\nabla]\!]\,[\![\nabla]\!]\,\boldsymbol{z} \;=\; \nabla\,\mathrm{div}\,\boldsymbol{z} - \triangle\boldsymbol{z} \tag{A.23}$$

Cartesian coordinates on $\mathbb{R}^3$ are usually denoted by x, y, z; we use the notations $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ for the corresponding unit vectors.

Often used on $\mathbb{R}^3$ are also the cylindrical coordinates R, τ, z. As we wrote above, different coordinate systems are defined starting from the cartesian one; in this case

$$R = \sqrt{x^2+y^2}, \quad \tan\tau = \frac{y}{x}, \quad z = z$$

The cylindrical unit vectors are denoted by $\boldsymbol{e}_R, \boldsymbol{e}_\tau, \boldsymbol{e}_z$. One has $\lambda_R = 1$, $\lambda_z = 1$ and $\lambda_\tau = R$ so that $\nabla\tau \;=\; R^{-1}\boldsymbol{e}_\tau$.

B

A WKB ansatz for equation (3.20)

In this appendix we describe our efforts to try to solve equation (4.14), by the so-called "WKB approximation". The latter is a method to cope with *singular perturbations*, conceived in the 1930ies to compute the semiclassical limit of Schrödinger equation [82].

We speak of singular perturbation when a differential equation has a small parameter which multiplies the derivative of highest order; when the parameter goes to 0, the dimension of the space of solutions is reduced[1]. In the case of equation (4.14), the small parameter is $q \stackrel{\text{def}}{=} 1/(eB_0)$, where B_0 is a constant and stands here for the "typical" magnetic field strength, so that the limit $eB_0 \to \infty$ means a very strong field, so that the particle sticks to a field line: the usual picture of translation plus gyration is reduced only to translation.

In quantum mechanics, Schrödinger equation is also a singular perturbation (the small parameter being $\hbar$). However, we have two extra difficulties with respect to the original development of the method: we deal with a partial differential equation, which is also a vectorial equation, and the derivative of highest order is highly non-linear (being the convective derivative). In the literature, a WKB theory for partial differential equations was proposed in [62]. Instead, we don't know about a general method for a WKB solution for vectorial equations.

We start by considering the following decomposition

$$\boldsymbol{W} \;=\; \boldsymbol{b}\Phi \;+\; \mu\, e^{\lambda\mathcal{B}/q}\,\mathcal{B}^2\boldsymbol{u} \tag{B.1}$$

where $\boldsymbol{u}$ is a unit norm vector, in the plane perpendicular to $\boldsymbol{b}$, and $\mathcal{B} \stackrel{\text{def}}{=} [\![\boldsymbol{b}]\!]$; μ, λ and Φ are scalar quantities, which for the moment are left free. A projection of equation (4.14) on the direction $\boldsymbol{b}$ of the field, and on the plane perpendicular to it, is achieved by multiplying the equation on the left respectively by $\overline{\boldsymbol{b}}$ and by $\mathcal{B}$. The results are

$$\begin{cases} \overline{\boldsymbol{b}}(\overline{\boldsymbol{W}}\partial_{\overline{\boldsymbol{x}}})\boldsymbol{W} \;=\; 0 \\ q\mathcal{B}(\overline{\boldsymbol{W}}\partial_{\overline{\boldsymbol{x}}})\boldsymbol{W} \;=\; B\mu e^{\lambda\mathcal{B}/q}\mathcal{B}^2\boldsymbol{u} \end{cases} \tag{B.2}$$

[1] A very simple analogy exists with algebraic equations. The solutions of a second order algebraic equation $ax^2 + bx + c \;=\; 0$ are

$$x_\pm \;=\; \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

In the limit $a \to 0$, one finds $x_+ \to -c/b$, which in fact is the solution of $bx \;+\; c \;=\; 0$. However, $x_- \to \infty$.

By injecting the decomposition (B.1) in the r.h.s. of both equations we get

$$\begin{cases} \Phi(\overline{\boldsymbol{b}}\partial_{\overline{\boldsymbol{x}}})\Phi \,+\, \mu(\overline{\boldsymbol{r}}e^{\lambda\mathcal{B}/q}\mathcal{B}^2\partial_{\overline{\boldsymbol{x}}})\Phi \,=\, 0 \\ \Phi\mathcal{B}(\overline{\boldsymbol{W}}\partial_{\overline{\boldsymbol{x}}})\boldsymbol{b} \,-\, e^{\lambda\mathcal{B}/q}\Big(\mathcal{B}\boldsymbol{u}(\overline{\boldsymbol{W}}\partial_{\overline{\boldsymbol{x}}})\mu \,-\, \dfrac{\mu}{q}\mathcal{B}^2\boldsymbol{u}(\overline{\boldsymbol{W}}\partial_{\overline{\boldsymbol{x}}})\lambda\Big) \,=\, \dfrac{B\mu}{q}e^{\lambda\mathcal{B}/q}\mathcal{B}^2\boldsymbol{u} \end{cases} \tag{B.3}$$

To solve the equations above, we expand $\Phi = \sum_n \Phi_n q^n$ and $\mu = \sum_n \mu_n q^n$,

$$\begin{cases} \displaystyle\sum_{n=0}^{N} \Phi_n(\overline{\boldsymbol{b}}\partial_{\overline{\boldsymbol{x}}})\Phi_{N-n} + \mu_n(\overline{\boldsymbol{r}}\mathcal{B}^2 e^{-\lambda\mathcal{B}/q}\partial_{\overline{\boldsymbol{x}}})\Phi_{N-n} = 0 \\ \displaystyle\sum_{n=0}^{N}\Big(\Phi_n\mathcal{B}(\overline{\boldsymbol{W}}_{N-n}\partial_{\overline{\boldsymbol{x}}})\boldsymbol{b} - e^{\lambda\mathcal{B}/q}\mathcal{B}\boldsymbol{u}(\overline{\boldsymbol{W}}_{N-n}\partial_{\overline{\boldsymbol{x}}})\mu_n\Big) + \\ \qquad\qquad\qquad\qquad \displaystyle+\sum_{n=0}^{N+1} e^{\lambda\mathcal{B}/q}\mu_n\mathcal{B}^2\boldsymbol{u}(\overline{\boldsymbol{W}}_{N+1-n}\partial_{\overline{\boldsymbol{x}}})\lambda = B\mu_{N+1}e^{\lambda\mathcal{B}/q}\mathcal{B}^2\boldsymbol{u} \end{cases} \tag{B.4}$$

Can we get a recursive algorithm to compute all the quantities involved? We explicitly write the first orders. The lowest order corresponds to $N = -1$,

$$\mu_0 e^{\lambda\mathcal{B}/q}\big((\overline{\boldsymbol{W}}_0\partial_{\overline{\boldsymbol{x}}})\lambda \,-\, B\big)\mathcal{B}^2\boldsymbol{u} \,=\, 0$$

which is a scalar equation with possible solutions $\mu_0 = 0$ or $(\overline{\boldsymbol{W}}\partial_{\overline{\boldsymbol{x}}})\lambda = B$. Then for $N = 0$ the equations (B.6) read

$$\begin{aligned} (\overline{\boldsymbol{b}}\,\partial_{\overline{\boldsymbol{x}}})\,\frac{\Phi_0^2}{2} \,+\, \mu_0\,\overline{\boldsymbol{r}}\,\mathcal{B}^2\,e^{\lambda\mathcal{B}/q}\,\Phi_0 \,=\, 0 \\ \Phi_0\,\mathcal{B}\,(\overline{\boldsymbol{W}}\,\partial_{\overline{\boldsymbol{x}}})\,\boldsymbol{b} \,-\, \mu_0\,e^{\lambda\mathcal{B}/q}\big(\mathcal{B}\,\boldsymbol{u}(\overline{\boldsymbol{W}}_0\,\partial_{\overline{\boldsymbol{x}}})\mu_0 \,+\, \mathcal{B}^2\,\boldsymbol{u}\,(\overline{\boldsymbol{W}}_1\,\partial_{\overline{\boldsymbol{x}}})\,\lambda\big) + \\ + e^{\lambda\mathcal{B}/q}\mu_1\,\mathcal{B}^2\boldsymbol{u}\,(\overline{\boldsymbol{W}}_0\,\partial_{\overline{\boldsymbol{x}}})\,\lambda \,=\, B\mu_1\,e^{\lambda\mathcal{B}/q}\mathcal{B}^2\,\boldsymbol{u} \end{aligned}$$

which means 3 more constraints: a scalar equation and a vectorial equation in two dimensions. With the first two orders, we have 4 equations in the 7 unknowns: λ, μ_0, μ_1, Φ_0, Φ_1, r_1 and r_2. Apparently we have 3 degrees of freedoms left, but actually there are 3 more constraints coming from the initial data (equation (4.15)). Going to next order $N = 1$,

$$\Phi_0\,(\overline{\boldsymbol{b}}\partial_{\overline{\boldsymbol{x}}})\Phi_1 \,+\, \Phi_1\,(\overline{\boldsymbol{b}}\partial_{\overline{\boldsymbol{x}}})\Phi_0 \,-\, \mu_0\,(\overline{\boldsymbol{r}}\mathcal{B}^2 e^{-\lambda\mathcal{B}/q}\partial_{\overline{\boldsymbol{x}}})\Phi_1 \,-\, \mu_1(\overline{\boldsymbol{r}}\mathcal{B}^2 e^{-\lambda\mathcal{B}/q}\partial_{\overline{\boldsymbol{x}}})\Phi_0 \,=\, 0$$

$$\Phi_0\mathcal{B}(\overline{\boldsymbol{W}}_1\partial_{\overline{\boldsymbol{x}}})\boldsymbol{b} + \Phi_1\mathcal{B}(\overline{\boldsymbol{W}}_0\partial_{\overline{\boldsymbol{x}}})\boldsymbol{b} + e^{\lambda\mathcal{B}/q}\mathcal{B}^2\boldsymbol{u}(\mu_0\overline{\boldsymbol{W}}_2 + \mu_1\overline{\boldsymbol{W}}_1 + \mu_0\overline{\boldsymbol{W}}_0)\partial_{\overline{\boldsymbol{x}}}\lambda = B\mu_2 e^{\lambda\mathcal{B}/q}\mathcal{B}^2\boldsymbol{u}$$

We have added three more equations, but only two unknowns, Φ_2 and μ_2, so the problem seems to be overdetermined. Moreover, the same issue is found at every order: we add three equations but only two unknowns.

Alternatively, instead of (B.1), one may try an expansion like

$$\boldsymbol{W} \,=\, \Phi\boldsymbol{b} \,+\, e^{\lambda\mathcal{B}/q}\mathcal{B}^2\boldsymbol{u}\,, \qquad \Phi \,=\, \sum_{n=0}^{\infty}\Phi_n q^n\,, \qquad \boldsymbol{u} \,=\, \sum_{n=0}^{\infty}\boldsymbol{u}_n q^n$$

by which the system (B.2) gives

$$\sum_{n=0}^{N} \Phi_n(\overline{\boldsymbol{b}}\partial_{\overline{\boldsymbol{x}}})\Phi_{N-n} + (\overline{\boldsymbol{r}}_n\mathcal{B}^2 e^{-\lambda\mathcal{B}/q}\partial_{\overline{\boldsymbol{x}}})\Phi_{N-n} = 0 \tag{B.5}$$

$$\sum_{n=0}^{N}\Big(\Phi_n\mathcal{B}(\overline{\boldsymbol{W}}_{N-n}\partial_{\overline{\boldsymbol{x}}})\boldsymbol{b} - e^{\lambda\mathcal{B}/q}(\overline{\boldsymbol{W}}_{N-n}\partial_{\overline{\boldsymbol{x}}})\mathcal{B}\boldsymbol{u}_n\Big) + \\ + \sum_{n=0}^{N+1} e^{\lambda\mathcal{B}/q}\mathcal{B}^2\boldsymbol{u}_n(\overline{\boldsymbol{W}}_{N+1-n}\partial_{\overline{\boldsymbol{x}}})\lambda = Be^{\lambda\mathcal{B}/q}\mathcal{B}^2\boldsymbol{u}_{N+1} \tag{B.6}$$

The first two orders, for $N=-1$ and $N=0$ give

$$e^{\lambda\mathcal{B}/q}\mathcal{B}^2\boldsymbol{u}_0\big(\overline{\boldsymbol{W}}_0\partial_{\overline{\boldsymbol{x}}}\lambda - B\big) \;=\; 0$$

$$\tfrac{1}{2}(\overline{\boldsymbol{b}}\partial_{\overline{\boldsymbol{x}}})\Phi_0 \;+\; \overline{\boldsymbol{r}}_0\mathcal{B}^2 e^{-\lambda\mathcal{B}/q}\partial_{\overline{\boldsymbol{x}}}\Phi_0 \;=\; 0$$

$$\Phi_0\,\mathcal{B}(\overline{\boldsymbol{W}}_0\partial_{\overline{\boldsymbol{x}}})\boldsymbol{b} \;-\; e^{\lambda\mathcal{B}/q}\Big((\overline{\boldsymbol{W}}_0\partial_{\overline{\boldsymbol{x}}})\mathcal{B}\boldsymbol{u}_0 \;+\; \mathcal{B}^2\boldsymbol{u}_0(\overline{\boldsymbol{W}}_1\partial_{\overline{\boldsymbol{x}}})\lambda \;+\; \mathcal{B}^2\boldsymbol{u}_1\big((\overline{\boldsymbol{W}}_0\partial_{\overline{\boldsymbol{x}}})\lambda - B\big)\Big) \;=\; 0$$

which is a system of 4 equations in the 7 unknowns λ, Φ_0, Φ_1, $\boldsymbol{u}_0$, $\boldsymbol{u}_1$. We recall that we have exactly 3 more (algebraic) equations (4.15), so the system may be solvable. Moreover, at each $N-$th order we add three equations and three unknowns Φ_{N+1} and $\boldsymbol{u}_{N+1}$. However, a number of issues remain:

- The constraints (4.15). How can they be applied? It seems that one needs to write the full series to be able to apply them. In this case, they can't be used to fix some terms of the series itself.
- At each $N-$th order, the unknown $\boldsymbol{W}_{N+1}$ appears into equation (B.6) in a very nonlinear way, namely in the form of the differential operator $\overline{\boldsymbol{W}}_{N+1}\partial_{\overline{\boldsymbol{x}}}$, so it is really difficult to figure a way to invert the equation and build a recursive scheme.

Unfortunately, we have not yet figured out a way to overcome these difficulties. We finally mention that very recently a connection between the WKB method and guiding centre theory has been reconsidered by Burby in [13].

C

The Lagrangian description of dynamics

The Lagrangian descrition is one possible way of building dynamical systems, and a very common one in classical mechanics. Here we follow a bit unusual geometric description of the method, following Souriau [87].

It is unusual in the name, as it is often called "non-canonical Hamiltonian", for instance by the Guiding Centre community, or "pre-symplectic", or "Cartan", or "Poincaré-Cartan". However, we prefer to keep the name "Hamiltonian" for a dynamical system, may it be smooth or not, which possesses a Lie algebraic structure, which is not the case for the present approach (even though a relation with Hamiltonian mechanics exists, as we are going to see).

On the other hand, the name "Lagrangian" usually refers to an analytical description of the dynamics, where the phase space is a smooth manifold, and a dynamical system is built by computing the Euler-Lagrange equations associated to a given "Lagrangian function" (see for instance [51], [5]). This is not the case here.

The Lagrangian description of dynamics requires a smooth structure: the *presymplectic manifold*, which is a manifold $\mathbb{Y}$ with a (differentiable field of) 2-form l such that $\ker(l)$ has constant dimension, and $dl = 0$. Souriau names l the Lagrange form, and that's why we are calling this formalism the "Lagrangian description".

The principle is straighforward: the dynamics is a motion along $\ker(l)$. More explicitly, a vector field C on $\mathbb{Y}$ is determined by $lC = 0$, and then a dynamical system by[1] $\dot{y} = C$, for any coordinate set $\overline{\boldsymbol{y}}$ on $\mathbb{Y}$.

C.1 Relation with Hamiltonian formulation

Starting from a presymplectic manifold, it is possible to build a Hamiltonian dynamical system. Indeed, the manifold[2] $\mathbb{X} \stackrel{\text{def}}{=} \mathbb{Y}/\ker(l)$ has the structure of a *symplectic manifold*. This means that the restriction of l on $\mathbb{X}$, that we will denote by σ, is closed and with zero kernel. This conditions ensures the existence of a bivector π which is the inverse of σ in the sense that, for any vector field X on $\mathbb{X}$,

$$(\sigma\, X)\, \pi \;=\; X$$

[1] the overdot means derivation with respect to some parameter τ, which is not necessarily a physical time

[2] A geometrical note: the manifold $\mathbb{X}$ is called the *leaf* of $\mathbb{Y}$ by the foliation determined by l.

σ is called the symplectic form while π is called the *Poisson byvector*. The Poisson bivector defines a bracket on $\mathcal{C}^\infty(\mathbb{X} \to \mathbb{R})$ by

$$\{f\}g \;=\; dg(df\,\pi)$$

This bracket is alternating because of the alternance of π, while the Jacobi identity is a consequence of $d\sigma = 0$.

The "time" of the Hamiltonian system is the coordinate parametrizing $\ker(l)$, let us call it t. The difference $l - \sigma$ is interpreted as the (wedge) product of dt by a Hamiltonian 1-form $\mathfrak{h}$,

$$l \;=\; \sigma \;-\; \mathfrak{h} \wedge dt$$

If we call $\overline{\boldsymbol{x}}$ a set of coordinates on $\mathbb{X}$, the Hamiltonian system is given by

$$\tfrac{d}{dt}x \;=\; \mathfrak{h}\pi$$

Finally we observe the following: let again $\overline{\boldsymbol{x}}$ be a coordinate set on $\mathbb{X}$, and $\overline{\boldsymbol{y}}$ be a coordinate set on $\mathbb{Y}$. Then we may write $\overline{\boldsymbol{y}} = \{t, \overline{\boldsymbol{x}}\}$. The two-form l is closed in $\mathbb{Y}$, which means that

$$d_{\mathbb{Y}} l \;=\; d\overline{\boldsymbol{y}}\partial_{\overline{\boldsymbol{y}}} l \;=\; 0$$

Instead, σ is closed on $\mathbb{X}$,

$$d_{\mathbb{X}}\sigma \;=\; d\overline{\boldsymbol{x}}\partial_{\overline{\boldsymbol{x}}}\sigma \;=\; 0$$

C.2 Example: Relativistic Mechanics

The dynamics of a charged particle, as described by system (4.1), on the 7-dimensional manifold of coordinates $\{t, \boldsymbol{x}, \boldsymbol{w}\}$ is determined by the following Lagrangian,

$$l_r \;=\; d\overline{\boldsymbol{w}} \wedge d\boldsymbol{x} \;+\; e\,\overline{\boldsymbol{E}}\,d\boldsymbol{x} \wedge dt \;-\; \tfrac{\overline{\boldsymbol{w}}}{m\gamma} d\boldsymbol{w} \wedge dt \;+\; e\,d\overline{\boldsymbol{x}}\big[\!\big[\boldsymbol{B}\big]\!\big]\,d\boldsymbol{x} \tag{C.1}$$

where $\gamma \;=\; \sqrt{1 + \overline{\boldsymbol{w}}\boldsymbol{w}/m}$ (as usual).

A vector field $\mathbf{C}$ on $\{t, \boldsymbol{x}, \boldsymbol{w}\}$ has the general form[3] $\mathbf{C} \;=\; \dot{t}\,\partial_t \;+\; \dot{\boldsymbol{x}}\,\partial_{\overline{\boldsymbol{x}}} \;+\; \dot{\boldsymbol{w}}\,\partial_{\overline{\boldsymbol{w}}}$, and we impose it to lie in $\ker(l)$,

$$\begin{aligned}
0 \;&\stackrel{!}{=}\; l_r\mathbf{C} \\
&=\; d\overline{\boldsymbol{w}}\,\dot{\boldsymbol{x}} \;-\; d\overline{\boldsymbol{x}}\,\dot{\boldsymbol{w}} \;-\; e\,dt\,\overline{\boldsymbol{E}}\,\dot{\boldsymbol{x}} \;+\; e\dot{t}\,\overline{\boldsymbol{E}}\,d\boldsymbol{x} \;+\; \frac{\dot{t}}{m\gamma}\,\overline{\boldsymbol{w}}\,d\boldsymbol{w} \;+\; \frac{dt}{m\gamma}\,\overline{\boldsymbol{w}}\,\dot{\boldsymbol{w}} \;-\; e\,d\overline{\boldsymbol{x}}\big[\!\big[\boldsymbol{B}\big]\!\big]\,\dot{\boldsymbol{x}} \\
&=\; d\overline{\boldsymbol{w}}\left(\dot{\boldsymbol{x}} \;-\; \frac{\dot{t}\boldsymbol{w}}{m\gamma}\right) \;+\; d\overline{\boldsymbol{x}}\left(e\boldsymbol{E}\dot{t} \;-\; e\big[\!\big[\boldsymbol{B}\big]\!\big]\,\dot{\boldsymbol{x}} \;-\; \dot{\boldsymbol{w}}\right) \;+\; dt\left(\frac{\overline{\boldsymbol{w}}\dot{\boldsymbol{w}}}{m\gamma} \;-\; e\overline{\boldsymbol{E}}\dot{\boldsymbol{x}}\right)
\end{aligned}$$

Now we equate each term to zero. If we assume that $\dot{t} > 0$ we can replace the overdot with a derivative by t. This tantamounts to divide everything by $\dot{t}$ and making the substitutions $\dot{x}/\dot{t} \to d\boldsymbol{x}/dt$, $\dot{\boldsymbol{w}}/\dot{t} \to d\boldsymbol{w}/dt$. Then one gets the following 7 equations,

$$\begin{cases}
\dfrac{d\boldsymbol{x}}{dt} \;=\; \dfrac{\boldsymbol{w}}{m\gamma} \\[2ex]
\dfrac{d\boldsymbol{w}}{dt} \;=\; e\boldsymbol{E} \;-\; \dfrac{e}{m\gamma}\big[\!\big[\boldsymbol{B}\big]\!\big]\,\boldsymbol{w} \\[2ex]
-\,e\overline{\boldsymbol{E}}\boldsymbol{w} \;+\; \overline{\boldsymbol{w}}\,\dfrac{d\boldsymbol{w}}{dt} \;=\; 0
\end{cases}$$

[3]Denoting the components of the field by the dotted coordinate name is a shortcut. To be precise we should write $C = C_t\partial_t + \ldots$ and then put $\dot{t} = C_t, \ldots$.

The third equation follows from the second one after multiplying on the left by $\overline{\boldsymbol{w}}$. So we have a 6-dimensional system, as one would expect: we started from a 2-form on a 7-dimensional manifold, but we assumed that $\dim(\ker(l)) = 1$.

If we kept t as a coordinate, we would have a free degree of freedom. If we use it to add to our system the equation $\dot{t} = \gamma$, then the system is parametrized with respect to the socalled "proper time", the physical time measured in a reference frame where the observer is at rest.

Finally, it is possible to show that imposing $dl = 0$ to (C.1), leads to the inhomogeneous Maxwell equations,

$$\partial_{\boldsymbol{x}} \boldsymbol{B} \,=\, 0 \qquad \big[\!\big[\partial_{\overline{\boldsymbol{x}}}\big]\!\big]\, \boldsymbol{E} \,+\, \partial_t \boldsymbol{B} \,=\, 0$$

This is called *Maxwell's principle.*

C.3 The Poincaré-Cartan 1-form

If $\mathbb{X}$ has the structure of a cotangent bundle, then $l = d\varpi$, then ϖ is called the Poincaré-Cartan one-form. This is the case in classical mechanics.

Any chart on the phase space of type $\{\overline{\boldsymbol{p}}, \overline{\boldsymbol{x}}, t\}$ which turns the Poincaré-Cartan 1-form into

$$\varpi \,=\, \overline{\boldsymbol{p}} d\boldsymbol{x} \,-\, H(\overline{\boldsymbol{p}}, \overline{\boldsymbol{q}})\, dt$$

is called a *canonical set of coordinates* (see for instance [5]; this definition is equivalent to the one that was given in the introduction of the dissertation). Instead if one "mess up" ϖ before differentiating, the resulting equations of motion are Hamiltonian but non-canonical. This is what is commonly done among the Guiding Centre community after the works of Littlejohn [16], [56], [15], [23].

The Poincaré-Cartan (henceforth PC) description of electrodynamics requires the potentials $\boldsymbol{A}, \Phi$ instead of the fields[4] $\boldsymbol{E}, \boldsymbol{B}$. In fact it can be shown that $l_r = d\varpi_r$ with

$$\varpi_r \,=\, \big(\overline{\boldsymbol{p}} \,-\, e\,\overline{\boldsymbol{A}}\big)\, d\boldsymbol{x} \,-\, d\big(\varepsilon \,+\, e\,\Phi\big)\, dt$$

The Poincaré-Cartan 1-form introduces a new freedom, for the transformation

$$\varpi_r \to \hat{\varpi}_r \,=\, \varpi \,+\, dg \tag{C.2}$$

gives the same l_r and thus the same dynamics. This is called a gauge tranformation, and indeed it includes as a particular case the gauge invariance of classical electrodynamics: if we choose as g the gauge function of equations (C.2) then $\hat{\varpi}_r$ would have the same functional form as if the hatted potentials of equations (C.2) were directly replaced into ϖ_r.

As a final note, in classical mechanics canonical transformations are also defined by equation (C.2), but the function g is usually denoted by S and called Hamilton's principal function. The equivalence of Gauge transforms and a canonical transforms was discussed by Elsasser in [26].

[4] We recall that, as a consequence of the inhomogeneous Maxwell equations,

$$\boldsymbol{E} \,=\, -\partial_{\overline{\boldsymbol{x}}}\Phi \,-\, \partial_t \boldsymbol{A} \qquad \boldsymbol{B} \,=\, \big[\!\big[\partial_{\overline{\boldsymbol{x}}}\big]\!\big]\, \boldsymbol{A}$$

So defining the fields by the potentials is equivalent to require the inhomogeneous Maxwell equations. However, the potentials are gauge dependent (see section 4.1).

Bibliography

[1] H. Alishah. *KAM Theory, Presymplectic Dynamics and Lie algebroids*. PhD thesis, Universidade Técnica de Lisboa, 2012.

[2] H. Alishah and R. De La Llave. Tracing KAM Tori in Presymplectic Dynamical Systems. *Journal of Dynamics and Differential Equations*, 24(4):685–711, 2012.

[3] V. I. Arnol'd. Proof of a theorem by A. N. Kolmogorov on the persistence of quasi-periodic motions under small perturbations of the Hamiltonian. *Russian Mathematical Survey*, 18(5), 1963.

[4] V. I. Arnol'd. Small Denominators and problems of stability of motion in Classical and Celestial Mechanics. *Russ. Math. Surv.*, 18(6):85, 1963.

[5] V. I. Arnol'd. *Mathematical Methods of Classical Mechanics*. Graduate Texts in Mathematics. Springer-Verlag, New York, 2 edition, 1989.

[6] Barut. *Electrodynamics and Classical Theory of Fields and Particles*. Dover Publications, 2010.

[7] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn. A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method. *Il Nuovo Cimento B (1971-1996)*, 79(2):201–223, 1984.

[8] J. B. Bost. Tores invariants des systèmes dynamiques hamiltoniens. In *Astérisque*, volume 133-134, pages 113–157, 1986.

[9] A. Bounemoura and S. Fischler. The classical KAM theorem for Hamiltonian systems via rational approximations. *Regular and Chaotic Dynamics*, 19(2):251–265, 2014.

[10] A. Bravetti, H. Cruz, and D. Tapias. Contact Hamiltonian Mechanics. *Annals of Physics*, 376:17–39, 2017. arXiv: 1604.08266.

[11] A. J. Brizard and N. Tronko. Equivalent Higher-order Guiding-center Hamiltonian Theories. *arXiv:1606.06534 [physics]*, 2016. arXiv: 1606.06534.

[12] H. W. Broer. KAM theory: The legacy of Kolmogorov's 1954 paper. *Bulletin of the American Mathematical Society*, 41(04):507–522, 2004.

[13] J. W. Burby. Guiding center dynamics as motion on a slow manifold in loop space. *arXiv:1905.04410 [math-ph]*, 2019. arXiv: 1905.04410.

[14] B. Cambon, X. Leoncini, M. Vittot, R. Dumont, and X. Garbet. Chaotic motion of charged particles in toroidal magnetic configurations. *Chaos*, 24(3):033101, 2014. arXiv: 1402.2625.

[15] J. R. Cary and A. J. Brizard. Hamiltonian theory of guiding-center motion. *Rev. Mod. Phys.*, 81(2):693–738, 2009.

[16] J. R. Cary and R. G. Littlejohn. Noncanonical Hamiltonian mechanics and its application to magnetic field line flow. *Annals of Physics*, 151(1):1–34, 1983.

[17] L. Chierchia and C. Koudjinan. V.I. Arnold's "pointwise" KAM Theorem. *arXiv:1908.02523 [math]*, 2019. arXiv: 1908.02523.

[18] J. J. DaCunha and J. M. Davis. A unified Floquet theory for discrete, continuous, and hybrid periodic linear systems. *Journal of Differential Equations*, 251(11):2987–3027, 2011.

[19] R. De La Llave. A tutorial on KAM theory. In *Smooth ergodic Theory & its applications*, volume 69, pages 175–292, Providence, 2001. American Math Society.

[20] R. De La Llave, A. González, A. Jorba, and J. Villanueva. KAM theory without action-angle variables. *Nonlinearity*, 18(2):855–895, 2005.

[21] A. Deprit. Free Rotation of a Rigid Body Studied in the Phase Plane. *American Journal of Physics*, 35(5):424–428, 1967.

[22] C. Di Troia. From charge motion in general magnetic fields to the non perturbative gyrokinetic equation. *Physics of Plasmas*, 22(4):042103, 2015. arXiv: 1501.04353.

[23] C. Di Troia. Non-perturbative guiding center and stochastic gyrocenter transformations: gyro-phase is the Kaluza-Klein 5ˆth dimension also for reconciling General Relativity with Quantum Mechanics. *Journal of Modern Physics*, 09(04):701–752, 2018. arXiv: 1610.00504.

[24] P. Duclos, P. Stovicek, and M. Vittot. Perturbation of an eigen-value from a dense point spectrum: A general Floquet Hamiltonian. *Ann. Inst. H. Poincaré Phys. Théor.*, 71, 1998.

[25] M. S. P. Eastham. *The spectral theory of periodic differential equations*. Scottish Academic Press [distributed by Chatto & Windus, London, 1973.

[26] K. Elsasser. Magnetic field line flow as a Hamiltonian problem. *Plasma Phys. Control. Fusion*, 28(12A):1743–1752, 1986.

[27] D. F. Escande. Stochasticity in classical Hamiltonian systems: universal aspects. *Physics Reports*, 121:165–261, 1985.

[28] D. F. Escande. Contributions of plasma physics to chaos and nonlinear dynamics. *Plasma Phys. Control. Fusion*, 58(11):113001, November 2016. arXiv: 1604.06305.

[29] D. F. Escande and F. Doveil. Renormalization method for computing the threshold of the large-scale stochastic instability in two degrees of freedom Hamiltonian systems. *J Stat Phys*, 26(2):257–284, 1981.

[30] D. F. Escande and F. Doveil. Renormalization method for the onset of stochasticity in a hamiltonian system. *Physics Letters A*, 83(7):307–310, 1981.

[31] A. Fasano and S. Marmi. *Analytical mechanics: an introduction.* Oxford graduate texts. Oxford University Press, Oxford ; New York, 2006. OCLC: ocm61761659.

[32] F. Fassò. The Euler-Poinsot top: A non-commutatively integrable system without global action-angle coordinates. *ZAMP Zeitschrift fur angewandte Mathematik und Physik*, 47(6):953–976, 1996.

[33] C. J. Fewster and K. Rejzner. Algebraic Quantum Field Theory - an introduction. *arXiv:1904.04051 [hep-th, physics:math-ph]*, 2019. arXiv: 1904.04051.

[34] R. P. Feynman, R. Leighton, and M. Sands. *The Feynman Lectures on Physics, Vol. II: The New Millennium Edition: Mainly Electromagnetism and Matter.* Basic Books, New York, 2011.

[35] G. Floquet. Sur les équations différentielles linéaires à coefficients périodiques. *Ann. Sci. École Norm. Sup.*, 12:47–88, 1883.

[36] J. Féjoz. Introduction to KAM Theory, 2011.

[37] J. Féjoz. A proof of the invariant torus theorem of Kolmogorov. *Regul. Chaot. Dyn.*, 17(1):1–5, 2012. arXiv: 1102.0923.

[38] G. Gallavotti. Quasi-integrable mechanical systems. 1986.

[39] C. S. Gardner. Adiabatic invariants of classical periodic systems. *Physical Review*, 115(791), 1959.

[40] G. Giachetta, L. Mangiarotti, and G. Sardanashvily. Differential Geometry of Time-Dependent Mechanics. *arXiv:dg-ga/9702020*, 1997. arXiv: dg-ga/9702020.

[41] A. Giorgilli. Quantitative Methods in Classical Perturbation Theory. In *From Newton to chaos: modern techniques for understanding and coping with chaos in N–body dynamical system*, pages 21–38. Plenum Press, New York, a.e. roy e b.d. steves edition, 1995.

[42] John M. Greene. A method for determining a stochastic transition. *Journal of Mathematical Physics*, 20(6):1183–1201, June 1979.

[43] L. Guillebon and M. Vittot. Gyro-gauge-independent formulation of the guiding-center reduction to arbitrary order in the Larmor radius. *Plasma Phys. Control. Fusion*, 55(10):105001, 2013.

[44] P. Gurfil, A. Elipe, W. Tangren, and M. Efroimsky. The Serret–Andoyer Formalism in Rigid-Body Dynamics: I. Symmetries and Perturbations. *Regular and Chaotic Dynamics*, 12(4):389–425, 2007.

[45] R. S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc. (N.S.)*, 7(1):65–222, 1982.

[46] N. Jacobson. *Lie Algebras*, volume 10 of *Interscience Act in Pure And Applied Mathematics.* John Wiley & Sons, 1962.

[47] A. N. Kolmogorov. On the preservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk, SSSR*, 98:527–530, 1954.

[48] Y. Kosmann-Schwarzbach. Poisson Manifolds, Lie Algebroids, Modular Classes: a Survey. In *Proceedings of the 2007 Midwest Geometry Conference in honor of Thomas P. Branson*, volume 4 of *SIGMA 4*, 2008.

[49] M. Kruskal. Asymptotic Theory of Hamiltonian and other Systems with all Solutions Nearly Periodic. *Journal of Mathematical Physics*, 3(4):806–828, 1962.

[50] L. D. Landau and E. Lifshitz. *The Classical Theory of Fields*, volume 2. Pergamon Press, 1969.

[51] L. D. Landau and E. Lifshitz. *Mechanics*, volume 1. Pergamon Press, 1969.

[52] Y. Li and Y. Yi. Persistence of invariant tori in generalized Hamiltonian systems. *Ergod. Th. Dynam. Sys.*, 22(04), 2002.

[53] R. G. Littlejohn. A guiding center Hamiltonian: A new approach. *Journal of Mathematical Physics*, 20(12):2445–2458, 1979.

[54] R. G. Littlejohn. Hamiltonian formulation of guiding center motion. *The Physics of Fluids*, 24(9):1730–1749, 1981.

[55] R. G. Littlejohn. Hamiltonian perturbation theory in noncanonical coordinates. *Journal of Mathematical Physics*, 23(5):742–747, 1982.

[56] R. G. Littlejohn. Variational principles of guiding centre motion. *Journal of Plasma Physics*, 29(1):111–125, 1983.

[57] R. S. MacKay. A renormalization approach to invariant circles in area-preserving maps. *Physica D: Nonlinear Phenomena*, 7(1):283–300, May 1983.

[58] J. E Marsden, R Montgomery, P. J Morrison, and W. B Thompson. Covariant poisson brackets for classical fields. *Annals of Physics*, 169(1):29–47, June 1986.

[59] J. E. Marsden and T. S. Ratiu. *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems*, volume 17. Springer Science & Business Media, 2013.

[60] H. Marthinsen and B. Owren. Geometric integration of non-autonomous Hamiltonian problems. *arXiv:1409.5058 [math]*, 2014. arXiv: 1409.5058.

[61] M. Materassi and P. J Morrison. Metriplectic torque for rotation control of a rigid body. *(preprint)*, page 8, 2018.

[62] R. Miura and M. Kruskal. Application of a Non Linear WKB Method to the Korteweg-DeVries Equation. *SIAM Journal on Applied Mathematics*, 26(2):376–395, 1974.

[63] P. J. Morrison. The Maxwell-Vlasov equations as a continuous hamiltonian system. *Physics Letters. A*, 80(5/6):383–386, 1980.

[64] P. J. Morrison. Poisson brackets for fluid and plasmas. In *Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems*, volume 88, page 36, New York, 1982. American Institute of Physics.

[65] P. J. Morrison. Hamiltonian description of the ideal fluid. *Rev. Mod. Phys.*, 70(2):55, 1998.

[66] P. J. Morrison and J. Vanneste. Weakly nonlinear dynamics in noncanonical Hamiltonian systems with applications to fluids and plasmas. *Annals of Physics*, 368:117–147, 2016.

[67] P.J. Morrison. Hamiltonian field description of the one-dimensional Poisson-Vlasov equations. Technical Report PPPL-1788, 6423520, Princeton University, 1981.

[68] J. Moser. On Invariant Curves of Area-Preserving Mappings of an Annulus. *Nachr. Akad. Wiss. Gottingen*, II:1–20, 1962.

[69] J. Moser. A rapidly convergent iteration method and non-linear differential equations = II. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, 20(3):499–535, 1966.

[70] J. Moser. A rapidly convergent iteration method and non-linear partial differential equations - I. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, 20(2):265–315, 1966.

[71] J. Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.*, 169(1):136–176, 1967.

[72] A. Neishtadt and A. Artemyev. Hamiltonian for guiding center motion: symplectic structure approach. *arXiv:1905.12316 [math-ph, physics:physics]*, 2019. arXiv: 1905.12316.

[73] T. Northrop. *The Adiabatic Motion of Charged Particles.* Interscience Publishers, Inc, 1963.

[74] T. G. Northrop. Adiabatic charged-particle motion. *Reviews of Geophysics*, 1(3):283–304, 1963.

[75] T. G. Northrop. Adiabatic Theory of Charged Particle Motion. In B. McCormac, editor, *Radiation Trapped in the Earth's Magnetic Field*, Astrophysics and Space Science Library, pages 26–44. Springer Netherlands, 1966.

[76] Michal Pavelka, Václav Klika, Oğul Esen, and Miroslav Grmela. A hierarchy of Poisson brackets in non-equilibrium thermodynamics. *Physica D: Nonlinear Phenomena*, 335:54–69, November 2016.

[77] G. Pinzari and L. Chierchia. Properly-degenerate KAM theory (following V. I. Arnold). *Discrete and Continuous Dynamical Systems - Series S*, 3(4):545–578, 2010.

[78] J. Pöschel. Integrability of hamiltonian systems on cantor sets. *Communications on Pure and Applied Mathematics*, 35(5):653–696, 1982.

[79] X. Raymond. Simple Nash-Moser Implicit Function Theorem. *L'Einsegnement Mathematique*, 35, 1989.

[80] M. Reed and B. Simon. *Functional Analysis, Volume 1.* Academic Press, 1981.

[81] H. Rüssmann. Invariant tori in non-degenerate nearly integrable Hamiltonian systems. *Regular and Chaotic Dynamics*, 6, 2001.

[82] J. J. Sakurai. *Modern Quantum Mechanics*. Cambridge University Press, 2nd edition, 2017.

[83] D. Salamon. The Kolmogorov-Arnold-Moser theorem. *Math. Phys. Electron. J*, 2004.

[84] D. Salamon and E. Zehnder. KAM theory in configuration space. *Commentarii Mathematici Helvetici*, 64(1):84–132, 1989.

[85] N. Sato and M. Yamada. Local Representation and Construction of Beltrami Fields. *Physica D: Nonlinear Phenomena*, 391:8–16, 2019. arXiv: 1809.03136.

[86] B. D. Scott. Gyrokinetic Field Theory as a Gauge Transform or: gyrokinetic theory without Lie transforms. *arXiv:1708.06265 [physics]*, 2017.

[87] J. M. Souriau. *Structure of Dynamical Systems*. Birkhauser, 1997.

[88] D. P. Stern. Euler Potentials. *American Journal of Physics*, 38(4):494–501, 1970.

[89] P. Tantalo. *Geometric Phases for Free Rigid Body with Variable Inertia Tensor*. PhD thesis, University of California Santa Cruz, 1993.

[90] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*. American Math Society, 2012.

[91] T. F. Viscondi, I. L. Caldas, and P. J. Morrison. Beatification: Flattening the Poisson bracket for two-dimensional fluid and plasma theories. *Physics of Plasmas*, 24(3):032102, 2017.

[92] M. Vittot. Perturbation Theory and Control in Classical or Quantum Mechanics by an Inversion Formula. *Journal of Physics A: Mathematical and General*, 37(24):6337–6357, 2004. arXiv: math-ph/0303051.

[93] M. Vittot. *Une version "Algèbre de Lie" de la Théorie de Perturbation Hamiltonienne Classique ou Quantique et du Contrôle Hamiltonien*. HDR Thesis, Aix-Marseille Université, Marseille, 2007.

[94] J.-C. Yoccoz. An Introduction To Small Divisors Problems. In M. Waldschmidt, P. Moussa, J.-M. Luck, and C. Itzykson, editors, *From Number Theory to Physics*, pages 659–679. Springer, Berlin, Heidelberg, 1992.

[95] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems, I. *Communications on Pure and Applied Mathematics*, 28(1):91–140, 1975.

[96] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems, II. *Communications on Pure and Applied Mathematics*, 1976.

Abstract

The Hamiltonian formulation of classical mechanics reveals an underlying Lie algebraic structure which is a key element for developing an efficient perturbation theory. But Lie structures are met in a wider class of dynamical systems, called Poisson systems; some examples are, among others, fluid dynamics, electrodynamics, kinetic theory. In the first part of this thesis, we propose a purely algebraic approach to classical perturbation theory to extend its scope to any Poisson system. In this method, introduced in [Vittot, 2004], a (Lie) transform allows to split the perturbation into a term preserving the unperturbed flow, and a smaller correction, quadratic in the original perturbation strength.

The second part of the dissertation is about the dynamics of a Throbbing Top (a non-autonomous Rigid Body). Being probably the most basic example of Poisson system, the Top was a natural choice for an application of our theory. We consider first a symmetric Top with periodically dependent momenta of inertia; by introducing a suitable set of coordinates, the system is reduced to a nearly classical description; indeed we show that our theorem applies and reproduce the KAM theorem of classical mechanics. Then we switch to a non symmetric Top with non-periodically fluctuating momenta of inertia: in this case we study for which conditions the static trajectories give a good approximation to those of the non-autonomous system.

In the third part of this work we study the dynamics of a magnetically confined particle. In this case the unperturbed flow is the dynamics in an arbitrary given electromagnetic field; then by perturbation theory one may reduce the dimensionality of the dynamics, or study the retroaction of the particle on the field. However, providing an efficient description of the unperturbed flow is a formidable task, related to the long-standing issue of Guiding Centre Theory in plasma physics. Recently a novel relativistic and non-perturbative approach to Guiding Centre theory has been proposed [Di Troia, 2018]. We derive the equations of motion and their Poisson structure in this description.

Résumé

La formulation Hamiltonienne de la mécanique classique révèle une structure algébrique de Lie sous-jacente qui est un élément clé pour développer une théorie de perturbation efficace. Mais on trouve des structures de Lie dans une classe plus grande de systèmes dynamiques, appelé systèmes de Poisson; certains exemples sont, entre autres, la dynamique des fluides, l'électrodynamique, la théorie cinétique. Dans la première partie de cette thèse, on propose une approche purement algébrique à la théorie classique des perturbations, qui s'applique donc à tout les système de Poisson. Dans cette méthode, introduite dans [Vittot, 2004], une transformation (de Lie) permet de diviser la perturbation en un terme préservant le flot non perturbé, et une correction plus petite, quadratique par rapport à la perturbation originale.

Dans la deuxième partie de la thèse on considere la dynamique d'une Toupie Pulsante (un Corps Rigide non autonome). S'agissant probablement de l'exemple le plus basique de système de Poisson, la Toupie était un choix naturel pour une application de notre théorie. Nous considérons d'abord une toupie symétrique àvec des moments d'inertie qui oscille periodiquement. En introduisant des coordonnées appropriées, le système est réduit à un systeme presque classique: en effet, on montre que notre théorème s'applique et reproduit le théorème de KAM de la mécanique classique. Puis on considere une Toupie non symétrique avec moments d'inertie qui presentent des fluctuations quelconques: dans ce cas, on etudie sous quelles conditions les trajectoires du systeme sont proches de celle du système statique.

Dans la troisième partie de ce mémoire, on étudie la dynamique d'une particule chargée confinée magnétiquement. Dans ce cas le flot non perturbé est la dynamique dans un champ électromagnétique donné arbitraire. Alors par la théorie des perturbations on peut réduire la dimensionnalité de la dynamique, ou étudier la rétroaction de la particule sur le champ. Cependant, fournir une description du flot non perturbé est une tâche redoutable, liée à la question de longue date de la théorie du centre-guide en physique des plasmas. Récemment une version relativiste et non perturbative de la théorie des centres guides a été proposée [Di Troia, 2018]. Nous dérivons les équations du mouvement et leur structure de Poisson dans cette description.