

CAN ONE HEAR WHISTLER WAVES?

CHRISTOPHE CHEVERRY

ABSTRACT. The aim of this article is to provide a mathematical framework giving access to a better understanding of whistler-mode chorus emissions in space plasmas. There is presently a general agreement [10, 18, 22, 24, 26] that the emissions of whistler waves involve a mechanism of wave-particle interaction which can be described in the framework of the relativistic Vlasov-Maxwell equations, with a penalized skew-symmetric term where the inhomogeneity of the magnetic field plays an essential part. The description of the related phenomena is achieved in two stages. The first provides a new approach allowing to extend in longer times the classical insights on fast rotating fluids [6, 8, 11, 12, 14, 23]. The second is based on a study of oscillatory integrals implying special phases.

Keywords. Rotating fluids; dynamical and hamiltonian systems; non linear geometric optics; WKB calculus; oscillatory integrals; plasma physics; relativistic Vlasov-Maxwell equations; whistler waves.

1. INTRODUCTION

Whistlers [1] are very low frequency (VLF) electromagnetic (radio) waves generated by lightning, described in 1953 by Storey [25] as the sound of the *dawn chorus*. It was then suggested by Dungey [10] in 1963 - see also [22, 24] - that the natural or artificial emissions of whistler waves may be due to the oscillations in density of the charged particles trapped in the **Van Allen belts** [4]. Since, many works [20] have been put forth to explain what happens. Arising probably in connection with progresses in satellite observations (whistler waves are regularly detected by Cluster and Themis during substorms), there is now a number of contributions on this subject, coming from various scientific fields, giving rise to different interpretations of the phenomena. The recent PhD thesis in applied mathematics of **S. Le Bourdieu** [18], in astrophysics of **A. Tenerani** [26] and in plasma physics of **P. Pfannmöller** [21], as well as many references cited therein, furnish interesting investigations by means of observations, theoretical modelings and numerical studies.

The evolution of whistler waves, in time $\tau \in \mathbb{R}_+$ (with $\tau \sim 1$ representing 10^3 s), in position $x \in \mathbb{R}^3$, and in velocity $v \in \mathbb{R}^3$, is governed by the relativistic Vlasov-Maxwell system. The equations are presented below directly in dimensionless variables. The reader is referred to Paragraph 3.1.1 for a definition of ε , ϵ_r and κ , and a justification of the following model:

$$(1.1) \quad \begin{cases} \partial_\tau f + \varepsilon^{-1} v \cdot \nabla_x f + \varepsilon^{-2} (v \times \tilde{B}) \cdot \nabla_p f + (E + v \times B) \cdot \nabla_p f = 0, & f|_{\tau=0} \equiv f, \\ \partial_\tau (\epsilon_r E) - \varepsilon^{-1} \operatorname{curl} B = -\varepsilon^{-1} \kappa j, & \operatorname{div} (\epsilon_r E) = \kappa \rho, & E|_{\tau=0} \equiv 0, \\ \partial_\tau B + \varepsilon^{-1} \operatorname{curl} E = 0, & \operatorname{div} B = 0, & B|_{\tau=0} \equiv 0. \end{cases}$$

Date: 15/11/2013.

It is wellknown that the particles gyrate at frequencies much greater than the typical rate of change of the magnetic field. This means introducing a small parameter, the electron gyrofrequency, of size $\varepsilon \sim 10^{-5}$. The symbol ϵ_r ($\neq \varepsilon!$) is the relative permittivity (space plasmas are inhomogeneous media). The velocity v is limited by the speed of light, 1 in dimensionless variables. It is linked to the relativistic impulsion p through the relations:

$$(1.2) \quad v(p) = p (1 + |p|^2)^{-1/2}, \quad p(v) = v (1 - |v|^2)^{-1/2}.$$

The electric current density j and the electric charge density ρ are given by:

$$(1.3) \quad j := \int_{\mathbb{R}^3} f v dp, \quad \rho := \int_{\mathbb{R}^3} f dp, \quad dp := (1 - |v|^2)^{-5/2} dv.$$

In (1.1), the letter \tilde{B} stands for the dipole model of the Earth's magnetic field. It bears noting that $\tilde{B}(x)$ is a non-constant function given by formulas (2.7) and (3.2). In fact, the contribution $v \times \tilde{B}(x)$ reflects how, at position x , the external geomagnetic field $\tilde{B}(x)$ deflects the charged particles through the Lorentz force. The physical quantities to determine are the distribution function f (which indicates the particle density), together with the self-consistent electromagnetic field (E, B) .

Whistler waves are inherently a matter of creation and propagation of oscillations. These two stages do not occur on the same period. The relevant oscillations are developed at the level of the Vlasov part after a long preparation, from $t := \varepsilon^{-1} \tau \sim \varepsilon$ to $t \sim 1$ with fast rotations, and then from $t \sim 1$ to $t \sim \varepsilon^{-1}$ with again fast rotations (of a different kind). During this maturation process, the features of \tilde{B} have a full impact. The oscillations thus generated inherit special structures, stable in the sense that they are not affected by the self-consistent electromagnetic field (E, B) . They are viewed in long times $\tau \sim 1$ as source terms by the Maxwell equations. They are handled and propagated there as singularities. Then, what matters are the relative positions of the wave-fronts induced by the oscillations and the directions selected by the characteristic variety of Maxwell equations.

This article is organized in accordance with the preceding guidelines. The Part 2 is devoted to the study of the dynamical system issued from the Vlasov equation. It is followed by Part 3 focusing on particular Lagrangian distributions, build at the level of the Maxwell equation with the oscillations exhibited in Part 2.

1.1. Analysis of the Vlasov equation. The Vlasov part of (1.1) is the first equation of (1.1), the one on f . It controls the dynamics of the particles. Assuming that the function $(E, B)(\varepsilon, \tau, x)$ is known, it can be solved by following the characteristic flow:

$$(1.4) \quad f_\varepsilon(\tau, x, v) = f(x_\varepsilon(-\tau, x, v), v_\varepsilon(-\tau, x, v)), \quad f_\varepsilon(0, x, v) := f(x, v),$$

where $(x_\varepsilon, v_\varepsilon)(\tau, x, v) \in \mathbb{R}^3 \times \mathbb{R}^3$ is obtained through:

$$(1.5) \quad \begin{cases} \partial_\tau x_\varepsilon = \varepsilon^{-1} v_\varepsilon, & x_\varepsilon(0, x, v) = x, \\ \partial_\tau v_\varepsilon = \varepsilon^{-2} c(|v_\varepsilon|) v_\varepsilon \times \tilde{B} + \tilde{N}(v_\varepsilon, E, B), & v_\varepsilon(0, x, v) = v, \end{cases}$$

with $|v| < 1$ and coefficients adjusted by taking into account (1.2):

$$(1.6) \quad c(v) := (1 - |v|^2)^{1/2}, \quad \tilde{N}(v, E, B) := (1 - |v|^2)^{1/2} (E - (E \cdot v) v + v \times B).$$

◇ *Physical observations.* The trajectories associated with (1.5) are the combination of three motions, see [2] for illustrative pictures and basic heuristical descriptions:

(P1) *Fast times* $t := \varepsilon^{-1} \tau \sim \varepsilon \sim 10^{-3} s$: The particles oscillate at the cyclotron frequency around the field lines. They move approximately in circles, whose centers are along the field lines. These circles are contained in the orthogonal planes to the field lines, and they have Larmor radius $\sim \varepsilon \ll 1$ (hypothesis of cold plasma);

(P2) *Intermediate times* $t \sim \varepsilon \sim 1 s$: The centers follow the field lines. The particles bounce back and forth between two mirror points located just above the Earth's poles, crossing each time the equatorial plane;

(P3) *Long times* $t \sim \varepsilon^{-1} \sim 10^3 s$ (or $\tau \sim 1$) : A longitudinal drift (initiation of a rotation around the Earth) becomes perceptible. It is associated with the creation of a ring current. Positive particles go towards west, negative towards east. ◇

These motions have long been of interest to both mathematicians and physicists. Their description raises difficulties known under the name of the *Störmer problem*. This subject has been intensively studied through the hamiltonian approach, with tools issued from KAM theory. There are actually continuations in the direction of chaos. To have an overview, we refer to the survey articles [4, 9]. The whole picture gives rise to toroidal regions surrounded the earth, the so-called *Van-Allen belts*, where the particles are trapped and oscillate with a lifetime from a few hours to 10 years.

The hamiltonian approach is more concerned with the Lagrangian point of view. It does not offer much insight about the variations which can occur (at fixed times and at given spatial scales) when passing from a trajectory to one another. Such information is encoded elsewhere, in the Eulerian concept of what is the phase of an oscillation. Now, to capture the phases generated by (1.5) during long times $\tau \sim 1$, it is necessary to conveniently separate the scales. This means to revisit the Störmer problem by the way of modern averaging methods. It is this program that is achieved in Part 2.

Section 2.1 starts with the Assumption 2.1 on \tilde{B} , allowing to straighten out the field lines in an orthogonal way. Once that is accomplished, and as long as $t \sim 1$, the asymptotic analysis of (1.5) relies on adaptations of the filtering method [6, 23], in the spirit of what has been done in [12]. The situation during intermediate times $t \sim 1$ is similar to the usual reports in rotating fluids [6, 11, 12, 14, 23]. It is a starting point.

Changing t into τ produces new problematical terms of size ε^{-1} . To partially absorb them, a rectification procedure is adopted in Section 2.2. This needs to consider the homological equation (2.27). This amounts to solve the modulation equation governing the evolution during intermediate times. There is a corresponding flow map Ξ . In a convenient choice of coordinates, the phase space (of dimension 6) is foliated by planes which are globally invariant under the action of Ξ . Each of these planes is equipped for Ξ with an energy function \mathcal{H} , the *reduced hamiltonian* (Definition 2.6). When the mean curvature of the surfaces orthogonal to the field lines changes its sign along the field lines (this can take place only if \tilde{B} undergoes variations in directions), as it is the case (when passing the equatorial plane) for the geomagnetic field, the function \mathcal{H} is a potential well.

It follows that the trajectories are periodic. Here is an explanation for the step (P2). Here lies also a key to derive uniform estimates in t (eliminating, for instance, problems of secular growth). We refer the reader to the paragraphs 2.2.1 for more comments.

After rectification, there remains terms of size ε^{-1} but they act with a sort of decoupling. In Section 2.3, the discussion falls within the field of supercritical WKB analysis. This includes formal computations, beginning in Paragraph 2.3.2, with the derivation of the long time modulation equation (2.78). It is only then that a precise description of the step (P3) becomes accessible. The induction and the justification (with a loss of precision by a factor ε^{-2}) come into play after, in Paragraphs 2.3.3 and 2.3.4. In the end of Section 2.3, see Paragraph 2.3.5, special emphasis is given to the important notions of secondary phase ψ_s and of principal phase $\psi_{p\varepsilon}$.

The statement below, proven in Paragraph 2.3.6, highlights some key results of Part 2.

Theorem 1. *[global and long time modeling of radiation belt dynamics] Assume that the self-consistent electromagnetic field $(E, B)(\varepsilon, \tau, x) \in \mathcal{C}^\infty([0, 1] \times \mathbb{R}_+ \times \mathbb{R}^3; \mathbb{R}^6)$ is a known given function. Consider the centred dipole model (2.7) for the geomagnetic field \tilde{B}_e . There is a time $\mathcal{T} \in \mathbb{R}_+^*$ and a domain \mathcal{EBelt} of the phase space $(x, v) \in \mathbb{R}^3 \times \mathbb{R}^3$, which is relatively compact in $(\mathbb{R}^3 \setminus \{0\})^2$, such that :*

- 1) *the projection of \mathcal{EBelt} onto the x coordinate is a toroidal region of \mathbb{R}^3 which can be made conform to the shape of the Van Allen belt, designated by \mathcal{Belt} ;*
- 2) *for all $(x, v) \in \mathcal{EBelt}$, the solution $(x_\varepsilon, v_\varepsilon)(\tau, x, v)$ to (1.5) remains for all $\tau \in [0, \mathcal{T}]$ in a neighbourhood of \mathcal{EBelt} , that is compact in $(\mathbb{R}^3 \setminus \{0\})^2$;*
- 3) *there are profiles $X_j(\tau, x, v, \vartheta, \theta)$ and $V_j(\tau, x, v, \vartheta, \theta)$ in $\mathcal{C}^\infty([0, \mathcal{T}] \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{T}^2; \mathbb{R}^3)$ with $j \in \mathbb{N}$ such that, for all $N \in \mathbb{N}^*$, the asymptotic behaviour when ε goes to zero of the solution $(x_\varepsilon, v_\varepsilon)(\tau, x, v)$ to (1.5) can be described with a precision $O(\varepsilon^{N-2})$ in $L^\infty([0, \mathcal{T}] \times \mathcal{EBelt}; \mathbb{R}^6)$ through the following multiscale and multiphase expansion:*

$$(1.7) \quad \begin{pmatrix} x_\varepsilon \\ v_\varepsilon \end{pmatrix} (\tau, x, v) \underset{\varepsilon \rightarrow 0}{\sim} \sum_{j=0}^N \varepsilon^j \begin{pmatrix} X_j \\ V_j \end{pmatrix} \left(\tau, x, v, \frac{\psi_s^e(\tau, x, v)}{\varepsilon}, \frac{\psi_{p\varepsilon}^e(\tau, x, v)}{\varepsilon^2} \right).$$

The secondary phase $\psi_s^e(\tau, x, v)$ does not depend on (E, B) . It is specified in Definition 2.6, together with (2.109). The principal phase $\psi_{p\varepsilon}^e(\tau, x, v)$ is an oscillation of small amplitude:

$$(1.8) \quad \psi_{p\varepsilon}^e(\tau, x, v) := \psi_{p0}^e(\tau, x, v) + \varepsilon \psi_{p1}^e \left(\tau, x, v, \frac{\psi_s^e(\tau, x, v)}{\varepsilon} \right).$$

Explicit formulas for the scalar functions $\psi_{p0}^e(\tau, x, v)$ and $\psi_{p1}^e(\tau, x, v, \vartheta)$ are provided in the Definitions 2.6 and 2.7. For all $k \in \mathbb{N}$, the \mathcal{C}^k -norms of the profiles (X_j, V_j) with $j \in \mathbb{N}$, of ψ_{p0}^e and of ψ_{p1}^e can be controlled by the \mathcal{C}^k -norm of (E, B) .

The notions of extended Van Allen belt \mathcal{EBelt} and of Van Allen belt \mathcal{Belt} will be clarified more specifically later on, see (2.10), (2.55) and Paragraph 2.3.6 to get an overview. Both phases ψ_s^e and $\psi_{p\varepsilon}^e$ emerge already during intermediate times $t \sim 1$. But it is not until long times $\tau \sim 1$ that these objects operate as indicated in (2.94) and (1.8). The ingredients ψ_s^e and $\psi_{p\varepsilon}^e$ are the cornerstones of understanding whistler waves.

1.2. Analysis of the Maxwell part. The Maxwell part of (1.1) consists of the two last equations in (1.1). It gives access to the self-consistent electromagnetic field (E, B) , which is in fact completely determined by the knowledge of the electric current density j . The approach developed in Part 3 is based on a common sense principle. The impact of solar flares is first of all reflected in the input of localized initial data f . This generates exact solutions $(f_\varepsilon^e, E_\varepsilon^e, B_\varepsilon^e)$ to (1.1). The oscillating wave front of f_ε^e is forced by the fast dynamics contained in the Vlasov part, that is by (1.5). In so doing, the inhomogeneities of the penalization term $\tilde{B}_\varepsilon(\cdot)$ play a central part, whereas $(E_\varepsilon^e, B_\varepsilon^e)$ has little effect.

It follows that the functions f_ε^e can be compared to the expressions f_ε obtained by solving the Vlasov equations with $(E, B) = (0, 0)$. The influence of the coupling does not affect the main high frequency aspects. From this point of view, the coupling is secondary or it comes after, during very long times $\tau \gg 1$. With this in mind, apply Theorem 1 in the case $(E, B) = (0, 0)$, to get some distribution function f_ε and therefore, through (1.3), some electric current density j_ε . Then, consider the Cauchy problem built with the two last equations of (1.1), where the source term j is the known function j_ε . Interesting information about whistler waves can be extracted just by looking at the resulting solution $(E_\varepsilon, B_\varepsilon)$. Whistler waves are audio frequency electromagnetic waves produced by lightning, detected through *spectrograms*, giving a visual representation of the spectrum of frequencies in the signal as they vary with time.

◇ *Physical observations.* It is worth pointing out that the spectrograms recorded share the following main characteristics:

(C1) A broad (VLF) frequency range (from 100 Hz to 10 kHz) is implied. Within this range, chorus waves can occur in distinct frequency bands;

(C2) The precise description of a chorus element can show rising frequencies, see the sample produced in figure 1 p. 623 of [20].

(C3) Whistler-mode chorus emissions consist of a succession of discrete elements (chorus elements), occurring with a repetition period of about or less than 1 second, marked by light-coloured vertical lines in the related pictures, see for instance [1]. ◇

The present analysis is based on this concrete information. From (C1), we will deduce that some semiclassical regime is at stake, with an adequate dispersion relation. From (C2), we will infer the presence of harmonics, indicating nonlinear effects. But our foremost concern is to explain (C3). Though the generation mechanism of chorus alluded in (C3) has been extensively studied in the past, it is not yet well understood. Please refer to [20] for recent results obtained during investigation in the Earth's magnetosphere.

Part 3 is aimed to propose a mathematical interpretation of what happens concerning (C3). Subsection 3.1 is devoted to the modelling of whistler waves. It begins in Paragraph 3.1.1 with the nondimensionalization of the original equations, those considered in [18, 21, 26]. Paragraph 3.1.2 outlines the way whistler waves are usually tackled, in mathematical terms. It is followed in Paragraph 3.1.3 by a presentation of a quite different perspective, based on the propagation of singularities.

On the one hand, the fast dynamics underlying the system of ODEs (1.5) serve to force the oscillating structures of the electric current density j_ε . In accordance with (C1), they select a particular frequency range, as well as very specific geometrical features. On the other hand, as explained in Paragraph 3.1.4, the frequencies thus micolocalized are managed by a **dispersion relation** (Figure 10 of [3]) which must take into account the presence of a non-homogeneous magnetized plasma. Combining these two aspects leads to a class of oscillatory integrals, introduced in Paragraph 3.1.5, involving phases $\Psi(\pm, q; t, x; s, y, \varsigma, \xi)$ whose principal characteristics can be viewed at the level of (3.27).

At this stage, the main question to address is to obtain stationary phase expansions. As noted in the introduction of Section 3.2, this relies on a study of the (interior and boundary) critical points associated with the scalar function $\Psi(\pm, q; t, x; \cdot)$. The behaviour of Ψ with respect to the time variable s has a remarkable property. Indeed, the preliminary estimates of Paragraph 3.2.1 reveals that $\Psi(\pm, q; t, x; \cdot, y, \varsigma, \xi)$ is composed of a linear part in s (with a nonzero coefficient) plus some periodic part.

From that, it becomes possible to show (at least for sufficiently large times t) the absence of boundary critical points (Paragraph 3.2.2), to sort the pertinent harmonics $q \in \mathbb{Z}$ (see Paragraph 3.2.3), and to focus the discussion on the spatial phase $\Psi(\pm, q; t, x; s, y, \varsigma, \pi/2)$ by pointing out that the signals can only emanate from the mirror points. Then, a distinction must be drawn between two different situations. In the end, what matters are the respective positions of two quantities, which are relevant from a physics point of view. The first quantity is the maximum amplitudes b_M of the geomagnetic field along trajectories (see Paragraph 3.2.1 for a precise definition); the second quantity is the threshold limit λ_M issued from the dispersion relation (see Assumption 3.1).

The case $c(w) b_M < \lambda_M$ is investigated in Paragraph 3.2.5. The other case $c(w) b_M > \lambda_M$ is examined in Paragraph 3.2.6. In both situations, the aim is to identify the successive times s^j (with $j \in \mathbb{N}$) when the signals composing a chorus element are emitted. Depending on the situation, the episodes of whistler-mode chorus emissions may be shorter or longer. But, they all have one thing in common. As reflected in (C3), the signals appear again and again, over consecutive periods of almost equal length. This aspect is roughly highlighted below, in Theorem 2. It is present in the background of Propositions 3.1 and 3.2, which contain the complete and rigorous statements.

Theorem 2 (whistler-mode chorus emissions). *The phase Ψ involve several critical points which can be classified by order of appearance s^j , with $s^j \leq s^{j+1}$, the majority of which are separated by a uniform interval of time.*

Let us close this introduction with two brief concluding remarks. Beyond the case of earth's magnetic field, the analysis of Part 2 should be the entry point into many new, interesting and challenging problems in physics, since it furnishes a systematic method to study very long-time effects induced by the cross product with a non constant vector field. Similarly, the Part 3 must be understood as a preliminary step towards more refined studies. There remain a number of important mathematical issues that need to be addressed, about the geometry of propagation, the effects of coupling, the stability estimates, etc.

TABLE OF CONTENTS

1. Introduction	1
1.1. Analysis of the Vlasov equation.	2
1.2. Analysis of the Maxwell part.	5
2. Multiscale analysis of the Vlasov part	8
2.1. The averaging method during intermediate times $t \sim 1$.	8
2.1.1. Assumptions on the external magnetic field.	9
2.1.2. Straightening out the field lines.	11
2.1.3. Filtering out the main oscillations.	12
2.2. The approach to reach long times $t \sim \varepsilon^{-1}$.	13
2.2.1. Preliminaries and guidelines.	13
2.2.2. Approximate rectification of the vector field \mathcal{A} .	15
2.2.3. The autonomous change of variables $\Xi_1(\mathfrak{z})$.	16
2.2.4. The reduced hamiltonian.	17
2.2.5. The autonomous change of variables $\Xi_2(\mathfrak{z})$.	21
2.2.6. The time dependent change of variables $\Xi_3(t, \mathfrak{z})$.	22
2.3. The long time WKB calculus.	24
2.3.1. Initialization.	24
2.3.2. Modulation equations.	25
2.3.3. The induction.	27
2.3.4. Justification of the WKB approach.	28
2.3.5. Description of the phases in long times.	28
2.3.6. Proof of Theorem 1.	31
3. Propagation of singularities for the Maxwell part	32
3.1. From Vlasov-Maxwell equations to oscillatory integrals.	32
3.1.1. The model.	32
3.1.2. A brief foray into stability and instability issues.	33
3.1.3. A new approach to whistler waves.	34
3.1.4. Dispersion relation for whistler waves.	36
3.1.5. A class of oscillatory integrals.	38
3.2. Stationary phase expansions.	40
3.2.1. Preliminary estimates.	41
3.2.2. Absence of boundary critical points for sufficiently large times.	43
3.2.3. Sorting of the harmonics.	43
3.2.4. Reduction to a spatial phase.	44
3.2.5. Harmonics q for which the variations of $ q c(w) b_e \circ \bar{Y}_0$ are below the threshold limit.	45
3.2.6. Emission points in the case of overlapping values.	49
3.2.7. Quick comments about the propagation of singularities.	51
References	51

2. MULTISCALE ANALYSIS OF THE VLASOV PART

Many works are devoted to the mathematical analysis of fast rotating fluids. In dimension three, the penalization matrix can always be represented by the cross product with some vector field \tilde{B} . The most studied case is when \tilde{B} is a constant vector. For a discussion and for a number of references, we can refer to the book [6]. Look also at the contribution [11] for specific development related to rarefied plasmas.

A physically more relevant situation is when \tilde{B} undergoes variations in the spatial position. Decompose such a $\tilde{B}(x)$ into $\tilde{B}(x) = \tilde{b}(x) \tilde{r}^1(x)$, where $\tilde{b}(x)$ and $\tilde{r}^1(x)$ are respectively the amplitude and the unitary direction of $\tilde{B}(x)$. As long as $\tilde{b} \neq 0$ and $t \sim 1$, without inducing special difficulties, a rotation brings back to the case of a vector field \tilde{B} having a constant direction. This means that, without limiting the generality of the foregoing, the discussion can be achieved with $\tilde{r}^1 \equiv \vec{e}$, where $\vec{e} \in \mathbb{R}^3$ is a fixed vector of norm 1. Then, the analysis is already more complicated than in the constant case. There are special phenomena at stake, see for instance the articles [8, 12].

When looking at the long time evolution, that is for $t \sim \varepsilon^{-1}$ instead of $t \sim 1$, the changes in the unit directions generated by \tilde{B} , that is concerning \tilde{r}^1 , also appear to play a crucial role. The reduction to $\tilde{r}^1 \equiv \vec{e}$ is no more possible. Other tools must be developed, and other effects are expected. The aim of this Part 2 is precisely to fill this gap.

The starting point is (1.5), where it is assumed that (E, B) is a known smooth function. The matter is to obtain a good description of the asymptotic behaviour of $(x_\varepsilon, v_\varepsilon)(\tau, x, v)$ when ε goes to 0. Keeping in mind the physical observations (Pj), the discussion will be systematically tested in the earth's context. This will appear throughout the text inside paragraphs that will be introduced by the word: o Application. The subscript "e" will be added each time the geomagnetic field is implied.

The preliminary stage is to well understand what happens during intermediate times $t \sim 1$. This is achieved in Section 2.1. This allows to explain the contents of Step (P1), and to catch basic features about (P2). Then, in Section 2.2, we consider longer times $t \sim \varepsilon^{-1}$. There, the approach is completely new. The related results are in coherence with the observations reported in Step (P2) and beyond, they give a good insight into Step (P3).

2.1. The averaging method during intermediate times $t \sim 1$. Changing the time variable τ into $t := \varepsilon^{-1}\tau$, the dynamical system (1.5) still contains fast rotating terms:

$$(2.1) \quad \begin{cases} \partial_t x_\varepsilon = v_\varepsilon, & x_\varepsilon(0, x, v) = x, \\ \partial_t v_\varepsilon = \varepsilon^{-1} c(|v_\varepsilon|) v_\varepsilon \times \tilde{B} + \varepsilon \tilde{N}(v_\varepsilon, E, B), & v_\varepsilon(0, x, v) = v. \end{cases}$$

Before applying at the level of (2.1) the classical filtering method (Paragraph 2.1.3), it is necessary to straighten out the field lines (Paragraph 2.1.2). To this end, Assumptions on the vector field \tilde{B} are needed (Paragraph 2.1.1). These hypotheses provide significant advantages with a view to achieving the above objectives. Most importantly, they enable to argue globally in space. Technically, their implementation allows in the next Paragraph 2.1.2 to straighten out the field lines in an orthogonal way. Also, they give a direct access to a geometrical interpretation of quantities that will be introduced later on.

2.1.1. *Assumptions on the external magnetic field.* We work under the following hypotheses, which are both intrinsic [27] and quite general.

Assumption 2.1. [structure of the three-dimensional skew-symmetric matrix] Given some open set $\tilde{\mathcal{O}} \subset \mathbb{R}^3$, the vector field $\tilde{B} : \tilde{\mathcal{O}} \rightarrow \mathbb{R}^3$ satisfies the following conditions:

(i) The function \tilde{B} is a smooth nonzero function: $\tilde{B} \in C^\infty(\tilde{\mathcal{O}}; \mathbb{R}^3 \setminus \{0\})$. More precisely, there is a constant $c \in \mathbb{R}_+^*$ and a unit vector field $\tilde{r}^1 \in C^\infty(\tilde{\mathcal{O}}; \mathbb{S}^2)$ such that:

$$(2.2) \quad \tilde{B}(x) = -\tilde{b}(x) \tilde{r}^1(x), \quad |\tilde{r}^1(x)| = 1, \quad \tilde{b}(x) := |\tilde{B}(x)| > c, \quad \forall x \in \tilde{\mathcal{O}}.$$

(ii) There is a global smooth diffeomorphism $y : \tilde{\mathcal{O}} \rightarrow \mathcal{O}$, defined through the curvilinear coordinates $y^j : \tilde{\mathcal{O}} \rightarrow \mathbb{R}$ with $j \in \{1, 2, 3\}$, such that $\nabla_x y^1$ is everywhere collinear to \tilde{B} , and giving rise to a triply orthogonal system.

Here and later δ_k^j is the Kronecker symbol: $\delta_k^j = 0$ if $j \neq k$ and $\delta_k^j = 1$ if $j = k$. Locally, the condition (ii) amounts to the same thing as:

$$(2.3) \quad \nabla_x y^j = \tilde{d}^j \tilde{r}^j, \quad \tilde{r}^j \cdot \tilde{r}^k = \delta_k^j, \quad \tilde{d}^j := |\nabla_x y^j|, \quad \forall (j, k) \in \{1, 2, 3\}^2.$$

Exchanging the indices $j = 2$ and $j = 3$ if necessary, we can always make sure that the basis $(\tilde{r}^1, \tilde{r}^2, \tilde{r}^3)$ is right-handed. The contravariant metric tensor (g^{jj}) associated with the change y is diagonal, with diagonal coefficients $g^{jj} = (\tilde{d}^j)^2$. Recall that $\tilde{d}^j = (\tilde{d}_j)^{-1}$, where the \tilde{d}_j are the Lamé coefficients. Adopt the conventions:

$$(2.4) \quad r^j := \tilde{r}^j \circ y^{-1}, \quad d^j := \tilde{d}^j \circ y^{-1}, \quad \forall j \in \{1, 2, 3\}, \quad b := \tilde{b} \circ y^{-1}.$$

As a consequence of (2.3)-(2.4), we have:

$$(2.5) \quad D_x y \circ y^{-1} = D O, \quad D = \begin{pmatrix} d^1 & 0 & 0 \\ 0 & d^2 & 0 \\ 0 & 0 & d^3 \end{pmatrix}, \quad O = {}^t O^{-1} = \begin{pmatrix} t_{r^1} \\ t_{r^2} \\ t_{r^3} \end{pmatrix}.$$

We will note y the current variable in the domain \mathcal{O} . This variable y must not be confused with the diffeomorphism $y : \tilde{\mathcal{O}} \rightarrow \mathcal{O}$. From now on, a symbol V representing a vector in \mathbb{R}^d with $d \in \mathbb{N}^*$ will be denoted by $V = {}^t(V^1, \dots, V^d)$. The geometric interpretation of (ii) is discussed in [27]. The condition (ii) can be tested through differential equations on \tilde{B} . It allows to incorporate most physical situations, for instance the geomagnetic case.

◦ Application 1 [the geomagnetic field]. A **Van Allen belt** is one of at least two layers of charged particles (plasma) that is held in place [4] around the planet Earth by the geomagnetic field \tilde{B}_e . Whistler mode waves are principally issued from the outer radiation belt, denoted by \mathcal{Belt} in what follows, and simply referred to as *the* Van Allen belt. In the vicinity of \mathcal{Belt} , the function \tilde{B}_e can be adequately approximated as a dipole field, whose dipole axis D_e is tilted at an angle of 10 degrees with respect to the earth's rotational axis. Let $R \simeq 6,371 \times 10^6$ m be the earth radius. The cartesian coordinate system x , with dimensionless spatial variable $x := x/R$, may be adjusted in such a way that D_e and the (magnetic) equatorial plane $\mathcal{E}q$ are given by:

$$D_e := \{x \in \mathbb{R}^3; x^1 = x^2 = 0\}, \quad \mathcal{E}q := \{x \in \mathbb{R}^3; x^3 = 0\}.$$

The domain \mathcal{Belt} does not intersect D_e . It can be viewed as a volume of revolution with axis D_e , which is symmetric with respect to $\mathcal{E}q$. Moreover, the region \mathcal{Belt} is contained in a spherical shell which extends from an altitude of about three to ten Earth radii. This last property leads to the following framework:

$$(2.6) \quad \mathcal{Belt} \subset \tilde{\mathcal{O}}_e, \quad \tilde{\mathcal{O}}_e := \{x \in \mathbb{R}^3; 3 < |x| < 9, (x^1)^2 + (x^2)^2 \neq 0\}.$$

Let $\tilde{b}_0 \simeq 3, 1 \times 10^{-5} T$ be the size of $|\tilde{B}_e|$, as it is measured at the surface of the Earth. In rescaled units, the expression \tilde{B}_e is replaced by $\tilde{B}_e := \tilde{B}_e / \tilde{b}_0$. In the conventional cylindrical coordinates ρ , ϕ and z issued from x , the application $\tilde{B}_e(\cdot)$ does not depend on $\phi \in \mathbb{R}$. It can be described by [4]:

$$(2.7) \quad \tilde{B}_e(\rho, z) = \nabla \times \tilde{A} = \frac{1}{(\rho^2 + z^2)^{5/2}} \begin{pmatrix} -3\rho z \\ 0 \\ \rho^2 - 2z^2 \end{pmatrix}, \quad \tilde{A} = \frac{-\rho}{(\rho^2 + z^2)^{3/2}} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

The amplitude \tilde{b}_e of \tilde{B}_e is the function:

$$\tilde{b}_e(\rho, z) := (\rho^2 + z^2)^{-5/2} [(\rho^2 + z^2)^2 + 3z^2(\rho^2 + z^2)]^{1/2}, \quad \tilde{b}_e(\rho, 0) = \rho^{-3}.$$

It is clearly homogeneous of degree -3 . This remark combined with (2.6) gives rise to:

$$(2.8) \quad 0 < b_m := 9^{-3} \leq \tilde{b}_e(x) \leq b_M := 2 \times 3^{-3}, \quad \forall x \in \tilde{\mathcal{O}}_e.$$

In view of (2.8), the part (i) of Assumption 2.1 becomes obvious. On the other hand, we have $\text{div } \tilde{B}_e = 0$, but also $\text{curl } \tilde{B}_e = 0$. Thus, the Helmholtz decomposition of \tilde{B}_e can be achieved in the form $\tilde{B}_e = \nabla_x y_e^1$. This special choice of y_e^1 implies that $\tilde{d}_e^1 \equiv \tilde{b}_e$. To see why the remaining conditions in (ii) are verified, it suffices to produce:

$$(2.9) \quad y_e^1 := z(\rho^2 + z^2)^{-3/2} \in \mathbb{R}, \quad y_e^2 := -\rho^4(\rho^2 + z^2)^{-3} \in \mathbb{R}_-^*, \quad y_e^3 := \phi \in \mathbb{R}.$$

With these conventions, the coordinates y^1 , y^2 and y^3 can be interpreted respectively as a sort of latitude, (negative) altitude and longitude. From (2.6), we can infer that:

$$\mathcal{O}_e = y_e(\tilde{\mathcal{O}}_e) \subset \{(y^1, y^2, y^3) \in \mathbb{R} \times \mathbb{R}_-^* \times \mathbb{R}; |y^1| < 3^{-1}, -9^{-1} < y^2 < 0\}.$$

From (2.9), we can also compute $\nabla_x y_e^j = \tilde{d}_e^j \cdot {}^t \tilde{r}_e^j$ with $\tilde{d}_e^j := |\nabla_x y_e^j|$ for all $j \in \{1, 2, 3\}$, and check that the conditions ${}^t \tilde{r}_e^j \cdot \tilde{r}_e^k = \delta_k^j$ are satisfied. It appears here that a convenient choice for $y_e : \tilde{\mathcal{O}}_e \rightarrow \mathcal{O}_e$ can be obtained as the composition of two orthogonal change of variables. First, the one from x to (ρ, ϕ, z) . Then, the one from (ρ, ϕ, z) to the toroidal coordinates $y_e = ({}^t y_e^1, y_e^2, y_e^3)$ defined in (2.9). In the coordinate system y , the field lines become a family of parallel straight lines:

$$\mathcal{A}(\hat{y}) := \{(y^1, \hat{y}) \in \mathbb{R} \times \mathbb{R}_-^* \times \mathbb{R}; y^1 \in \mathbb{R}\}, \quad \hat{y} := (y^2, y^3).$$

The inclusion (2.6) is only indicative. The minimal distance ρ_m and the maximal distance ρ_M , which can be measured from a particle to the earth's center knowing that this particle is located in $\mathcal{Belt} \cap \mathcal{E}q$, can be measured more accurately, with $3 < \rho_m < \rho_M < 9$. Remark that $y^2 = -\rho^{-2}$ when $y^1 = 0$. Accordingly, define $y_m^2 := -\rho_m^{-2}$ and $y_M^2 := -\rho_M^{-2}$. The altitude y^2 being fixed, mark the maximal latitude $h_e(y^2)$ which can be reached by a particle trapped along the field line $\mathcal{A}(y^2, y^3)$. The value $h_e(y^2)$ corresponds also to the extremal position of a mirror point having the altitude y^2 .

All ingredients y_m^2 , y_M^2 and $h_e(\cdot)$ can be adjusted on the basis of concrete observations. They allow to determine precisely the shape of the Van Allen belt, leading to:

$$(2.10) \quad \mathcal{Belt} = \{y; |y^1| \leq h_e(y^2)\}, \quad h_e \in C_0^\infty(\mathbb{R}^*; \mathbb{R}_+), \quad \text{supp } h_e = [y_m^2, y_M^2].$$

Lemma 2.1. *The functions $r_e^j := \tilde{r}_e^j \circ y_e^{-1}$, $d_e^j := \tilde{d}_e^j \circ y_e^{-1}$ and $b_e := \tilde{b}_e \circ y_e^{-1}$ satisfy:*

$$(2.11a) \quad \partial_{y^j} r_e^3 = 0, \quad \forall j \in \{1, 2\},$$

$$(2.11b) \quad \partial_{y^3} d_e^j = 0, \quad \forall j \in \{1, 2, 3\}, \quad (\partial_{y^3} r_e^1) \cdot r_e^2 = 0,$$

$$(2.11c) \quad \partial_{y^3} b_e = 0, \quad \partial_{y^1} b_e(0, y^2) = 0, \quad \partial_{y^1 y^1}^2 b_e(0, y^2) > 0, \quad \partial_{y^2} b_e < 0.$$

The function $b_e(\cdot, y^2)$ is strictly increasing on \mathbb{R}_+^* . On the other hand, for all $y^2 \in \mathbb{R}_-^*$, the partial applications $b_e(\cdot, y^2)$ and $d_e^j(\cdot, y^2)$ with $j \in \{1, 2, 3\}$ are even in $y^1 \in \mathbb{R}$.

Proof. By construction $r_e^3 = {}^t(-\sin y^3, \cos y^3, 0)$, explaining (2.11a). From the definitions in (2.9), we can easily deduce that \tilde{d}_e^1 and \tilde{d}_e^2 depend only on (ρ, z) , whereas $\tilde{d}_e^3 = \rho^{-1}$. This gives rise to the first part of (2.11b). Knowing what is r_e^1 for $\phi = 0$, the values of r_e^1 for $\phi \neq 0$ are obtained by a rotation of angle ϕ . The vector $\partial_{y^3} r_e^1$ is collinear to r_e^3 , and therefore $(\partial_{y^3} r_e^1) \cdot r_e^2 = 0$. Obviously, b_e does not depend on $y^3 \equiv \phi$. In view of (2.9), we have the implicit relation:

$$(2.12) \quad R^4 (y^1)^2 + R \sqrt{-y^2} = 1, \quad R := (\rho^2 + z^2)^{1/2}.$$

With (2.7) and (2.12), we can compute:

$$(2.13) \quad b_e^2 = \frac{1 + 3(y^1)^2 R^4}{R^6}, \quad \partial_{y^1} R = \frac{-2y^1 R^5}{1 + 3(y^1)^2 R^4}, \quad \partial_{y^2} R = \frac{R^2}{2\sqrt{-y^2} [1 + 3(y^1)^2 R^4]}.$$

From (2.13), it is easy to verify that $\partial_{y^2} b_e(y) < 0$, and that $\partial_{y^1} b_e(y) > 0$ when $y^1 > 0$. The information about $\partial_{y^1 y^1}^2 b_e(0, y^2)$ comes from $\partial_{y^1 y^1}^2 (b_e)^2(0, y^2) = 18 R^{-2}$. The symmetry with respect to $\mathcal{E}q$ guarantees that $b_e(\cdot, y^2)$ and $d_e^j(\cdot, y^2)$ are even in $y^1 \in \mathbb{R}$. \square

The requirements on \tilde{B} made in Assumption 2.1 are fulfilled by other magnetic fields, such as Jupiter's (Voyager spacecraft detected whistler-like activity in the vicinity of Jupiter). Even in the case of the earth, they are compatible with more accurate models, which for instance incorporate solar effects. However, for the sake of simplicity, these aspects will not be taken into account.

2.1.2. *Straightening out the field lines.* Change (x, v) into $(y, w) := (y(x), O \circ y(x) v)$. Write the system (2.1) in terms of these new variables (y, w) to get:

$$(2.14) \quad \begin{cases} \partial_t y_\varepsilon = D(y_\varepsilon) w_\varepsilon, & y_\varepsilon(0, y, w) = y, \\ \partial_t w_\varepsilon = \varepsilon^{-1} c(|w_\varepsilon|) B(\varepsilon, y_\varepsilon, w_\varepsilon) \times w_\varepsilon + \varepsilon \mathcal{N}(y_\varepsilon, w_\varepsilon, E, B), & w_\varepsilon(0, y, w) = w. \end{cases}$$

By construction, we have $B(\varepsilon, \cdot) = B_0(\cdot) + \varepsilon B_1(\cdot)$ with:

$$(2.15) \quad B_0 = \begin{pmatrix} B_0^1 \\ B_0^2 \\ B_0^3 \end{pmatrix} = \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} B_1^1 \\ B_1^2 \\ B_1^3 \end{pmatrix} = \begin{pmatrix} [(D w \cdot \nabla_y)^t r^3] \cdot r^2 \\ [(D w \cdot \nabla_y)^t r^1] \cdot r^3 \\ [(D w \cdot \nabla_y)^t r^2] \cdot r^1 \end{pmatrix}.$$

The nonlinear perturbation \mathcal{N} can be deduced from $\tilde{\mathcal{N}}$ as follows:

$$(2.16) \quad \mathcal{N}(y, w, E, B) := O(y) \tilde{\mathcal{N}}({}^t O(y) w, E \circ y^{-1}(y), B \circ y^{-1}(y)).$$

The dynamical system (2.14) is of the same type as (2.1) except that the singular part of the penalization matrix, which is represented above by B_0 , points now in the constant direction $e^1 = {}^t(1, 0, 0)$. Thus, the effect and the interest of the step 2.1.2 is to straighten out the direction of the vector field $B(\cdot)$. This is done only at leading order. Since the diffeomorphism $y : \tilde{\mathcal{O}} \rightarrow \mathcal{O}$ is an orthogonal change of variables, this operation creates at the level of the equation on w_ε a skew-symmetric $O(1)$ contribution that is materialized by the cross product $B_1 \times$. Observe that the action of B_1 is absent if and only if r_1 (and therefore r_2 and r_3) can be taken as a constant function.

2.1.3. Filtering out the main oscillations. The filtering method has been first introduced by S. Schochet [23], and then by E. Grenier [14]. It is today well-known. It has been applied in the article [12] to deal with a two-dimensional version of the present framework. Below, it is implemented to start the analysis. Introduce:

$$\Lambda := e_1 \times = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad e^{\theta \Lambda} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}, \quad \theta \in \mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z}).$$

Consider a smooth scalar function $\phi_\varepsilon(t, y, w)$, and replace w_ε by $u_\varepsilon := e^{-\phi_\varepsilon \Lambda / \varepsilon} w_\varepsilon$. Define $z_\varepsilon := {}^t(y_\varepsilon, u_\varepsilon) \in \mathbb{R}^3 \times \mathbb{R}^3$, which is a function of the variables (t, y, w) . When expressing the preceding quantities in terms of z , oscillations are introduced. For instance:

$$c(|w_\varepsilon|) B^1(\varepsilon, y_\varepsilon, w_\varepsilon) = c(|w_\varepsilon|) B_0^1(y_\varepsilon) + \varepsilon c(|w_\varepsilon|) B_1^1(y_\varepsilon, w_\varepsilon) = \mathcal{B}(\varepsilon, \varepsilon^{-1} \phi_\varepsilon(t, y, w), z_\varepsilon).$$

We have ${}^t \Lambda = -\Lambda$. The matrix $e^{\theta \Lambda}$ is orthogonal so that $|w_\varepsilon| = |u_\varepsilon|$ and:

$$\mathcal{B}(\varepsilon, \theta, z) = \mathcal{B}_0(z) + \varepsilon \mathcal{B}_1(\theta, z), \quad \mathcal{B}_0(z) := c(|u|) b(y), \quad \mathcal{B}_1(\theta, z) := c(|u|) B_1^1(y, e^{\theta \Lambda} u).$$

Definition 2.1 (eikonal equation). *The phase ϕ_ε is adjusted in such a way that:*

$$(2.17) \quad \partial_t \phi_\varepsilon = \mathcal{B}(\varepsilon, \varepsilon^{-1} \phi_\varepsilon(t, y, w), z_\varepsilon), \quad \phi_\varepsilon(0) = 0.$$

The right-hand side of (2.17) is based on $B_0^1 \equiv b$. Note that, for subsequent important reasons, the apparently small term εB_1^1 is also incorporated inside \mathcal{B} . The equation (2.17) is not self-contained. It must be coupled with a condition allowing to determine z_ε . Taking into account (2.17), the time evolution of z_ε is governed by a differential equation that can be formalized according to:

$$(2.18) \quad \partial_t z_\varepsilon = \mathcal{A}(\varepsilon, \varepsilon^{-1} \phi_\varepsilon(t, y, w), z_\varepsilon), \quad z_\varepsilon(0) = {}^t(y, w).$$

At this stage, we are faced with the system (2.17)-(2.18). In comparison with (2.14), the source terms inside (2.17) and (2.18) are no more of size ε^{-1} . The two applications $\mathcal{B}(\cdot, \theta, z)$ and $\mathcal{A}(\cdot, \theta, z)$ are now smooth in $\varepsilon \in [0, 1]$. On the other hand, the source terms inside (2.17) and (2.18) involve many oscillations at frequencies of the order ε^{-1} . Indeed, both functions $\mathcal{B}(\varepsilon, \cdot, z)$ and $\mathcal{A}(\varepsilon, \cdot, z)$ are periodic in θ , with period 2π .

In practice, we find that $\mathcal{A}(\varepsilon, \cdot) = \mathcal{A}_0 + \varepsilon \mathcal{A}_1$ with:

$$(2.19a) \quad \begin{pmatrix} \mathcal{A}_0^1 \\ \mathcal{A}_0^2 \\ \mathcal{A}_0^3 \end{pmatrix}(\theta, z) = D(y) e^{\theta\Lambda} u = \begin{pmatrix} d^1(y) u^1 \\ d^2(y) (\cos \theta u^2 - \sin \theta u^3) \\ d^3(y) (\sin \theta u^2 + \cos \theta u^3) \end{pmatrix},$$

$$(2.19b) \quad \begin{pmatrix} \mathcal{A}_0^4 \\ \mathcal{A}_0^5 \\ \mathcal{A}_0^6 \end{pmatrix}(\theta, z) = c(|u|) e^{-\theta\Lambda} \begin{pmatrix} 0 \\ B_1^2 \\ B_1^3 \end{pmatrix} \times e^{\theta\Lambda} u \\ = c(|u|) \begin{pmatrix} (B_1^2 \sin \theta - B_1^3 \cos \theta) u^2 + (B_1^2 \cos \theta + B_1^3 \sin \theta) u^3 \\ -(B_1^2 \sin \theta - B_1^3 \cos \theta) u^1 \\ -(B_1^2 \cos \theta + B_1^3 \sin \theta) u^1 \end{pmatrix},$$

$$(2.19c) \quad \begin{pmatrix} \mathcal{A}_1^1 \\ \mathcal{A}_1^2 \\ \mathcal{A}_1^3 \end{pmatrix}(\theta, z) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \mathcal{A}_1^4 \\ \mathcal{A}_1^5 \\ \mathcal{A}_1^6 \end{pmatrix}(\theta, z) = e^{-\theta\Lambda} \mathcal{N}(y, e^{\theta\Lambda} u, E, B).$$

The contribution issued from the part ${}^t(\mathcal{A}_0^4, \mathcal{A}_0^5, \mathcal{A}_0^6)$ is now nonlinear since the functions B_1^2 and B_1^3 must be evaluated at the point $(y, e^{\theta\Lambda} u)$. Still, some skew-symmetric property has been preserved in so far as:

$${}^t(\mathcal{A}_0^4, \mathcal{A}_0^5, \mathcal{A}_0^6)(\theta, z) = \Delta(\theta, z) u, \quad {}^t\Delta(\theta, z) = -\Delta(\theta, z), \quad \forall (\theta, z) \in \mathbb{T} \times \mathbb{R}^6.$$

Define $\bar{\mathcal{A}}_0$ as being the mean value of \mathcal{A}_0 with respect to θ . Observe that $\bar{\mathcal{A}}_0$ is composed of the diagonal part ${}^t(d^1(y)u^1, 0, 0)$ and of the skew-symmetric action $\bar{\Delta}(z)u$.

2.2. The approach to reach long times $t \sim \varepsilon^{-1}$. Change the variable t into $\tau := \varepsilon t$. Then, the system (2.17)-(2.18) is translated into:

$$(2.20) \quad \partial_\tau \begin{pmatrix} z_\varepsilon \\ \phi_\varepsilon \end{pmatrix} = \frac{1}{\varepsilon} \begin{pmatrix} \mathcal{A} \\ \mathcal{B} \end{pmatrix} \left(\varepsilon, \frac{\phi_\varepsilon}{\varepsilon}, z_\varepsilon \right), \quad \begin{pmatrix} z_\varepsilon \\ \phi_\varepsilon \end{pmatrix}(0) = \begin{pmatrix} {}^t(y, w) \\ 0 \end{pmatrix}.$$

The long time evolution ($t \sim \varepsilon^{-1}$ or $\tau \sim 1$) is a delicate question for many reasons. There is the presence of singular terms, like $\varepsilon^{-1}(\mathcal{A}_0^1, \mathcal{A}_0^2, \mathcal{A}_0^3)$ and $\varepsilon^{-1}\mathcal{B}_0$, which at first sight inherit no special structure. And, even concerning the skew-symmetric part $\varepsilon^{-1}\Delta(\theta, z)u$, the filtering method cannot be repeated along the same lines as before. Objections come from the presence already of oscillations ($\partial_\theta \Delta \neq 0$) and from the nonlinearity of the penalization matrix ($\nabla_z \Delta \neq 0$). To go further, other arguments must be implemented.

2.2.1. Preliminaries and guidelines. Looking at (2.20) gives the impression that the fast dynamics in long times are controlled by the unworkable term $(\mathcal{A}_0, \mathcal{B}_0)$. But it would be without paying attention to the presence of two scales, playing different roles. In fact, the oscillations inside $(\mathcal{A}_0, \mathcal{B}_0)$ must be handled separately. As we will see, what principally matters is the influence of $\bar{\mathcal{A}}_0$.

◇ A first idea when dealing with (2.20) is to adequately separate the different scales. To this end, we have to introduce notations. Given $T \in \mathbb{R}_+^*$, note $\mathbb{T}_T := \mathbb{R}/(T\mathbb{Z})$ and, as usual, note simply $\mathbb{T} := \mathbb{T}_{2\pi} \equiv \mathbb{R}/(2\pi\mathbb{Z})$. Let $Z(\tau, t, \theta) \in \mathcal{C}^0(\mathbb{R} \times \mathbb{T}_T \times \mathbb{T})$. Introduce:

$$\bar{Z}(\tau, t) := \frac{1}{2\pi} \int_0^{2\pi} Z(\tau, t, \theta) d\theta, \quad \langle \bar{Z} \rangle(\tau) := \frac{1}{2\pi T} \int_0^T \int_0^{2\pi} Z(\tau, t, \theta) dt d\theta.$$

We can further decompose the Fourier serie of Z according to:

$$Z(\tau, t, \theta) = \underbrace{c_{(0,0)}(\tau)}_{\langle \bar{Z} \rangle(\tau)} + \underbrace{\sum_{p \in \mathbb{Z}^*} c_{(p,0)}(\tau) e^{ipt/T}}_{\bar{Z}^*(\tau, t)} + \underbrace{\sum_{(p,q) \in \mathbb{Z} \times \mathbb{Z}^*} c_{(p,q)}(\tau) e^{i(pt/T + q\theta)}}_{Z^*(\tau, t, \theta)}.$$

To better set the scene, let us comment the above conventions. The symbols $\bar{\cdot}$ and $\langle \cdot \rangle$ mean that we extract the mean value of $Z(\tau, t, \theta)$ in θ , and of $\bar{Z}(\tau, t)$ in t . The marks \cdot^* and \cdot^* at the top right of Z indicate that we deal with the oscillating part of $Z(\tau, t, \theta)$ in θ , and of $\bar{Z}(\tau, t)$ in t . The actions ∂_θ^{-1} and ∂_t^{-1} will correspond to the inverse operators of ∂_θ and ∂_t , that is $\partial_\theta^{-1} \partial_\theta = Id$ and $\partial_t^{-1} \partial_t = Id$, with values in the set of functions with zero mean. For instance:

$$\partial_t^{-1} \bar{Z}^*(\tau, t) := \int_0^t \bar{Z}(\tau, s) ds - \frac{1}{T} \int_0^T \left(\int_0^t \bar{Z}(\tau, s) ds \right) dt, \quad \langle \partial_t^{-1} \bar{Z}^* \rangle = 0.$$

The property (2.21) below is obvious. Still, as it will be often used, it is emphasised.

Lemma 2.2. *Let $Z \in C^0(\mathbb{T}_T)$. The integral of Z is the sum of a linear function plus a periodic (bounded) function:*

$$(2.21) \quad \int_0^t Z(s) ds = \langle Z \rangle t + \int_0^t Z^*(s) ds, \quad \sup_{t \in \mathbb{R}} \left| \int_0^t Z^*(s) ds \right| < +\infty. \quad \diamond$$

◇ A second idea is to remove \mathcal{A}_0 from the system (2.20). This requires a rectification procedure whose principle is exposed in Paragraph 2.2.2, and whose effective construction is achieved from Paragraph 2.2.3 to 2.2.6. This rectification is based on two operations. First, a conjugaison of the system (2.20) with a flow induced by $\bar{\mathcal{A}}_0$. This pertains to a nonlinear conjugaison since $\bar{\mathcal{A}}_0$ is a nonlinear function, see (2.35) and (2.37). Secondly, while bearing in mind the presence of oscillations, an absorption of the oscillating terms through a sort of elliptic equation. ◇

◇ A third idea is to exploit the properties of the geomagnetic field to show that the flow generated by $\bar{\mathcal{A}}_0$ is globally defined in time, and in fact periodic in time (Proposition 2.1). It is at this level that the variations of the vector field \tilde{r}^1 , that is of the unit directions generated by \tilde{B} , come into play. Depending on the situation, these variations can induce destabilizing effects or, inversely, confinement properties. The geometrical criterion giving rise to a potential well is exposed in Assumption 2.2. It is satisfied in the case of magnetic dipoles, see Lemma 2.5. ◇

The main purpose of this Section 2.2 is to account for the observations which have been reported in Steps (P2) and (P3). Throughout the discussion, the hamiltonian structure underlying the dynamical system (1.5) will be omnipresent. In particular, it will manifest in terms of a *reduced hamiltonian*, see Definition 2.3. In Paragraph 2.2.4, we also wish to emphasize two new Mathematical interpretations (1 and 2) of physical phenomena: the first explains the bouncing back through some geometrical argument; the second highlights the notion of extended Van Allen belt, from which the Van Allen belt is derived.

2.2.2. *Approximate rectification of the vector field \mathcal{A} .* What is called here "approximate rectification", it is a (time dependent) oscillatory transformation of z_ε (given by a lift) allowing to get rid of \mathcal{A}_0 . Concretely, consider two maps:

$$\begin{aligned} \Xi : \mathbb{R}_+ \times \mathbb{R}^6 &\longrightarrow \mathbb{R}^6 & \tilde{\Xi} : \mathbb{R}_+ \times \mathbb{T} \times \mathbb{R}^6 &\longrightarrow \mathbb{R}^6 \\ (t, \mathfrak{z}) &\longmapsto \Xi(t, \mathfrak{z}), & (t, \theta, \mathfrak{z}) &\longmapsto \tilde{\Xi}(t, \theta, \mathfrak{z}). \end{aligned}$$

Then, look at the system:

$$(2.22) \quad \partial_\tau \begin{pmatrix} \mathfrak{z}_\varepsilon \\ \varphi_\varepsilon \end{pmatrix} (\tau) = \frac{1}{\varepsilon} \begin{pmatrix} \mathbf{A} \\ \mathbf{b} \end{pmatrix} \left(\varepsilon, \frac{\tau}{\varepsilon}, \frac{\varphi_\varepsilon}{\varepsilon}, \mathfrak{z}_\varepsilon \right), \quad \begin{pmatrix} \mathfrak{z}_\varepsilon \\ \varphi_\varepsilon \end{pmatrix} (0) = \begin{pmatrix} \mathfrak{z}_{i\varepsilon} \\ 0 \end{pmatrix},$$

where the initial data $\mathfrak{z}_{i\varepsilon}$ is adjusted such that $\Xi(0, \mathfrak{z}_{i\varepsilon}) + \varepsilon \tilde{\Xi}(0, 0, \mathfrak{z}_{i\varepsilon}) = (y, w)$, and where, assuming that the matrix $D_{\mathfrak{z}}\Xi(t, \mathfrak{z})$ is invertible for all (t, \mathfrak{z}) , the applications \mathbf{A} and \mathbf{b} are given (for ε small enough) by the chain rules:

$$(2.23) \quad \mathbf{A}(\varepsilon, t, \theta, \mathfrak{z}) := [D_{\mathfrak{z}}\Xi(t, \mathfrak{z}) + \varepsilon D_{\mathfrak{z}}\tilde{\Xi}(t, \theta, \mathfrak{z})]^{-1} [\mathcal{A}(\varepsilon, \theta, \Xi(t, \mathfrak{z}) + \varepsilon \tilde{\Xi}(t, \theta, \mathfrak{z})) - \partial_t \Xi(t, \mathfrak{z}) - \mathcal{B}(\varepsilon, \theta, \Xi(t, \mathfrak{z}) + \varepsilon \tilde{\Xi}(t, \theta, \mathfrak{z})) \partial_\theta \tilde{\Xi}(t, \theta, \mathfrak{z}) - \varepsilon \partial_t \tilde{\Xi}(t, \theta, \mathfrak{z})],$$

$$(2.24) \quad \mathbf{b}(\varepsilon, t, \theta, \mathfrak{z}) := \mathcal{B}(\varepsilon, \theta, \Xi(t, \mathfrak{z}) + \varepsilon \tilde{\Xi}(t, \theta, \mathfrak{z})) = \mathcal{B}_0(\Xi(t, \mathfrak{z})) + O(\varepsilon).$$

There is a link between (2.20) and (2.22). Indeed, any solution $(\mathfrak{z}_\varepsilon, \varphi_\varepsilon)$ to (2.22) gives rise to a solution $(z_\varepsilon, \phi_\varepsilon)$ to (2.20), just by posing:

$$(2.25) \quad z_\varepsilon(\tau) := \Xi(\varepsilon^{-1} \tau, \mathfrak{z}_\varepsilon(\tau)) + \varepsilon \tilde{\Xi}(\varepsilon^{-1} \tau, \varepsilon^{-1} \varphi_\varepsilon(\tau), \mathfrak{z}_\varepsilon(\tau)), \quad \phi_\varepsilon(\tau) := \varphi_\varepsilon(\tau).$$

Since the functions \mathbf{A} and \mathbf{b} are smooth with respect to $\varepsilon \in [0, 1]$, they can be expanded in powers of ε , say $\mathbf{A}(\varepsilon, \cdot) = \mathbf{A}_0 + \varepsilon \mathbf{A}_1 + \varepsilon^2 \mathbf{A}_2 + \dots$ and $\mathbf{b}(\varepsilon, \cdot) = \mathbf{b}_0 + \varepsilon \mathbf{b}_1 + \varepsilon^2 \mathbf{b}_2 + \dots$ with:

$$(2.26) \quad \mathbf{A}_0(t, \theta, \mathfrak{z}) := D_{\mathfrak{z}}\Xi(t, \mathfrak{z})^{-1} [\mathcal{A}_0(\theta, \Xi(t, \mathfrak{z})) - \partial_t \Xi(t, \mathfrak{z}) - \mathcal{B}_0(\Xi(t, \mathfrak{z})) \partial_\theta \tilde{\Xi}(t, \theta, \mathfrak{z})].$$

In view of (2.22), there is little possibility to get on \mathfrak{z}_ε a uniform control in $\varepsilon \in]0, 1]$, which could be valid for long times $\tau \sim 1$, except if $\mathbf{A}_0 \equiv 0$. Looking at the mean part of (2.26), this last condition amounts to the same thing as imposing the *first modulation equation*:

$$(2.27) \quad \partial_t \bar{Z}_0 - \bar{\mathcal{A}}_0(\bar{Z}_0) = 0, \quad \bar{Z}_0(t) \equiv \Xi(t, \mathfrak{z}).$$

The positivity of $\mathcal{B}_0(z) = c(|u|)b(y)$ is guaranteed by (2.2) and the definition of $c(\cdot)$. This allows to divide by \mathcal{B}_0 . Then, looking at the oscillating part of (2.26), we can give a sense to the condition:

$$(2.28) \quad \tilde{\Xi}(t, \theta, \mathfrak{z}) = \mathcal{B}_0 \circ \Xi(t, \mathfrak{z})^{-1} (\partial_\theta^{-1} \mathcal{A}_0^*)(\theta, \Xi(t, \mathfrak{z})).$$

Assuming that $\Xi(t, \mathfrak{z})$ can be identified through (2.27), the relation (2.28) determines $\tilde{\Xi} \equiv \tilde{\Xi}^*$ without ambiguity. By the way, we recover the predominance of $\bar{\mathcal{A}}_0$ over \mathcal{A}_0^* . The problem is about (2.27). It is to show that the autonomous differential equation (2.27) has global bounded solutions. For clarity and to exhibit important underlying features, it is better to detail the three intermediate changes $\Xi_1(\mathfrak{z})$, $\Xi_2(\mathfrak{z})$ and $\Xi_3(t, \mathfrak{z})$ leading to:

$$(2.29) \quad \Xi(t, \mathfrak{z}) := \Xi_1 \circ \Xi_2 \circ \Xi_3(t, \mathfrak{z}).$$

There will be a Paragraph devoted to the construction of each change Ξ_j , for $j \in \{1, 2, 3\}$.

2.2.3. *The autonomous change of variables $\Xi_1(\mathfrak{z})$.* Use spherical coordinates with radial distance \mathbf{u} , azimuth angle ς and polar angle \mathbf{a} to represent the velocity vector u according to the decomposition $u = {}^t(u^\parallel, u^\perp)$ with $u^\parallel \equiv u^1 \in \mathbb{R}$, $u^\perp \equiv {}^t(u^2, u^3) \in \mathbb{R}^2$, and:

$$(2.30) \quad u^\parallel := \mathbf{u} \cos \varsigma, \quad u^\perp = \mathbf{u} \sin \varsigma {}^t(\cos \mathbf{a}, \sin \mathbf{a}), \quad (\varsigma, \mathbf{u}, \mathbf{a}) \in \mathbb{R} \times \mathbb{R}_+^* \times \mathbb{R}.$$

Similarly, decompose $w = {}^t(w^\parallel, w^\perp)$ with $w^\parallel \equiv w^1 \in \mathbb{R}$ and $w^\perp \equiv {}^t(w^2, w^3) \in \mathbb{R}^2$ into:

$$(2.31) \quad w^\parallel := \mathbf{w} \cos \varsigma_w, \quad w^\perp = \mathbf{w} \sin \varsigma_w {}^t(\cos \mathbf{a}_w, \sin \mathbf{a}_w), \quad (\varsigma_w, \mathbf{w}, \mathbf{a}_w) \in \mathbb{R} \times \mathbb{R}_+^* \times \mathbb{R}.$$

Definition 2.2. *[definition of Ξ_1] Take $\mathfrak{z} = {}^t(y, \varsigma, \mathbf{u}, \mathbf{a})$ and define:*

$$\Xi_1(\mathfrak{z}) = {}^t(\Xi_1^1, \Xi_1^2, \Xi_1^3, \Xi_1^4, \Xi_1^5, \Xi_1^6)(\mathfrak{z}) := (y^1, y^2, y^3, \mathbf{u} \cos \varsigma, \mathbf{u} \sin \varsigma \cos \mathbf{a}, \mathbf{u} \sin \varsigma \sin \mathbf{a}).$$

Under the action of Ξ_1 , the identity (2.26) reduces to:

$$(2.32) \quad A_0(\theta, \mathfrak{z}) = D_3 \Xi_1(\mathfrak{z})^{-1} \mathcal{A}_0(\theta, \Xi_1(\mathfrak{z})), \quad \bar{A}_0(\mathfrak{z}) = D_3 \Xi_1(\mathfrak{z})^{-1} \bar{\mathcal{A}}_0(\Xi_1(\mathfrak{z})),$$

and the equation (2.27) is translated into:

$$(2.33) \quad \partial_t \bar{\mathfrak{z}}_0 - \bar{A}_0(\bar{\mathfrak{z}}_0) = 0, \quad \bar{\mathfrak{z}}_0(0) = \bar{\mathfrak{z}}_{i0}^{s1} := (y, \varsigma_w, \mathbf{w}, \mathbf{a}_w), \quad \bar{Z}_0 = \Xi_1(\bar{\mathfrak{z}}_0).$$

We have also:

$$(2.34) \quad A(\varepsilon, \theta, \mathfrak{z}) = D_3 \Xi_1(\mathfrak{z})^{-1} \mathcal{A}(\varepsilon, \theta, \Xi_1(\mathfrak{z})) = A_0(\theta, \mathfrak{z}) + \varepsilon A_1(\theta, \mathfrak{z}),$$

and we can obtain:

$$(2.35) \quad \begin{pmatrix} A_0^1 \\ A_0^2 \\ A_0^3 \end{pmatrix}(\theta, \mathfrak{z}) = \mathbf{u} \begin{pmatrix} d^1(y) \cos \varsigma \\ d^2(y) \cos(\theta + \mathbf{a}) \sin \varsigma \\ d^3(y) \sin(\theta + \mathbf{a}) \sin \varsigma \end{pmatrix}, \quad \begin{pmatrix} \bar{A}_0^1 \\ \bar{A}_0^2 \\ \bar{A}_0^3 \end{pmatrix}(\mathfrak{z}) = \mathbf{u} \begin{pmatrix} d^1(y) \cos \varsigma \\ 0 \\ 0 \end{pmatrix}.$$

From above, we deduce that $\bar{\mathfrak{z}}_0^2$ and $\bar{\mathfrak{z}}_0^3$ do not depend on t . This is in coherence with [4] and gyrokinetic theory [5]. We recover here that the main spatial effect (in position $x \in \tilde{\mathcal{O}}$) during intermediate times ($t \sim 1$) is a drift along the field lines.

Lemma 2.3. *[conservation of the kinetic energy] Assume that $E \equiv 0$. Then, the quantities $|v_\varepsilon|$, $|w_\varepsilon|$, $|u_\varepsilon|$ or \mathbf{u}_ε are constant for all times $\tau \in \mathbb{R}_+$, equal to $|v| = |w| = \mathbf{w}$.*

Proof. The Definition 2.30 of \mathbf{u}_ε says that $\mathbf{u}_\varepsilon = |\mathbf{u}_\varepsilon|$. Since the transformations allowing to pass from v_ε to w_ε , and then from w_ε to \mathbf{u}_ε , are orthogonal, we find that:

$$|v_\varepsilon|^2 = |O \circ \mathbf{y}(x_\varepsilon) v_\varepsilon|^2 = |w_\varepsilon|^2 = |e^{\phi_\varepsilon \Lambda / \varepsilon} u_\varepsilon|^2 = |u_\varepsilon|^2 = \mathbf{u}_\varepsilon^2.$$

It suffices to argue at the level of v_ε . From (1.5) and (1.6), we get:

$$\partial_\tau |v_\varepsilon|^2 = 2 v_\varepsilon \cdot \tilde{\mathcal{N}}(v_\varepsilon, 0, B) = 2 (1 - |v_\varepsilon|^2)^{1/2} v_\varepsilon \cdot (v_\varepsilon \times B) = 0. \quad \square$$

The traduction at the level of (2.33) of this result 2.3 is that $\bar{\mathfrak{z}}_0^5(t) = \mathbf{w}$ for all $t \in \mathbb{R}_+^*$, or equivalently that $A^5(\varepsilon, \cdot) \equiv 0$, that is $A_0^5(\cdot) \equiv 0$ and $A_1^5(\cdot) \equiv 0$. Now, when $E \neq 0$, the proof indicates that $\partial_t |v_\varepsilon|^2 = O(\varepsilon)$. This means that the conservation of the kinetic energy remains true at the main order ε^0 . As a consequence, we still have $A_0^5 \equiv 0$, as confirmed by a direct computation. By construction:

$$(2.36) \quad \begin{pmatrix} A_0^4 \\ A_0^6 \end{pmatrix}(\theta, \mathfrak{z}) = c(\mathbf{u}) \begin{pmatrix} -B_1^2 \sin(\theta + \mathbf{a}) + B_1^3 \cos(\theta + \mathbf{a}) \\ -\cotan \varsigma [B_1^2 \cos(\theta + \mathbf{a}) + B_1^3 \sin(\theta + \mathbf{a})] \end{pmatrix}.$$

In (2.36), there is some hidden dependence in θ because, in view of (2.15), the functions $B_1^2(y, w)$ and $B_1^3(y, w)$ must be evaluated at the point $w = e^{\theta\Lambda} u$. Taking this remark into account and averaging the expressions A_0^4 and A_0^6 with respect to θ gives rise to:

$$(2.37) \quad \begin{pmatrix} \bar{A}_0^4 \\ \bar{A}_0^6 \end{pmatrix}(\mathfrak{z}) = \mathbf{u} c(\mathbf{u}) \begin{pmatrix} \sin \varsigma \mathcal{C}(y) \\ \cos \varsigma \mathcal{E}(y) \end{pmatrix}, \quad \begin{aligned} \mathcal{C}(y) &:= [(d^2 \partial_{y^2}{}^t r^2 + d^3 \partial_{y^3}{}^t r^3) \cdot r^1](y)/2, \\ \mathcal{E}(y) &:= [(d^2 \partial_{y^2}{}^t r^3 - d^3 \partial_{y^3}{}^t r^2) \cdot r^1](y)/2. \end{aligned}$$

Consider the surfaces $\mathcal{S}_c^k := \{x \in \tilde{\mathcal{O}}; y^k(x) = c\}$, with $k \in \{1, 2, 3\}$.

Lemma 2.4. *[geometrical interpretation of the coefficients \mathcal{C} and \mathcal{E}] We have $\mathcal{E}(y) = 0$. Moreover, $\mathcal{C}(y)$ coincides with the mean curvature of the surface $\mathcal{S}_{y^1}^1$, computed at the point x such that $y(x) = y$.*

Proof. Let γ^j be some integral curve of r^j . It can be parametrized by y^j or by its arc length s , with the correspondance $d^j \partial_{y^j} \equiv \partial_s$. By Dupin's theorem, γ^j is necessarily a line of curvature for the surfaces \mathcal{S}_c^k with $k \neq j$. It follows that the quantities $(d^j \partial_{y^j}{}^t r^j \cdot r^1)(y)$ with $j \in \{2, 3\}$ can be interpreted as the principal curvatures of the surface $\mathcal{S}_{y^1}^1$, computed at the point x such that $y(x) = y$. By definition, the average of these two principal curvatures, that is $\mathcal{C}(y)$, is the mean curvature of $\mathcal{S}_{y^1}^1$ at $y^{-1}(y)$.

Another argument prevails concerning \mathcal{E} . The quantities $d_2(\partial_{y^2}{}^t r_3) \cdot r_1$ and $(d_3 \partial_{y^3}{}^t r_2) \cdot r_1$ can be viewed as the two extra-diagonal terms of the Gauss map $\Gamma : \mathcal{S}_c^1 \rightarrow \mathbb{S}^2$. Due to the symmetry of the Weingarten endomorphism, we are sure that $\mathcal{E} \equiv 0$. \square

From (2.33), it follows that $\bar{\mathfrak{z}}_0(t, y, w)$ satisfies:

$$(2.38) \quad \bar{\mathfrak{z}}_0^2 = y^2, \quad \bar{\mathfrak{z}}_0^3 = y^3, \quad \bar{\mathfrak{z}}_0^5 = w, \quad \bar{\mathfrak{z}}_0^6 = \mathbf{a}_w, \quad \forall t \in \mathbb{R},$$

together with a planar motion directed by:

$$(2.39) \quad \begin{cases} \partial_t \bar{\mathfrak{z}}_0^1 = w d^1(\bar{\mathfrak{z}}_0^1, y^2, y^3) \cos \bar{\mathfrak{z}}_0^4, & \bar{\mathfrak{z}}_0^1(0) = y^1, \\ \partial_t \bar{\mathfrak{z}}_0^4 = w c(w) \mathcal{C}(\bar{\mathfrak{z}}_0^1, y^2, y^3) \sin \bar{\mathfrak{z}}_0^4, & \bar{\mathfrak{z}}_0^4(0) = \varsigma_w. \end{cases}$$

2.2.4. *The reduced hamiltonian.* By shrinking the size of \mathcal{O} if necessary, we can always assume that the domain \mathcal{O} is given in the following form:

$$(2.40) \quad \mathcal{O} = \{y = {}^t(y^1, y^2, y^3) = {}^t(y^1, \hat{y}); y^1 \in]y_-^1(\hat{y}), y_+^1(\hat{y})[, \hat{y} := (y^2, y^3) \in \hat{\mathcal{O}}\}.$$

In (2.40), we assume that:

$$(2.41a) \quad \text{The set } \hat{\mathcal{O}} \subset \mathbb{R}^2 \text{ is open and bounded,}$$

$$(2.41b) \quad \text{The functions } y_{\pm}^1 \in \mathcal{C}^\infty(\hat{\mathcal{O}}; \mathbb{R}) \text{ are such that } y_-^1 \leq y_+^1.$$

Assumption 2.2. *[mean curvature with a unique oriented change of sign] The set \mathcal{O} can be described as in (2.40)-(2.41), and there is a smooth function $y_0^1 : \hat{\mathcal{O}} \rightarrow \mathbb{R}$ such that the zero set of \mathcal{C} is given by the graph of y_0^1 . In other words:*

$$(2.42) \quad \mathcal{C}^{-1}(0) = \{y; \mathcal{C}(y) = 0\} = \{(y_0^1(\hat{y}), \hat{y}); \hat{y} \in \hat{\mathcal{O}}\}.$$

Moreover, the sign of $\mathcal{C}(\cdot)$ changes from $-$ to $+$ when crossing $\mathcal{C}^{-1}(0)$:

$$(2.43) \quad \partial_{y^1} \mathcal{C}(y_0^1(\hat{y}), \hat{y}) > 0, \quad \forall \hat{y} \in \hat{\mathcal{O}}.$$

◦ Application 2 [change of sign along the equatorial plane]. For $k \in \{1, 2, 3\}$, introduce the surfaces $\mathcal{S}_{ec}^k := \{x \in \mathcal{O}_e; y_e^k(x) = c\}$. Let $\mathcal{C}_e(y)$ be the mean curvature of the surface \mathcal{S}_{ec}^1 , computed at the point x such that $y_e(x) = y$.

Lemma 2.5. *The function $\mathcal{C}_e(\cdot)$ does not depend on the variable y^3 . It is given by:*

$$(2.44) \quad \mathcal{C}_e(y^1, y^2) = \frac{3}{2} [1+3(R^2 y^1)^2]^{-3/2} [3+5(R^2 y^1)^2] R y^1, \quad R := (\rho^2 + z^2)^{1/2},$$

with R as in (2.12). Assumption 2.2 is verified with $y_{e0}^1 \equiv 0$. For all $y^2 \in \mathbb{R}_-^*$, the partial application $\mathcal{C}_e(\cdot, y^2)$ is odd on \mathbb{R} , and strictly increasing on \mathbb{R}_+^* .

Proof. Since the set \mathcal{S}_{ec}^1 is a surface of revolution with axis D_e , the function $\mathcal{C}_e(\cdot)$ does not depend on y^3 . The identity (2.44) comes from a computation based on (2.9) and on the classical identity $\mathcal{C}_e(y) = -2^{-1}(\operatorname{div}_x \tilde{r}_e^1) \circ y^{-1}(y)$. Let $a := R^2 y^1$. Exploit the relation (2.12) to extract:

$$\partial_{y^1} \mathcal{C}_e(y) = \frac{3}{2} \left\{ 2 a^2 \mathcal{G}'(a^2) + \frac{1+a^2}{1-a^2} \mathcal{G}(a^2) \right\} \frac{R^2 \sqrt{-y^2}}{1+3(y^1)^2 R^4}, \quad \mathcal{G}(\lambda) := \frac{(3+5\lambda)}{(1+3\lambda)^{3/2}}.$$

The quantity within the brackets is positive, so that $\partial_{y^1} \mathcal{C}_e(y) > 0$ with in particular:

$$(2.45) \quad \partial_{y^1} \mathcal{C}_e(0, y^2) = \frac{9}{2(-y^2)^{1/2}} > 0, \quad \partial_{y^1 y^2}^2 \mathcal{C}_e(0, y^2) = \frac{9}{4(-y^2)^{3/2}} > 0, \quad \forall y^2 \in \mathbb{R}_-^*. \quad \square$$

It is worth mentioning how to derive Assumption 2.2 from geometrical considerations. In view of (2.9), the surface \mathcal{S}_{e0}^1 is the plane $\{z = 0\}$. This gives $\mathcal{C}_e^{-1}(0) \supset \{z = 0\}$. When $c \neq 0$, the surface \mathcal{S}_{ec}^1 has everywhere two non-zero principal curvatures of the same sign, and therefore everywhere a non-zero mean curvature. This furnishes (2.42) with $y_0^1 \equiv 0$. Moreover, the surface \mathcal{S}_{ec}^1 is the mirror image of \mathcal{S}_{-c}^1 for the reflection with respect to the plane \mathcal{S}_{e0}^1 , indicating that $\mathcal{C}_e(\cdot, y^2)$ is an odd function. The convention adopted about the orientation implies that $y^1 \mapsto \mathcal{C}_e(y^1, y^2)$ is increasing, explaining the sign in (2.43). ◦

Definition 2.3. *Introduce the applications:*

$$(2.46) \quad \begin{aligned} \mathcal{P}_1 : \mathcal{O} &\longrightarrow \mathbb{R} & \mathcal{P}_2 :]0, \pi[\times]0, 1[&\longrightarrow \mathbb{R} \\ y &\longrightarrow \int_{y_0^1(\hat{y})}^{y^1} \frac{\mathcal{C}(s, \hat{y})}{d^1(s, \hat{y})} ds, & (\varsigma, \mathbf{u}) &\longrightarrow -\frac{\ln(\sin \varsigma)}{c(\mathbf{u})}. \end{aligned}$$

The sum \mathcal{H} of \mathcal{P}_1 and \mathcal{P}_2 is called the reduced hamiltonian:

$$(2.47) \quad \mathcal{H}(\mathfrak{z}) := \mathcal{P}_1(y) + \mathcal{P}_2(\varsigma, \mathbf{u}) = \int_{y_0^1(\hat{y})}^{y^1} \frac{\mathcal{C}(s, \hat{y})}{d^1(s, \hat{y})} ds - \frac{\ln(\sin \varsigma)}{c(\mathbf{u})}, \quad \mathfrak{z} = {}^t(y, \varsigma, \mathbf{u}, \mathbf{a}).$$

In contrast with the full hamiltonian associated to (1.5), the function \mathcal{H} is related to the modulation equation (2.27). Lemma 2.6 below says that it governs what happens at leading order during intermediate times $t \sim 1$. This is why \mathcal{H} is called the reduced hamiltonian.

Lemma 2.6. *Fix $(\hat{y}, \mathbf{w}) \in \hat{\mathcal{O}} \times]0, 1[$. The application $\mathcal{H}(\cdot, y^2, y^3, \cdot, \mathbf{w})$ plays the part of an energy function for the planar motion (2.39), meaning that:*

$$(2.48) \quad \partial_t(\mathcal{H} \circ \bar{\mathfrak{J}}_0)(t, y, w) = 0, \quad \forall (t, y, w) \in \mathbb{R}_+ \times \mathcal{O}.$$

Proof. Just compute:

$$\partial_t(\mathcal{H} \circ \bar{\mathfrak{z}}_0) = (\partial_{y^1} \mathcal{P}_1)(\bar{\mathfrak{z}}_0^1, y^2, y^3) \partial_t \bar{\mathfrak{z}}_0^1 + (\partial_\varsigma \mathcal{P}_2)(\bar{\mathfrak{z}}_0^4, w) \partial_t \bar{\mathfrak{z}}_0^4,$$

and use (2.38), (2.39) together with the definitions of \mathcal{P}_1 and \mathcal{P}_2 to find 0. \square

Select $\mathbf{u} \in]0, 1[$. The partial application $\mathcal{P}_2(\cdot, \mathbf{u}) :]0, \pi[\rightarrow \mathbb{R}$ is strictly convex. It has a global minimum at $\varsigma = \pi/2$, with $\mathcal{P}_2(\pi/2, \mathbf{u}) = 0$. This is a potential well. Define:

$$(2.49) \quad g(\varsigma, \mathbf{u}) := \operatorname{sgn} \left(\varsigma - \frac{\pi}{2} \right) \left(\frac{|\ln(\sin \varsigma)|}{c(\mathbf{u})} \right)^{1/2}, \quad g^2 \equiv \mathcal{P}_2, \quad (\varsigma, \mathbf{u}) \in]0, \pi[\times]0, 1[.$$

The function $g(\cdot, \mathbf{u}) :]0, \pi[\rightarrow \mathbb{R}$ is smooth and strictly increasing. Its inverse is denoted by $g_{\mathbf{u}}^{-1}$. Now what about \mathcal{P}_1 ? The structure of \mathcal{P}_1 depends on the properties of d^1 and \mathcal{C} . The condition (2.2) and the definition (2.3) imply that $d^1 > 0$. Thus, all confinement and geometrical properties of the flow (2.39) are encoded in the behaviour of \mathcal{C} .

Lemma 2.7. *Under Assumption 2.2, for all $\hat{y} \in \hat{\mathcal{O}}$, the application $\mathcal{P}_1(\cdot, \hat{y})$ is a potential well. There is a smooth and strictly increasing function $f(\cdot, \hat{y}) :]y_-^1, y_+^1[\rightarrow \mathbb{R}$, with inverse denoted by $f_{\hat{y}}^{-1}$, such that $f^2 \equiv \mathcal{P}_1$ and:*

$$(2.50) \quad \partial_{y^1} f(y) > 0, \quad f(y_0^1(\hat{y}), \hat{y}) = 0, \quad y = (y^1, \hat{y}) \in]y_-^1(\hat{y}), y_+^1(\hat{y})[\times \hat{\mathcal{O}}.$$

Proof. By definition:

$$\mathcal{P}_1(y_0^1(\hat{y}), \hat{y}) = 0, \quad (\partial_{y^1} \mathcal{P}_1)(y) = d^1(y)^{-1} \mathcal{C}(y), \quad d^1(y) > 0, \quad \forall y \in \mathcal{O}.$$

As a consequence of (2.42) and (2.43), we find that:

$$(\partial_{y^1} \mathcal{P}_1)(y_0^1(\hat{y}), \hat{y}) = 0, \quad (\partial_{y^1 y^1}^2 \mathcal{P}_1)(y_0^1(\hat{y}), \hat{y}) = d^1(y_0^1(\hat{y}), \hat{y})^{-1} (\partial_{y^1} \mathcal{C})(y_0^1(\hat{y}), \hat{y}) > 0.$$

Just take $f(y) := \operatorname{sgn}(y^1 - y_0^1(\hat{y})) \mathcal{P}_1(y)^{1/2}$. \square

The function $\mathcal{H} = f^2 + g^2$ has a strict minimum equal to zero, at $(y^1, \varsigma) = (y_0^1, \frac{\pi}{2})$. The level curves of \mathcal{H} are invariant sets for (2.39), diffeomorphic to a family of concentric circles. This indicates that the trajectories are periodic. This will be confirmed further.

◦ Application 3 [reason of the bouncing back] To our knowledge, the following explanation has not yet been produced in connection with Step (P2). This is why it is put forward.

Mathematical interpretation 1 (of the bouncing back). *The particles bounce back and forth between two mirror points because the mean curvature $\mathcal{C}_e(\cdot, y^2)$ of the surfaces which are orthogonal to the field lines changes its sign along the field lines $\mathcal{A}(y^2, y^3)$.* ◦

From now on, we will work under Assumption 2.2, with energy levels $\mathcal{H}(\mathfrak{z})$ below $h(\hat{y})^2$ where $h \in \mathcal{C}^\infty(\hat{\mathcal{O}}; \mathbb{R}_+)$ is such that:

$$(2.51) \quad 0 \leq h(\hat{y}) < \min \left(\limsup_{y^1 \rightarrow y_-^1(\hat{y})} |f(y^1, \hat{y})|; \limsup_{y^1 \rightarrow y_+^1(\hat{y})} f(y^1, \hat{y}) \right).$$

In view of (2.39), the amplitude \mathbf{u} of u influences the speeds at which the trajectories are traversed. Due to the relativistic correction $c(\mathbf{u})$ in (2.47), it also modifies the form of the trajectories. However, since \mathcal{P}_1 does not depend on \mathbf{u} , the projections in position of the trajectories are not affected by changes in the values of \mathbf{u} .

When $\mathbf{u} = 0$ or when $\mathcal{H}_e(\mathfrak{z}) = 0$, the particle is stationary. This situation has no dynamical interest. It can be avoided through the two conditions $0 < \mathbf{u}_m \leq \mathbf{u}$ and $0 < \mathcal{H}_m \leq \mathcal{H}_e(\mathfrak{z})$, imposed for some $\mathbf{u}_m \in]0, 1[$ and for some $\mathcal{H}_m \in \mathbb{R}_+^*$. The case $\mathbf{u} = 1$ is also not pertinent, because even the most energetic particles which are implied move below the speed of light. We will impose $\mathbf{u} \leq \mathbf{u}_M < 1$ for some \mathbf{u}_M with $\mathbf{u}_m < \mathbf{u}_M < 1$. All this amounts to focus on a bounded domain of the phase space of the following type:

$$(2.52) \quad \mathfrak{D} \equiv \mathfrak{D}(h) := \left\{ \mathfrak{z}; \mathfrak{z} = {}^t(y^1, \hat{y}, \varsigma, \mathbf{u}, \mathbf{a}), \hat{y} = {}^t(y^2, y^3) \in \hat{\mathcal{O}}, \right. \\ \left. 0 < \mathcal{H}_m \leq \mathcal{H}(\mathfrak{z}) < h(\hat{y})^2, 0 < \mathbf{u}_m < \mathbf{u} < \mathbf{u}_M < 1 \right\}.$$

Coming back in the variables (y, u) , and then in the original variables (x, v) , we get:

$$(2.53) \quad \Omega := \{ \Xi_1(\mathfrak{z}); \mathfrak{z} \in \mathfrak{D} \}, \quad \tilde{\Omega} := \{ (x, v); (y(x), O \circ y(x) v) \in \Omega \}.$$

The set $\tilde{\Omega}$ is the domain alluded in Theorem 1. Introduce also:

$$(2.54) \quad \check{\mathfrak{D}} := \{ \check{\mathfrak{z}} = {}^t(\mathfrak{h}, \hat{y}, \mathbf{u}); 0 \leq \mathfrak{h} < h(\hat{y}), \hat{y} \in \hat{\mathcal{O}}, \mathbf{u}_m < \mathbf{u} < \mathbf{u}_M \}.$$

◦ Application 4 [drawing the shape of Van Allen belts]. Consider a particle starting at time $t = 0$ from the position (y, w) . Define $\mathfrak{h} := \sqrt{\mathcal{H}_e(y, w)}$. During intermediate times, the trajectory of this particle is completely determined by $\check{\mathfrak{z}} = {}^t(\mathfrak{h}, \hat{y}, w)$. The value of \mathcal{H} is constant along the motion, equal to \mathfrak{h}^2 , but the values of \mathcal{P}_{e_1} and \mathcal{P}_{e_2} are not.

The potential well $\mathcal{P}_{e_1}(\cdot, y^2)$ is maximal when $|y^1|$ reaches its maximum value along the trajectory. This occurs when $y^1 = \pm y_m^1(y^2)$, with y_m^1 adjusted such that $y_m^1(y^2) = f_{ey^2}^{-1}(\mathfrak{h})$. This is when $\bar{\mathfrak{z}}_0^1(t) = \pm y_m^1$. Then, we have $\bar{\mathfrak{z}}_0^4(t) = \pi/2$. At this time t , there is no drift along the field line $\mathcal{F}(y^2, y^3)$. The particle is bouncing back.

The potential well $\mathcal{P}_{e_2}(\cdot, w)$ is maximal when $|\varsigma - \pi/2|$ reaches its maximum value along the trajectory. This occurs when $\varsigma = \varsigma_m$ (modulo π), with ς_m adjusted such that $\varsigma_m := g_w^{-1}(\mathfrak{h})$. This happens when $\bar{\mathfrak{z}}_0^1(t) = 0$, that is when the particle crosses the equatorial plane $\mathcal{E}q$. The value of ς_m is called the *equatorial pitch angle* by physicists.

To ensure that the particles will not strike the atmosphere (and therefore be lost), the velocity vectors $\bar{U}_0(t)$ must point outside the loss cone [22], meaning that $|\bar{U}_0^\perp(t)| \neq 0$, and most importantly, that the quotient $|\bar{U}_0^\parallel(t)|/|\bar{U}_0^\perp(t)| \equiv \cotan \varsigma$ is adequately bounded. This can be achieved through a control on the pitch-angles ς_m (away from 0 and π) or through a control on the mirror points y_m^1 . In fact, both y_m^1 and ς_m are completely determined by the value of $\mathfrak{h}_m := f_{ey^2}(y_m^1) = g_w(\varsigma_m)$. This brings to light the notion of extended Van Allen belt \mathcal{EBelt} , which is the region in position-velocity defined by:

$$(2.55) \quad \mathcal{EBelt} := \{ (y, \varsigma, \mathbf{u}, \mathbf{a}) \in \mathcal{O}_e \times \mathbb{T} \times [\mathbf{u}_m, \mathbf{u}_M] \times \mathbb{T}; \sqrt{\mathcal{H}_e(y, \varsigma, \mathbf{u})} \leq f_{ey^2} \circ h_e(y^2) \},$$

where the function $h_e(\cdot)$ was introduced in Paragraph 2.1.1.

Mathematical interpretation 2 (of the Van Allen belts). *The Van Allen belt \mathcal{Belt} may be seen as the spatial projection of an extended Van Allen belt \mathcal{EBelt} . The shape of \mathcal{EBelt} can be drawn by prescribing the values of \mathbf{u} and adequate bounds on the energy levels of the reduced hamiltonian \mathcal{H}_e , see (2.55).*

o

It is \mathcal{EBelt} which is pertinent from a dynamical point of view, for instance when dealing with localization properties of the initial data $(y, \varsigma_w, w, \mathbf{a}_w)$.

2.2.5. *The autonomous change of variables $\Xi_2(\mathfrak{z})$.* To avoid multiple notations, we will consider during this second step that the function \mathbf{A} obtained before, at the level of line (2.34), is the new starting point, and therefore rename it \mathcal{A} . Thus, in this paragraph 2.2.5, we note $z = {}^t(y, \varsigma, \mathbf{u}, \mathbf{a})$. For such a point z chosen in Ω , we can write:

$$(2.56) \quad (y^1, \varsigma) = (f_{\hat{y}}^{-1}(\mathfrak{h} \cos v), g_{\mathbf{u}}^{-1}(\mathfrak{h} \sin v)), \quad \mathfrak{h} := \sqrt{\mathcal{H}(y^1, \hat{y}, \varsigma, \mathbf{u})} \in [0, h(\hat{y})].$$

The system (2.39) reduces to $\partial_t \mathfrak{h} = 0$, together with a scalar equation on v , namely:

$$(2.57) \quad \partial_t v = w \mu_{\mathfrak{z}}(v), \quad \mu_{\mathfrak{z}}(v) := - \frac{\mathcal{C}(f_{\hat{y}}^{-1}(\mathfrak{h} \cos v), \hat{y})}{2 \mathfrak{h} \cos v} \frac{\cos g_{\mathbf{u}}^{-1}(\mathfrak{h} \sin v)}{\mathfrak{h} \sin v}.$$

Of course, the function $\mu_{\mathfrak{z}}(\cdot)$ is not defined when $\mathfrak{h} = 0$, $v = 0 \pmod{\pi}$ or $v = \frac{\pi}{2} \pmod{\pi}$, but it can be extended there by continuity. Then, it becomes a smooth function on the whole domain \mathbb{R} , which is periodic in v with period 2π . Moreover:

Lemma 2.8. *The speed of rotation μ is bounded and nowhere zero. More precisely:*

$$(2.58) \quad \exists (\mu_-, \mu_+) \in (\mathbb{R}_+^*)^2; \quad \mu_- \leq \mu_{\mathfrak{z}}(v) \leq \mu_+, \quad \forall (\mathfrak{z}, v) \in \check{\mathfrak{D}} \times \mathbb{R}.$$

Proof. By construction, $\mathcal{C}(f_{\hat{y}}^{-1}(\cdot), \hat{y})$ and $-\cos g_{\mathbf{u}}^{-1}(\cdot)$ are two strictly increasing functions, taking the value 0 at the point 0, with non zero derivatives at all points. It follows that $\mu_{\mathfrak{z}}(v) \in \mathbb{R}_+^*$ for all (\mathfrak{z}, v) . The uniform control in (2.58) comes from an argument of compactness. Indeed, $\check{\mathfrak{z}} \in \check{\mathfrak{D}}$ and $\check{\mathfrak{D}}$ is relatively compact in view of (2.41a) and (2.54). On the other hand, $v \in \mathbb{T}$ and the torus \mathbb{T} is simply compact. \square

The minoration inside (2.58) is very important in what follows. In particular, it allows to consider the function $1/\mu_{\mathfrak{z}}(v)$. This is again a bounded, strictly positive function, which is periodic in v with period 2π . For all $\check{\mathfrak{z}} \in \check{\mathfrak{D}}$, the application:

$$(2.59) \quad \mathfrak{K}_{\check{\mathfrak{z}}} : \mathbb{R} \longrightarrow \mathbb{R}, \quad \mathfrak{K}_{\check{\mathfrak{z}}}(v) := \int_0^v \frac{dr}{\mu_{\check{\mathfrak{z}}}(r)},$$

is a diffeomorphism. With Lemma 2.2 in mind, introduce the decomposition:

$$(2.60) \quad \mathfrak{K}_{\check{\mathfrak{z}}}(v) = k_{\check{\mathfrak{z}}} v + \mathfrak{K}_{\check{\mathfrak{z}}}^p(v), \quad k_{\check{\mathfrak{z}}} := \frac{1}{2\pi} \int_0^{2\pi} \frac{dr}{\mu_{\check{\mathfrak{z}}}(r)} \in [\mu_+^{-1}, \mu_-^{-1}],$$

where $\mathfrak{K}_{\check{\mathfrak{z}}}^p(\cdot)$ is a 2π -periodic function. The inverse of $\mathfrak{K}_{\check{\mathfrak{z}}}$, in the sense that $\mathfrak{K}_{\check{\mathfrak{z}}} \circ \mathfrak{K}_{\check{\mathfrak{z}}}^{-1}(\nu) = \nu$, can be put in the form:

$$(2.61) \quad \mathfrak{K}_{\check{\mathfrak{z}}}^{-1}(\nu) = k_{\check{\mathfrak{z}}}^{-1} \nu + \tilde{\mathfrak{K}}_{\check{\mathfrak{z}}}^p(\nu), \quad \tilde{\mathfrak{K}}_{\check{\mathfrak{z}}}^p(\nu + 2\pi k_{\check{\mathfrak{z}}}) = \tilde{\mathfrak{K}}_{\check{\mathfrak{z}}}^p(\nu), \quad \forall \nu \in \mathbb{R}.$$

Using $\mathfrak{K}_{\check{\mathfrak{z}}}$, it is possible to normalize the speed of rotation. To this end, it suffices to replace the variable v by $\nu := \mathfrak{K}_{\check{\mathfrak{z}}}(v)$. The motion on the invariant tori expressed in terms of the canonical coordinates (\mathfrak{h}, ν) becomes linear in the angle variable ν :

$$(2.62) \quad \partial_t \mathfrak{h} = 0, \quad \partial_t \nu = w.$$

Definition 2.4. [definition of Ξ_2] Take $\mathfrak{z} = {}^t(\mathfrak{h}, y^2, y^3, \mathbf{u}, \mathbf{a}, \nu)$ and define:

$$\begin{aligned} \Xi_2^1(\mathfrak{z}) &:= f_{\hat{y}}^{-1}(\mathfrak{h} \cos \mathfrak{K}_3^{-1}(\nu)), & \Xi_2^2(\mathfrak{z}) &:= y^2, & \Xi_2^3(\mathfrak{z}) &:= y^3, \\ \Xi_2^4(\mathfrak{z}) &:= g_{\mathbf{u}}^{-1}(\mathfrak{h} \sin \mathfrak{K}_3^{-1}(\nu)), & \Xi_2^5(\mathfrak{z}) &:= \mathbf{u}, & \Xi_2^6(\mathfrak{z}) &:= \mathbf{a}. \end{aligned}$$

Transforming (2.27) through $\Xi_1 \circ \Xi_2$ yields:

$$(2.63) \quad \partial_t \bar{\mathfrak{Z}}_0 - \bar{\mathbf{A}}_0(\bar{\mathfrak{Z}}_0) = 0, \quad \bar{\mathfrak{Z}}_0(0) = \bar{\mathfrak{Z}}_{i0}^{s2}, \quad \bar{Z}_0 = \Xi_1 \circ \Xi_2(\bar{\mathfrak{Z}}_0), \quad \bar{\mathbf{A}}_0 = {}^t(0, \dots, 0, \mathbf{u}).$$

The initial data $\bar{\mathfrak{Z}}_{i0}^{s2}$ can be determined through the relation $\Xi_1 \circ \Xi_2(\bar{\mathfrak{Z}}_{i0}^{s2}) = z = {}^t(y, w)$.

With $\mathfrak{h}_z := \sqrt{\mathcal{H}(y, \varsigma_w, w)}$ and $\check{\mathfrak{z}} = {}^t(\mathfrak{h}_z, \hat{y}, w)$, this means that:

$$(2.64) \quad \bar{\mathfrak{Z}}_{i0}^{s2} = (\mathfrak{h}_z, \hat{y}, w, \mathbf{a}_w, \nu_z), \quad \nu_z := \mathfrak{K}_3 \circ \arccos[\mathfrak{h}_z^{-1} f_{\hat{y}}(y^1)].$$

At this stage, the vector field $\bar{\mathbf{A}}_0$ is rectified. The components $\bar{\mathfrak{Z}}_0^j$ with $j \in \{1, 2, 3, 4, 5\}$ are fixed parameters, whereas $\bar{\mathfrak{Z}}_0^6(t)$ is simply linear: $\bar{\mathfrak{Z}}_0^6(t) = \bar{\mathfrak{Z}}_{i0}^6 + w t$.

2.2.6. *The time dependent change of variables $\Xi_3(t, \mathfrak{z})$.* Just modify the last component.

Definition 2.5. [definition of Ξ_3] Take $\mathfrak{z} = {}^t(\mathfrak{h}, y^2, y^3, \mathbf{u}, \mathbf{a}, \mathfrak{b})$ and define:

$$\Xi_3^j(t, \mathfrak{z}) := \mathfrak{z}^j, \quad \forall j \in \{1, \dots, 5\}, \quad \Xi_3^6(t, \mathfrak{z}) := \mathfrak{b} + w t.$$

This operation changes $\bar{\mathbf{A}}_0$ into $\bar{\mathbf{A}}_0 \equiv 0$. But it makes the variable t appear at the level of the function \mathbf{A} defined by (2.23), with $\Xi = \Xi_1 \circ \Xi_2 \circ \Xi_3$ and $\tilde{\Xi} \equiv 0$. Remarkably, since Ξ_2 is periodic in ν with period $2\pi k_3$, the linear growth of Ξ_3 in the variable t is converted to a periodic behaviour in t of $\Xi(\cdot, \mathfrak{z})$.

Proposition 2.1. [periodic properties of the solutions to the first modulation equation] Select $\bar{\mathfrak{Z}}_{i0}^{s1} = (y, \varsigma_w, w, \mathbf{a}_w) \in \mathfrak{D}$ with \mathfrak{D} as in (2.52). Define $\bar{\mathfrak{Z}}_{i0}^{s2} = (\mathfrak{h}_z, \hat{y}, w, \mathbf{a}_w, \nu_z)$ as indicated in (2.64). Mark the position $\check{\mathfrak{z}} := {}^t(\mathfrak{h}_z, \hat{y}, w) \in \check{\mathfrak{D}}$. Introduce $b_z := \nu_z$ as well as:

$$(2.65) \quad T \equiv T(\bar{\mathfrak{Z}}_{i0}^{s1}) \equiv T(\check{\mathfrak{z}}) := 2\pi k_3 w^{-1}.$$

The solution $\bar{\mathfrak{Z}}_0$ to (2.33), or equivalently to (2.39), remains for all times $t \in \mathbb{R}_+$ in the energy level $\mathcal{H}(\bar{\mathfrak{Z}}_{i0}^{s1})$. With $\mathfrak{h}_z = \sqrt{\mathcal{H}(\bar{\mathfrak{Z}}_{i0}^{s1})}$, it can be described by:

$$(2.66) \quad \bar{\mathfrak{Z}}_0(t) = (f_{\hat{y}}^{-1}(\mathfrak{h}_z \cos \mathfrak{K}_3^{-1}(\mathfrak{b}_z + w t)), y^2, y^3, g_w^{-1}(\mathfrak{h}_z \sin \mathfrak{K}_3^{-1}(\mathfrak{b}_z + w t)), w, \mathbf{a}_w).$$

It is periodic in time t , with a period T depending on $\check{\mathfrak{z}}$ as in (2.65). The domain Ω , see (2.53), is invariant under the flow induced by (2.39). Given $F \in \mathcal{C}^0(\mathfrak{D}; \mathbb{R})$, we have:

$$(2.67) \quad \langle F \circ \bar{\mathfrak{Z}}_0(\cdot, \bar{\mathfrak{Z}}_{i0}^{s1}) \rangle = \int_{-\pi}^{\pi} F(f_{\hat{y}}^{-1}(\mathfrak{h}_z \cos v), y^2, y^3, g_w^{-1}(\mathfrak{h}_z \sin v), w, \mathbf{a}) P_3(v) dv,$$

where $P_3(v) := (2\pi k_3)^{-1} \mu_3(v)^{-1}$ is a probability density function on the interval $[-\pi, \pi]$.

Proof. A preliminary remark is that all components of $\check{\mathfrak{z}} = {}^t(\mathfrak{h}_z, \hat{y}, w)$ are constants of the motion. The explicit formula (2.66) is an outcome of the preceding analysis. To verify the periodicity, use (2.61) to obtain:

$$\mathfrak{K}_3^{-1}(\mathfrak{b}_z + w(t+T)) = k_3^{-1}(\mathfrak{b}_z + w t + 2\pi k_3) + \tilde{\mathfrak{K}}_3^p(\mathfrak{b}_z + w t + 2\pi k_3) = \mathfrak{K}_3^{-1}(\mathfrak{b}_z + w t) + 2\pi.$$

To get (2.67), just perform the change of variables $v = \mathfrak{K}_3^{-1}(\mathfrak{b}_z + w t)$, and use the definition of k_3 in (2.60) to see that P_3 is indeed a probability. \square

The identity (2.67) says that the time integral of the periodic function $F \circ \bar{\mathfrak{z}}_0(t, \bar{\mathfrak{z}}_{i0}^{s1})$ over its period T can be converted into a spatial integral. This is because the flow induced by (2.39) is completely ergodic when restricted to the energy levels of \mathcal{H} . Thus, the time average of the observable F along a trajectory of energy \mathfrak{h} can be viewed as the microcanonical ensemble average of F . It turns out that the resulting expression $\langle F \circ \bar{\mathfrak{z}}_0(\cdot, \bar{\mathfrak{z}}_{i0}^{s1}) \rangle$ depends only on $\check{\mathfrak{z}}$. This allows to define a map:

$$(2.68) \quad \begin{aligned} \mathcal{A} : \mathcal{C}^0(\mathfrak{D}; \mathbb{R}) &\longrightarrow \mathcal{C}^0(\check{\mathfrak{D}}; \mathbb{R}) \\ F(\mathfrak{z}) &\longrightarrow \mathcal{A}(F)(\check{\mathfrak{z}}) := \langle F \circ \bar{\mathfrak{z}}_0(\cdot, \mathfrak{z}) \rangle. \end{aligned}$$

The integrand in (2.67) is the pull-back of F in the action-angle variables (\mathfrak{h}, v) . When this integrand does not depend on v , we simply recover $\mathcal{A}(F) \equiv F$. Otherwise, the computation of $\mathcal{A}(F)$ can be more complicated. For instance, if F is a function of the variable y^1 only, the action of \mathcal{A} on the observable F produces a function $\mathcal{A}(F)(\check{\mathfrak{z}})$ which can depend on all the coordinates in $\check{\mathfrak{z}}$. Below, we remark that, under symmetry conditions, simplifications can occur when computing $\mathcal{A}(F)(\check{\mathfrak{z}})$.

Lemma 2.9. *Let $F(y^1, \hat{y}, \varsigma, \mathbf{u}, \mathbf{a}) \in \mathcal{C}^0(\mathfrak{D}; \mathbb{R})$. Assume that the functions $\mathcal{C}(\cdot, \hat{y})$ and $d^1(\cdot, \hat{y})$ are respectively odd and even. Then, the formula (2.67) can be further simplified into:*

$$(2.69) \quad \mathcal{A}(F)(\check{\mathfrak{z}}) = -\frac{1}{\pi k_{\check{\mathfrak{z}}}} \int_0^{\mathfrak{h}} r \frac{E(F)(f_{\hat{y}}^{-1}(\sqrt{\mathfrak{h}^2 - r^2}), \hat{y}, g_w^{-1}(r), \mathbf{w}, \mathbf{a}_w)}{\mathcal{C}(f_{\hat{y}}^{-1}(\sqrt{\mathfrak{h}^2 - r^2}), \hat{y}) \cos g_w^{-1}(r)} dr,$$

where $E(F)$ is the even part of $F(\cdot, \hat{y}, \cdot, \mathbf{w}, \mathbf{a}_w)$ centered at the point $(0, \pi/2)$, that is:

$$(2.70) \quad E(F)(y^1, \varsigma) := F(y^1, \varsigma) + F(-y^1, \varsigma) + F(y^1, \pi - \varsigma) + F(-y^1, \pi - \varsigma).$$

Proof. In (2.67), decompose the domain of integration into:

$$[-\pi, \pi] = [-\pi, -\pi/2] \cup [-\pi/2, 0] \cup [0, \pi/2] \cup [\pi/2, \pi].$$

On each of these subdomains, change v into $r = \mathfrak{h}_z \sin v$, with adequate formulas to express $\mathfrak{h}_z \cos v$ in terms of r . This yields four integral contributions. Under the hypothesis of Lemma 2.9, the functions $f(\cdot, \hat{y})$ and $\mathcal{C}(f_{\hat{y}}^{-1}(\cdot), \hat{y})$ are odd. This allows to recombine the four integrals into a single integral, with $E(F)$ as in (2.70). \square

◦ Application 5 [description of Step (P2)]. From Lemma 2.1, we can easily deduce that the potential \mathcal{P}_{e1} does not depend on y^3 , so that $\mathcal{H}_e(\mathfrak{z}) = \mathcal{H}_e(y^1, y^2, \varsigma, \mathbf{u})$. For similar reasons, in the earth's context, the formulas (2.64) and (2.65) become simply:

$$(2.71a) \quad T_e(\mathfrak{z}) = T_e(y^1, y^2, \varsigma_w, \mathbf{w}) = 2 \pi k_{(\mathfrak{h}, y^2, \mathbf{w})} w^{-1}, \quad \mathfrak{h} = \sqrt{\mathcal{H}_e(y^1, y^2, \varsigma_w, \mathbf{w})},$$

$$(2.71b) \quad \nu_e(\mathfrak{z}) = \nu_e(y^1, y^2, \varsigma_w, \mathbf{w}) = \mathfrak{K}_{(\mathfrak{h}, y^2, \mathbf{w})} \circ \arccos [\mathfrak{h}^{-1} f_{y^2}(y^1)].$$

Moreover, the application $y^1 \mapsto \mathcal{P}_{e1}(y^1, y^2)$ is even in $y^1 \in \mathbb{R}$. Fix $(y^2, \mathbf{u}) \in \mathbb{R}_+^* \times [\mathbf{u}_m, \mathbf{u}_M]$. Select $\mathfrak{h} \in \mathbb{R}_+^*$, and consider the energy curve $\mathcal{C}_e(\mathfrak{h}) := \{(y^1, \varsigma); \mathcal{H}_e(y^1, y^2, \varsigma, \mathbf{u}) = \mathfrak{h}^2\}$. As already observed, the minimal and maximal values of y^1 along $\mathcal{C}_e(\mathfrak{h})$ correspond to mirror points. They are obtained for $\varsigma = \pi/2$, when $g_{\mathbf{u}}(\varsigma)$ is minimal, equal to 0. They are opposite, located at the positions $\pm y_m^1(y^2)$ with $y_m^1(y^2)$ determined by the two conditions:

$$y_m^1(y^2) \in \mathbb{R}_+^*, \quad \mathcal{H}_e(y_m^1(y^2), y^2, \pi/2, \mathbf{u}) = \mathfrak{h}^2.$$

The phase portrait associated to the dynamical system (2.39), drawn in the plane $\{(y^1, \varsigma)\}$ contains $\mathcal{C}_e(\mathfrak{h})$ as a trajectory. Starting at time $t = 0$ from a position (y^1, ς_w) on $\mathcal{C}_e(\mathfrak{h})$, the solution to (2.39) follows the curve $\mathcal{C}_e(\mathfrak{h})$ with a non zero derivative. It turns in a direct sense around the point $(0, \pi/2)$. The function $\bar{\mathfrak{z}}_0^1(\cdot)$ is given by the first component in (2.66). It decreases up to $-y_m^1$ if $\varsigma_w \geq \pi/2$. It increases up to y_m^1 if $\varsigma_w \leq \pi/2$. Both positions $-y_m^1$ and y_m^1 are turning points. The informations given below explain what happens next, starting from $(y_m^1, \pi/2)$.

Lemma 2.10. *[periodic properties and special features of the solutions to the geomagnetic first modulation equation] Let $\mathfrak{h} \in \mathbb{R}_+^*$ and $y_m^1 \in \mathbb{R}_+^*$ such that $\mathcal{H}_e(y_m^1, y^2, \pi/2, w) = \mathfrak{h}^2$. Consider the solution $(\bar{\mathfrak{z}}_0^1, \bar{\mathfrak{z}}_0^4)$ to (2.39), issued at time $t = 0$ from the position $(y_m^1, \pi/2)$. The function $\bar{\mathfrak{z}}_0^1(\cdot)$ is periodic with the period T given in (2.65). It is strictly decreasing on the interval $[0, T/2]$ with $\bar{\mathfrak{z}}_0^1(T/4) = 0$ and $\bar{\mathfrak{z}}_0^1(T/2) = -y_m^1$. It is strictly increasing on the interval $[T/2, T]$ with $\bar{\mathfrak{z}}_0^1(3T/4) = 0$ and $\bar{\mathfrak{z}}_0^1(T) = y_m^1$, and so on.*

Lemma 2.10 gives a mathematical justification to the step (P2), from which reference was made in the introduction. Indeed, the component $\bar{Z}_0^1(t) = \Xi_1^1 \circ \mathfrak{z}_0(t) = \bar{\mathfrak{z}}_0^1(t)$ oscillates around the position $y_0^1(\hat{y}) = 0$, which is in the equatorial plane $\mathcal{E}q$. \circ

2.3. The long time WKB calculus. From now on, we work with $\mathfrak{z} = {}^t(\mathfrak{h}, y^2, y^3, \mathbf{u}, \mathbf{a}, \mathbf{b})$. We use (2.25) with Ξ as in (2.29) and $\tilde{\Xi}$ as in (2.28) to transform (2.20) into (2.22). As a consequence of Section 2.2, we can assert that $\mathbf{A}_0 \equiv 0$. Knowing that $\mathbf{A}_0 \equiv 0$, we can turn to the study of (2.22). What follows is inspired by two-timing tools and by supercritical WKB techniques, in the spirit of [7].

2.3.1. Initialization. Given $N \in \mathbb{N}^*$, we seek the solution $(\mathfrak{z}_\varepsilon, \varphi_\varepsilon)$ to (2.22) in the form:

$$(2.72) \quad \begin{pmatrix} \mathfrak{z}_\varepsilon \\ \varphi_\varepsilon \end{pmatrix}(\tau) = \begin{pmatrix} \mathfrak{z}_\varepsilon^a \\ \varphi_\varepsilon^a \end{pmatrix}(\tau) + \varepsilon^N \begin{pmatrix} r_\varepsilon^{\mathfrak{z}} \\ r_\varepsilon^\varphi \end{pmatrix}(\tau), \quad \begin{pmatrix} \mathfrak{z}_\varepsilon^a \\ \varphi_\varepsilon^a \end{pmatrix}(\tau) = \begin{pmatrix} \mathfrak{z}_\varepsilon \\ \Phi_\varepsilon \end{pmatrix}\left(\tau, \frac{\tau}{\varepsilon}, \frac{\varphi_\varepsilon(\tau)}{\varepsilon}\right),$$

where the profiles \mathfrak{z}_ε and Φ_ε depend on the three variables $(\tau, t, \theta) \in \mathbb{R}_+ \times \mathbb{T}_T \times \mathbb{T}$. We assume that $(\mathfrak{z}_\varepsilon, \Phi_\varepsilon)$ is given by an asymptotic expansion of the following type:

$$(2.73) \quad \begin{pmatrix} \mathfrak{z}_\varepsilon \\ \Phi_\varepsilon \end{pmatrix}(\tau, t, \theta) = \varepsilon^{-1} \begin{pmatrix} 0 \\ \langle \bar{\Phi}_p \rangle \end{pmatrix}(\tau) + \sum_{j=0}^{N-1} \varepsilon^j \begin{pmatrix} \bar{\mathfrak{z}}_j \\ \bar{\Phi}_j \end{pmatrix}(\tau, t) + \sum_{j=2}^N \varepsilon^j \begin{pmatrix} \mathfrak{z}_j^* \\ \Phi_j^* \end{pmatrix}(\tau, t, \theta).$$

Taking into account (2.22) to replace $\partial_\tau \varphi_\varepsilon$, the derivative ∂_τ applied with the substitution performed in (2.72) leads to the differential action:

$$Op(\mathfrak{z}; \partial) \equiv Op(\varepsilon, t, \theta, \mathfrak{z}; \partial_\tau, \partial_t, \partial_\theta) := \partial_\tau + \varepsilon^{-1} \partial_t + \varepsilon^{-2} b(\varepsilon, t, \theta, \mathfrak{z}) \partial_\theta.$$

Plug the ansatz (2.72) into (2.22) to infer the constraint:

$$(2.74) \quad Op(\mathfrak{z}_\varepsilon; \partial) \begin{pmatrix} \mathfrak{z}_\varepsilon \\ \Phi_\varepsilon \end{pmatrix}(\tau, t, \theta) - \frac{1}{\varepsilon} \begin{pmatrix} \mathbf{A} \\ \mathbf{b} \end{pmatrix}(\varepsilon, t, \theta, \mathfrak{z}_\varepsilon) = \varepsilon^N \begin{pmatrix} \mathcal{R}^{\mathfrak{z}} \\ \mathcal{R}^\varphi \end{pmatrix}(\varepsilon, \tau, t, \theta) = O(\varepsilon^N).$$

Note that φ_ε is not at all present at the level of (2.74). Thus, when discussing (2.74), the knowledge of φ_ε is not needed to move forward.

2.3.2. *Modulation equations.* The first equation of (2.74), about \mathfrak{Z}_ε , is decoupled from the second equation, about Φ_ε . Let us start with the formal analysis concerning \mathfrak{Z}_ε . To this end, substitute the expansion (2.73) into (2.74) and collect the terms with ε^{-1} in factor. By averaging in θ , we obtain the first modulation equation:

$$(2.75) \quad \partial_t \bar{\mathfrak{Z}}_0 - \bar{A}_0(t, \bar{\mathfrak{Z}}_0) = 0, \quad \bar{A}_0 \equiv 0.$$

We recognize here the rectified version of (2.27), implying that $\bar{\mathfrak{Z}}_0(\tau, t) = \langle \bar{\mathfrak{Z}}_0 \rangle(\tau)$. On the other hand, the examination of the oscillating part in θ leads to:

$$\mathfrak{Z}_1^* = \mathfrak{b}_0(t, \bar{\mathfrak{Z}}_0)^{-1} (\partial_\theta^{-1} A_0^*)(t, \theta, \bar{\mathfrak{Z}}_0), \quad \mathfrak{b}_0(t, \mathfrak{z}) := c(u) b(\Xi^1(t, \mathfrak{z}), y^2, y^3).$$

Recall that $A_0^* \equiv 0$ as a consequence of (2.28). Therefore, we find $\mathfrak{Z}_1^* \equiv 0$, as it has been already anticipated in (2.73). Then, examine what appears with ε^0 in factor to find:

$$(2.76) \quad \partial_\tau \langle \bar{\mathfrak{Z}}_0 \rangle - A_1(t, \theta, \langle \bar{\mathfrak{Z}}_0 \rangle) + \mathfrak{b}_0(t, \langle \bar{\mathfrak{Z}}_0 \rangle) \partial_\theta \bar{\mathfrak{Z}}_2 + \partial_t \bar{\mathfrak{Z}}_1 = 0.$$

To compute $A_1(t, \theta, \mathfrak{z})$, come back to (2.23) in order to extract:

$$(2.77) \quad \begin{aligned} A_1(t, \theta, \mathfrak{z}) = D_\mathfrak{z} \Xi(t, \mathfrak{z})^{-1} [& \mathcal{A}_1(\theta, \Xi(t, \mathfrak{z})) + (\tilde{\Xi}(t, \theta, \mathfrak{z}) \cdot \nabla_\mathfrak{z}) \mathcal{A}_0(\theta, \Xi(t, \mathfrak{z})) \\ & - \mathcal{B}_1(t, \theta, \mathfrak{z}) \partial_\theta \tilde{\Xi}(t, \theta, \mathfrak{z}) - \partial_t \tilde{\Xi}(t, \theta, \mathfrak{z}) \\ & - (\tilde{\Xi}(t, \theta, \mathfrak{z}) \cdot \nabla_\mathfrak{z}) \mathfrak{b}_0(t, \mathfrak{z}) \partial_\theta \tilde{\Xi}(t, \theta, \mathfrak{z})], \end{aligned}$$

where $\tilde{\Xi}$ can be replaced as indicated in (2.28). Take the average of (2.76) in both t and θ to exhibit the *second modulation equation*:

$$(2.78) \quad \partial_\tau \langle \bar{\mathfrak{Z}}_0 \rangle - \langle \bar{A}_1 \rangle(\langle \bar{\mathfrak{Z}}_0 \rangle) = 0, \quad \langle \bar{\mathfrak{Z}}_0 \rangle(0) = \mathfrak{z}_{i0} \equiv \bar{\mathfrak{Z}}_{i0}^{s2}, \quad \Xi(0, \mathfrak{z}_{i0}) = {}^t(y, w).$$

The non linear differential equation (2.78) allows to determine $\bar{\mathfrak{Z}}_0(\tau, t, \theta) \equiv \langle \bar{\mathfrak{Z}}_0 \rangle(\tau)$ on some interval $[0, \mathcal{T}]$ with $\mathcal{T} \in \mathbb{R}_+^*$. To understand the long time dynamic ($t \sim \varepsilon^{-1}$ or $\tau \sim 1$), as it appears at leading order in the original variables $z = {}^t(y, u)$, the procedure is to pass through (2.78), and then to exploit (2.25), giving rise to:

$$(2.79) \quad \bar{Z}_0(\tau, t) = \Xi(t, \langle \bar{\mathfrak{Z}}_0 \rangle(\tau)), \quad \langle \bar{Z}_0 \rangle(\tau) = \langle \Xi(\cdot, \langle \bar{\mathfrak{Z}}_0 \rangle(\tau)) \rangle.$$

Be careful, nothing guarantees that the map $\mathfrak{z} \longrightarrow \langle \Xi(\cdot, \mathfrak{z}) \rangle$ is a diffeomorphism. Thus, it may be that a self-contained evolution equation on $\langle \bar{Z}_0 \rangle$ is not available. Also, there is no evident link between $\langle \bar{Z}_0 \rangle(0) = \langle \Xi(\cdot, \mathfrak{z}_{i0}) \rangle$ and $(y, w) = \Xi(0, \mathfrak{z}_{i0})$.

◦ Application 6 [description of step (P3)]. Come back to the Vlasov part of (1.1). The influence of (E, B) on the second modulation equation is displayed through the term \mathcal{A}_1 inside A_1 , see (2.19c) and (2.77). Even if $(E, B) \equiv (0, 0)$, the source $\langle \bar{A}_1 \rangle$ is likely to bring a non trivial contribution, by a mechanism issued from the fast dynamics.

Lemma 2.11. [*self-induced longitudinal drift in the case of the geomagnetic field*] Consider the equation (2.78) with data issued from the earth's context. Assume that $(E, B) \equiv (0, 0)$. Then, the corresponding solutions $\langle \bar{\mathfrak{Z}}_0 \rangle(\tau, \mathfrak{z}_{i0})$ are defined for all times $\tau \in \mathbb{R}_+$. More precisely, they are given by a transport at constant speed:

$$(2.80) \quad \langle \bar{\mathfrak{Z}}_0 \rangle(\tau, \mathfrak{z}_{i0}) = (\mathfrak{z}_{i0}^1, \mathfrak{z}_{i0}^2, \mathfrak{z}_{i0}^3 + \langle \bar{A}_1^3 \rangle(\mathfrak{z}_{i0}) \tau, \mathfrak{z}_{i0}^4, \mathfrak{z}_{i0}^5 + \langle \bar{A}_1^5 \rangle(\mathfrak{z}_{i0}) \tau, \mathfrak{z}_{i0}^6).$$

In (2.80), the coefficients $\langle \bar{A}_1^3 \rangle$ and $\langle \bar{A}_1^5 \rangle$ depend only on $(\mathfrak{z}_{i0}^1, \mathfrak{z}_{i0}^2, \mathfrak{z}_{i0}^4)$. Recall (2.31). The coordinates $(\mathfrak{z}_{i0}^1, \mathfrak{z}_{i0}^2, \mathfrak{z}_{i0}^4)$ can be expressed in terms of (y, w) according to:

$$(2.81) \quad \mathfrak{z}_{i0}^1 = \sqrt{\mathcal{H}_e}(y^1, y^2, \varsigma_w, w), \quad \mathfrak{z}_{i0}^2 = y^2, \quad \mathfrak{z}_{i0}^4 = w.$$

Besides, the long time mean motion, observed in the original variables ${}^t(y, u)$, satisfies :

$$(2.82) \quad (\langle \bar{Z}_0^1 \rangle, \langle \bar{Z}_0^2 \rangle, \langle \bar{Z}_0^3 \rangle, \langle \bar{Z}_0^4 \rangle)(\tau, y, w) = {}^t(0, y^2, y^3 + \langle \bar{A}_1^3 \rangle(\mathfrak{z}_{i0}) \tau, 0).$$

Proof. The access to $\langle \bar{\mathfrak{Z}}_0 \rangle(\tau, \mathfrak{z}_{i0})$ is via (2.78). Solving (2.78) requires to compute $\langle \bar{A}_1 \rangle$. To this end, note that the formula (2.77) applies in any of the stages of the transformations. For instance, we can use it after the first transformation Ξ_1 . Remark that $(E, B) \equiv (0, 0)$ means that $\mathcal{A}_1 \equiv 0$. With (2.28), these overviews provide:

$$(2.83) \quad \begin{aligned} \mathbf{A}_1 = \mathbf{b}_0^{-1} (D_3 \Xi_2)^{-1} [& (\partial_\theta^{-1} \mathbf{A}_0^* \cdot \nabla_{\mathfrak{z}}) \mathbf{A}_0 - \mathbf{b}_1 \mathbf{A}_0^* - \mathbf{b}_0 \partial_t (\mathbf{b}_0^{-1} \partial_\theta^{-1} \mathbf{A}_0^*) \\ & - (\partial_\theta^{-1} \mathbf{A}_0^* \cdot \nabla_{\mathfrak{z}}) (\ln \mathbf{b}_0) \mathbf{A}_0^*] (\theta, \Xi_2 \circ \Xi_3(t, \mathfrak{z})), \end{aligned}$$

where the quantities in the right hand side are the expressions which have been obtained just before Lemma 2.4. Thus, we work here with $\mathbf{b}_j := \mathcal{B}_j \circ \Xi_1$ for $j \in \{0, 1\}$, and with \mathbf{A}_0^* given by (2.35) and (2.36). In the earth's context, there are many simplifications coming from (2.11) and from the special forms of \mathcal{B}_0 and Ξ . The term \mathbf{b}_1 is simply:

$$\begin{aligned} \mathbf{b}_1(t, \theta, \mathfrak{z}) &= \tilde{\Xi}(t, \theta, \mathfrak{z}) \cdot (\nabla_z \mathcal{B}_0)(\Xi(t, \mathfrak{z})) + \mathcal{B}_1(\theta, \Xi(t, \mathfrak{z})) \\ &= \mathbf{u} \sin \varsigma \sin(\theta + \mathbf{a}) [d^2 \partial_{y^2} (\ln b) + d^3 (\partial_{y^3} {}^t r^3) \cdot r^2]. \end{aligned}$$

On the other hand, (2.36) reduces to:

$$\begin{pmatrix} \mathbf{A}_0^{4*} \\ \mathbf{A}_0^{6*} \end{pmatrix} = \mathbf{u} c(\mathbf{u}) \begin{pmatrix} d^1 (\partial_{y^1} {}^t r^2) \cdot r^1 \cos \varsigma \cos(\theta + \mathbf{a}) + \mathcal{D} \sin \varsigma [\cos^2(\theta + \mathbf{a})]^* \\ -d^1 (\partial_{y^1} {}^t r^2) \cdot r^1 \cos^2 \varsigma (\sin \varsigma)^{-1} \sin(\theta + \mathbf{a}) - \mathcal{D} \cos \varsigma \cos(\theta + \mathbf{a}) \sin(\theta + \mathbf{a}) \end{pmatrix}$$

with $\mathcal{D} := d^2 (\partial_{y^2} {}^t r^2) \cdot r^1 - d^3 (\partial_{y^3} {}^t r^3) \cdot r^1$. This provides useful information for computing the contribution \mathbf{A}_1 , as indicated in line (2.83). Many simplifications occur when averaging \mathbf{A}_1 with respect to θ . For instance, since \mathbf{b}_0 does not depend on θ , we find that:

$$\langle (D_3 \Xi_2)^{-1} \partial_t (\mathbf{b}_0^{-1} \partial_\theta^{-1} \mathbf{A}_0^*) \rangle \equiv \langle \partial_\theta^{-1} [(D_3 \Xi_2)^{-1} \partial_t (\mathbf{b}_0^{-1} \mathbf{A}_0^*)] \rangle \equiv 0.$$

More generally, all contributions which are linear or cubic in $\cos(\theta + \mathbf{a})$ and $\sin(\theta + \mathbf{a})$, as well as the products $\cos(\theta + \mathbf{a}) \sin(\theta + \mathbf{a})$, are canceled. There remains $\bar{\mathbf{A}}_1 = {}^t(0, 0, \bar{\mathbf{A}}_1^3, 0, \bar{\mathbf{A}}_1^5, 0)$ with (there is a similar formula concerning $\bar{\mathbf{A}}_1^5$):

$$\begin{aligned} \bar{\mathbf{A}}_1^3(t, \mathfrak{z}) &= \mathbf{u}^2 \left\{ \frac{1}{2} \left[\frac{d^2 d^3}{d^1} \partial_{y^2} \ln \left(\frac{d^3}{d^1} \right) - \frac{(d^3)^2}{d^1} (\partial_{y^3} {}^t r^3) \cdot r^2 \right] (y^1, y^2) \sin^2 \varsigma \right. \\ &\quad \left. + [d^3 (\partial_{y^1} {}^t r^2) \cdot r^1] (y^1, y^2) \cos^2 \varsigma \right\} | (y^1, \varsigma, \mathbf{u}) = (\Xi_2^1, \Xi_2^4, \Xi_2^5) \circ \Xi_3(t, \mathfrak{z}). \end{aligned}$$

The expressions $\bar{\mathbf{A}}_1^j$ with $j \in \{3, 5\}$ do not depend on $(\mathfrak{z}^3, \mathfrak{z}^5) \equiv (y^3, \mathbf{a})$. We can exploit (2.67) to compute the average of $\bar{\mathbf{A}}_1$ with respect to t . This allows to remove the dependence in $\mathfrak{z}^6 \equiv \mathbf{b}$. The coefficients $\langle \bar{\mathbf{A}}_1^3 \rangle$ and $\langle \bar{\mathbf{A}}_1^5 \rangle$ depend only on $(\mathfrak{z}_{i0}^1, \mathfrak{z}_{i0}^2, \mathfrak{z}_{i0}^4)$.

Then, consider the equation (2.79). Since $\langle \Xi^j \rangle(\mathfrak{z}) = \mathfrak{z}_{i0}^j = y^j$ for $j \in \{2, 3\}$, the result (2.82) about $\langle \bar{Z}_0^2 \rangle$ and $\langle \bar{Z}_0^3 \rangle$ is a direct consequence of the formulas in (2.79). To get the information (2.82) concerning $\langle \bar{Z}_0^1 \rangle$ and $\langle \bar{Z}_0^4 \rangle$, it suffices to apply (2.69) with respectively $F(y^1, \varsigma) = y^1$ and $F(y^1, \varsigma) = \cos \varsigma$. \square

It is easy to show that $\bar{A}_1^3(t, \mathfrak{z})$ is a sum of positive terms as long as $|z|/\rho \leq \arctan \sqrt{2}$. In practice, the expression $\langle \bar{A}_1^3 \rangle(\mathfrak{z})$ is positive in a much larger region. The long time mean motion is a signed rotation around the earth, localized in the equatorial plane $\mathcal{E}q$, with a sign depending on the charge of the selected particles (electrons or protons). This must be correlated with the creation of a ring current.

The contents of Lemma 2.11 are consistent with the observations reported in step (P3) of the introduction. The long time conservation of the "energy levels" \mathfrak{h} and of the "heights" y^2 , represented by $\langle \bar{\mathfrak{Z}}_0^1 \rangle(\tau, \mathfrak{z}_{i0})$ and $\langle \bar{\mathfrak{Z}}_0^2 \rangle(\tau, \mathfrak{z}_{i0})$ at the level of (2.80), further supports the stability of the extended Van Allen belt $\mathcal{E}Belt$ (which is delimited by values of \mathfrak{h} and y^2) and, consequently, of the Van Allen belt. \circ

2.3.3. *The induction.* Use the equation (2.76) to derive:

$$(2.84) \quad \bar{\mathfrak{Z}}_1^*(\tau, t) = (\partial_t^{-1} \bar{A}_1^*)(t, \langle \bar{\mathfrak{Z}}_0 \rangle(\tau)), \quad \mathfrak{Z}_2^*(\tau, t, \theta) = b_0(t, \langle \bar{\mathfrak{Z}}_0 \rangle)^{-1} \partial_\theta^{-1} A_1^*(t, \theta, \langle \bar{\mathfrak{Z}}_0 \rangle).$$

We proceed to see whether it is possible to find the \mathfrak{Z}_j for $j \geq 2$. We argue by induction, with the inductive hypothesis:

$$(H_j) : \begin{cases} \text{The expressions } \mathfrak{Z}_k \text{ are known on the domain } [0, \mathcal{T}] \times \mathbb{T}_T \times \mathbb{T} \text{ for } 1 \leq k \leq j-2, \\ \text{The expressions } \bar{\mathfrak{Z}}_{j-1}^* \text{ and } \mathfrak{Z}_j^* \text{ are known on their domains of definition.} \end{cases}$$

We have already checked the validity of (H_2) . Assume (H_j) for all $j \in \{2, \dots, N\}$. Then, consider what comes with ε^{j-1} in factor. This is a linearized version of (2.76):

$$(2.85) \quad \partial_\tau \langle \bar{\mathfrak{Z}}_{j-1} \rangle - (\langle \bar{\mathfrak{Z}}_{j-1} \rangle \cdot \nabla_{\mathfrak{z}}) A_1 + b_0 \partial_\theta \mathfrak{Z}_{j+1}^* + \partial_t \bar{\mathfrak{Z}}_j^* = \mathcal{G}_{j-1},$$

where \mathcal{G}_{j-1} is a known function. From there, the procedure to deduce (H_{j+1}) is a repetition of what has been done in Paragraph 2.3.2. There is just a point that must be stressed. By averaging (2.85) in t and θ , we recover the linear equation :

$$(2.86) \quad \partial_\tau \langle \bar{\mathfrak{Z}}_{j-1} \rangle - (\langle \bar{\mathfrak{Z}}_{j-1} \rangle \cdot \nabla_{\mathfrak{z}}) \langle A_1 \rangle = \langle \bar{\mathcal{G}}_{j-1} \rangle, \quad \langle \bar{\mathfrak{Z}}_{j-1} \rangle(0) = -\bar{\mathfrak{Z}}_{j-1}^*(0, 0) - \mathfrak{Z}_{j-1}^*(0, 0, 0).$$

It follows that the Cauchy problem (2.86) has a solution $\langle \bar{\mathfrak{Z}}_{j-1} \rangle$ on the whole interval $[0, \mathcal{T}]$. It means that the life span of all the \mathfrak{Z}_j is the same as the one of (2.78). Stop the induction when $j = N$. With the \mathfrak{Z}^j , build \mathfrak{Z}_ε as indicated in (2.73).

To identify the profile $\Phi_\varepsilon(\tau, t, \theta)$, consider the asymptotic expansion proposed in (2.73), and perform a formal analysis at the level of (2.74). The study of what comes in factor of ε^{-1} furnishes with $\bar{Z}_0(\tau, t) = \Xi(t, \langle \bar{\mathfrak{Z}}_0 \rangle(\tau)) = {}^t(\bar{Y}_0, \bar{U}_0)(\tau, t) \in \mathbb{R}^3 \times \mathbb{R}^3$:

$$(2.87) \quad \langle \bar{\Phi}_p \rangle(\tau) = c(w) \int_0^\tau \langle b(\bar{Y}_0) \rangle(s) ds, \quad \bar{\Phi}_0^* = c(w) \partial_t^{-1} [b(\bar{Y}_0)^*].$$

The contribution with ε^j in factor, with $j \in \{0, \dots, N\}$, gives rise to:

$$(2.88) \quad \partial_\tau \Phi_j + \partial_t \Phi_{j+1} + \sum_{k=0}^j b_k \partial_\theta \Phi_{2-k+j} - b_{j+1} = 0.$$

This must be completed by what comes the constraint $\varphi_\varepsilon(0) \equiv 0$ in (2.22), that is:

$$(2.89) \quad \langle \bar{\Phi}_j \rangle(0) = -\bar{\Phi}_j^*(0, 0) - \Phi_j^*(0, 0, 0).$$

All coefficients b_k have been previously identified. Exploiting (2.88) and (2.89), the Φ_j can be determined successively, for $j = 1$ to N . By piecing the Φ_j together, we get Φ_ε .

2.3.4. *Justification of the WKB approach.* Fix any integer $N \in \mathbb{N}^*$. As long as the solution to (2.22) exists, it can be put in the form (2.72). As a by-product of the formal analysis of Paragraph 2.3, we can rely on (2.74). Since $(R^{\mathfrak{z}}, R^\phi)$ is a smooth function on the compact set $[0, 1] \times [0, \mathcal{T}] \times \mathbb{T}_T \times \mathbb{T}$, it is uniformly bounded. Taking into account the initial data of (2.18) and the preceding adjustments, we are faced with:

$$(2.90) \quad \partial_\tau \begin{pmatrix} r_\varepsilon^{\mathfrak{z}} \\ r_\varepsilon^\phi \end{pmatrix} = \begin{pmatrix} \mathcal{R}^{\mathfrak{z}} \\ \mathcal{R}^\phi \end{pmatrix} \left(\varepsilon, \tau, \frac{\tau}{\varepsilon}, \frac{\varphi_\varepsilon}{\varepsilon}, r_\varepsilon^{\mathfrak{z}} \right), \quad \begin{pmatrix} r_\varepsilon^{\mathfrak{z}} \\ r_\varepsilon^\phi \end{pmatrix} (0) = - \begin{pmatrix} Z_N^* \\ \Phi_N^* \end{pmatrix} (0, 0, 0),$$

where we have introduced:

$$\begin{aligned} (\mathcal{R}^{\mathfrak{z}}, \mathcal{R}^\phi)(\varepsilon, \tau, t, \theta, r) &:= \varepsilon^{-N-2} [\mathfrak{b}(\varepsilon, t, \theta, \mathfrak{Z}_\varepsilon) - \mathfrak{b}(\varepsilon, t, \theta, \mathfrak{Z}_\varepsilon + \varepsilon^N r)] \partial_\theta(\mathfrak{Z}_\varepsilon, \Phi_\varepsilon) \\ &\quad + \varepsilon^{-N-1} [(\mathfrak{A}, \mathfrak{b})(\varepsilon, t, \theta, \mathfrak{Z}_\varepsilon + \varepsilon^N r) - (\mathfrak{A}, \mathfrak{b})(\varepsilon, t, \theta, \mathfrak{Z}_\varepsilon)] - (R^{\mathfrak{z}}, R^\phi)(\varepsilon, \tau, t, \theta). \end{aligned}$$

Recall that $\partial_\theta \mathfrak{Z}_\varepsilon = O(\varepsilon^2)$ and $\mathfrak{A}_0 \equiv 0$. Therefore, there is a constant $C \in \mathbb{R}_+^*$ such that:

$$\sup \{ |\mathcal{R}^{\mathfrak{z}}(\varepsilon, \tau, t, \theta, r)|; (\varepsilon, \tau, t, \theta, r) \in [0, 1] \times [0, \mathcal{T}] \times \mathbb{T}_T \times \mathbb{T} \times B(0, 1) \} \leq C (|r| + 1).$$

Note $\mathcal{T}_\varepsilon \in \mathbb{R}_+^*$ the life span associated to (2.22). If $\mathcal{T}_\varepsilon > \mathcal{T}$, rename $\mathcal{T}_\varepsilon = \mathcal{T}$. By applying Gronwall's lemma to the first line of (2.90), we get easily:

$$(2.91) \quad \sup \{ |r_\varepsilon^{\mathfrak{z}}(\tau)|; (\varepsilon, \tau) \in]0, 1] \times [0, \mathcal{T}_\varepsilon] \} < +\infty.$$

Recall that $\partial_\theta \Phi_\varepsilon = O(\varepsilon^2)$. Thus, the information (2.91) can be exploited to infer that:

$$(2.92) \quad \sup \{ |\varepsilon r_\varepsilon^\phi(\tau)|; (\varepsilon, \tau) \in]0, 1] \times [0, \mathcal{T}_\varepsilon] \} < +\infty.$$

This last estimate is not as good as (2.91) but it prevents the solution of (2.90) to explode. We can assert that $\mathcal{T}_\varepsilon \equiv \mathcal{T}$ for all $\varepsilon \in]0, \varepsilon_0]$, and that $r_\varepsilon^{\mathfrak{z}}$ is indeed a remainder. Concerning the component r_ε^ϕ , by implementing (2.92) at the level of (2.90), we observe a loss of precision by a fixed factor ε^{-1} . Briefly:

Lemma 2.12. *For all $\varepsilon \in]0, 1]$, the solution to the ode (2.22) is defined for all $\tau \in [0, \mathcal{T}]$. Moreover, with $(\mathfrak{Z}_\varepsilon, \Phi_\varepsilon)$ as in (2.73), and with a precision in L^∞ that is uniform with respect to $\tau \in [0, \mathcal{T}]$ and initial data $\mathfrak{z}_{i\varepsilon}$ in a compact set, we can assert that:*

$$(2.93) \quad \begin{pmatrix} \mathfrak{z}_\varepsilon \\ \varepsilon \varphi_\varepsilon \end{pmatrix} (\tau) - \begin{pmatrix} \mathfrak{Z}_\varepsilon \\ \varepsilon \Phi_\varepsilon \end{pmatrix} \left(\tau, \frac{\tau}{\varepsilon}, \frac{\varphi_\varepsilon(\tau)}{\varepsilon} \right) = O(\varepsilon^N).$$

2.3.5. *Description of the phases in long times.* We can now assert that the asymptotic behaviour (when ε goes to zero) of $(x_\varepsilon, v_\varepsilon)$ during long times is given by (2.72) and (2.73). This reveals the presence of two types of oscillations.

◇ Oscillations at frequencies of the order ε^{-1} coming from the periodic behaviour in t of the profiles. When incorporating the (smooth) dependence on the initial data (y, w) , the period T becomes a function of (y, w) . To study the effects of the fast variations in (y, w) , it is necessary to put them at a uniform scale, that is to normalize T (say at the value 2π). This operation makes a phase appear.

Definition 2.6. *[notion of secondary phase] The function:*

$$(2.94) \quad \psi_s : (\tau, y, w) \longmapsto \psi_s(\tau, y, w) := 2\pi \tau / T = \tau w / k_{\mathfrak{z}},$$

where $\check{\mathfrak{z}} = {}^t(h_z, \hat{y}, w)$ and $h_z = \sqrt{\mathcal{H}}(y, \varsigma_w, w)$, is called the *secondary phase*. The expressions \mathcal{H} , $k_{\mathfrak{h}}$ and ς_w are given by (2.47), (2.59) and (2.31). In (2.94), the (smooth) dependence on (y, w) has been incorporated.

Introducing $\check{\mathfrak{z}}_\varepsilon(\tau, \vartheta, \theta) := \mathfrak{z}_\varepsilon(\tau, T\vartheta/2\pi, \theta)$, we find that:

$$(2.95) \quad \mathfrak{z}_\varepsilon(\tau) = \check{\mathfrak{z}}_\varepsilon\left(\tau, \frac{\psi_s(\tau)}{\varepsilon}, \frac{\varphi_\varepsilon(\tau)}{\varepsilon^2}\right) + O(\varepsilon^N), \quad \check{\mathfrak{z}}_\varepsilon \in \mathcal{C}^\infty([0, \mathcal{T}] \times \mathbb{T}^2; \mathbb{R}^6). \quad \diamond$$

\diamond Oscillations at frequencies of the order ε^{-2} associated with φ_ε . The access to φ_ε is possible through the differential equation (2.22), which may seem little descriptive, or through the relation (2.72), which may sound nested. In fact, apart from the existence of φ_ε on $[0, \mathcal{T}]$, the analysis has so far left aside a precise description of what is φ_ε . Information can be obtained by freezing a leading part of φ_ε , and by absorbing the rest inside the profiles. This can be done by seeking φ_ε in the form:

$$(2.96) \quad \varphi_\varepsilon(\tau) = \frac{\langle \bar{\Phi}_p \rangle(\tau)}{\varepsilon} + \bar{\Phi}_0\left(\tau, \frac{\tau}{\varepsilon}\right) + \varepsilon \hat{\Phi}_\varepsilon^c\left(\tau, \frac{\tau}{\varepsilon}, \frac{\langle \bar{\Phi}_p \rangle(\tau)}{\varepsilon^2} + \frac{\bar{\Phi}_0(\tau, \varepsilon^{-1}\tau)}{\varepsilon}\right).$$

Then, the relation (2.72) on φ_ε is equivalent to:

$$(2.97) \quad \left[\hat{\Phi}_\varepsilon^c - \bar{\Phi}_1 - \sum_{j=1}^{N-1} \varepsilon^j \Phi_{j+1}(\theta + \hat{\Phi}_\varepsilon^c) \right]_{\theta = \frac{\langle \bar{\Phi}_p \rangle}{\varepsilon^2} + \frac{\bar{\Phi}_0}{\varepsilon}} = \varepsilon^{N-1} r_\varepsilon^\varphi.$$

Consider the condition:

$$(2.98) \quad \mathcal{I}(\varepsilon, \tau, t, \theta, \hat{\Phi}_\varepsilon) := \hat{\Phi}_\varepsilon - \bar{\Phi}_1(\tau, t) - \sum_{j=1}^{N-1} \varepsilon^j \Phi_{j+1}(\tau, t, \theta + \hat{\Phi}_\varepsilon) = 0.$$

Since $\partial_{\hat{\Phi}} \mathcal{I} = 1 + O(\varepsilon)$, the implicit Theorem applied to (2.98) allows to uniquely determine the function $\hat{\Phi}_\varepsilon(\tau, t, \theta)$. It says that $\hat{\Phi}_\varepsilon$ is smooth in $(\varepsilon, \tau, t, \theta) \in [0, 1] \times [0, \mathcal{T}] \times \mathbb{T}_T \times \mathbb{T}$. Involved in the context of the relation (2.97), with the estimate (2.92) on r_ε^φ in mind, it also guarantees that $\hat{\Phi}_\varepsilon^c - \hat{\Phi}_\varepsilon = O(\varepsilon^{N-2})$, so that:

$$(2.99) \quad \varphi_\varepsilon(\tau) = \frac{\langle \bar{\Phi}_p \rangle(\tau)}{\varepsilon} + \bar{\Phi}_0\left(\tau, \frac{\tau}{\varepsilon}\right) + \varepsilon \hat{\Phi}_\varepsilon\left(\tau, \frac{\tau}{\varepsilon}, \frac{\langle \bar{\Phi}_p \rangle(\tau)}{\varepsilon^2} + \frac{\bar{\Phi}_0(\tau, \varepsilon^{-1}\tau)}{\varepsilon}\right) + O(\varepsilon^{N-1}).$$

At this stage, reasoning modulo $O(\varepsilon^{N-1})$, resorting to φ_ε is no more necessary.

Definition 2.7. [notion of principal phase] Introduce:

$\psi_{p0}(\tau) := \langle \bar{\Phi}_p \rangle(\tau) \in \mathcal{C}^\infty([0, \mathcal{T}]; \mathbb{R})$, $\psi_{p1}(\tau, \vartheta) := \bar{\Phi}_0(\tau, T\vartheta/2\pi) \in \mathcal{C}^\infty([0, \mathcal{T}] \times \mathbb{T}; \mathbb{R})$, where the functions $\langle \bar{\Phi}_p \rangle$ and $\bar{\Phi}_0^*$ are given by (2.87), whereas the part $\langle \bar{\Phi}_0 \rangle$ is determined by solving $\partial_\tau \langle \bar{\Phi}_0 \rangle - \langle b_1 \rangle = 0$ with initial data $\langle \bar{\Phi}_0 \rangle(0) = -\bar{\Phi}_0^*(0, 0)$. The expressions ψ_{p0} and ψ_{p1} are smooth with respect to (y, w) . The application:

$$\begin{aligned} \psi_{p\varepsilon} : [0, \mathcal{T}] \times \Omega &\longrightarrow \mathbb{R} \\ (\tau, y, w) &\longmapsto \psi_{p\varepsilon}(\tau, y, w) := \psi_{p0}(\tau, y, w) + \varepsilon \psi_{p1}\left(\tau, y, w, \frac{\psi_s(\tau, y, w)}{\varepsilon}\right), \end{aligned}$$

is called the *principal phase*.

The behaviour of the family $\{\mathfrak{z}_\varepsilon\}_\varepsilon$ when ε goes to zero can be described through a multiscale and multiphase expansion. More precisely:

Lemma 2.13. Define $\hat{\mathfrak{Z}}_\varepsilon(\tau, \vartheta, \theta) := \tilde{\mathfrak{Z}}_\varepsilon(\tau, \vartheta, \theta + \hat{\Phi}_\varepsilon(\tau, T\vartheta/2\pi, \theta))$. Then:

$$(2.100) \quad \mathfrak{z}_\varepsilon(\tau) = \hat{\mathfrak{Z}}_\varepsilon\left(\tau, \frac{\psi_s(\tau)}{\varepsilon}, \frac{\psi_{p\varepsilon}(\tau)}{\varepsilon^2}\right) + O(\varepsilon^N), \quad \hat{\mathfrak{Z}}_\varepsilon \in \mathcal{C}^\infty([0, T] \times \mathbb{T}^2; \mathbb{R}^6).$$

Proof. Replacing $\hat{\Phi}_\varepsilon^c$ with $\hat{\Phi}_\varepsilon$ gives rise to a loss of ε^{N-2} . But this is compensated by the fact that $\partial_\theta \tilde{\mathfrak{Z}}_\varepsilon = O(\varepsilon^2)$. We still have (2.100) in L^∞ . \square

In comparison with (2.95), the progress in (2.100) is that $\psi_{p\varepsilon}$ is a well identified function. Still, the form of the phase $\psi_{p\varepsilon}$ is more complicated as usual. For instance, the (long) time derivative of $\psi_{p\varepsilon}$ is an oscillation of large amplitude. Indeed, recalling (2.87), we have:

$$(2.101) \quad \partial_\tau \psi_{p\varepsilon}(\tau, y, w) = c(w) b \circ \bar{Y}_0\left(\tau, y, w, \frac{\tau}{\varepsilon}\right) + \varepsilon (\partial_\tau \bar{\Phi}_0)\left(\tau, y, w, \frac{\tau}{\varepsilon}\right).$$

Using again (2.87), and (2.89) with $j = 0$ (knowing that $\Phi_0^* \equiv 0$), we get:

$$(2.102) \quad \psi_{p\varepsilon}(0, y, w) = \langle \bar{\Phi}_p \rangle(0) + \varepsilon \langle \bar{\Phi}_0 \rangle(0) + \varepsilon \bar{\Phi}_0^*(0, 0) = 0. \quad \diamond$$

◦ Application 7 [about the principal phase induced by the geomagnetic field]. From (2.101) and (2.102), it is easy to deduce that:

$$(2.103) \quad \psi_{p\varepsilon}(\varepsilon t, y, w) = \varepsilon c(w) \int_0^t b_e \circ \bar{Y}_0(0, y, w, s) ds + O(\varepsilon^2).$$

The principal contribution of (2.103), the one with ε in factor, has a specific role to play in Part 3. This is why it is examined below.

Lemma 2.14. Recall that $T_e(\mathfrak{h}, y^2, w)$ and $\nu_e(\mathfrak{h}, y^1, y^2, w)$ are the two smooth functions which can be determined through the explicit formulas (2.71a) and (2.71b). There exists a scalar function $\Gamma_e(\cdot)$ such that, for all $(\tau, t) \in [0, T] \times \mathbb{R}_+$:

$$(2.104) \quad b_e \circ \bar{Y}_0(\tau, y, w, t) = \Gamma_e(\mathfrak{h}, y^2, w, \nu_e + wt), \quad \mathfrak{h} = \sqrt{\mathcal{H}_e}(y^1, y^2, \varsigma_w, w).$$

Proof. Recall (2.11c) and that $\Xi_1^j(\mathfrak{z}) = y^j$ for $j \in \{1, 2\}$. By construction:

$$b_e \circ \bar{Y}_0(\tau, t) = b_e(\tilde{\mathfrak{J}}_0^1(\tau, t), \tilde{\mathfrak{J}}_0^2(\tau, t)), \quad (\tilde{\mathfrak{J}}_0^1, \tilde{\mathfrak{J}}_0^2)(\tau, t) := (\Xi_2^1 \circ \Xi_3, \Xi_2^2 \circ \Xi_3)(t, \langle \bar{\mathfrak{J}}_0 \rangle(\tau, \mathfrak{z}_{i0}))$$

where $\langle \bar{\mathfrak{J}}_0 \rangle(\tau, \mathfrak{z}_{i0})$ is given by (2.80). In the case of \tilde{B}_e , the applications $\Xi_2^1 \circ \Xi_3$ and $\Xi_2^2 \circ \Xi_3$, given by Definitions 2.4 and 2.5, do not imply the variables y^3 and \mathfrak{a} :

$$\Xi_2^1 \circ \Xi_3(t, \mathfrak{z}) = f_{y^2}^{-1}(\mathfrak{h} \cos \mathfrak{K}_{(\mathfrak{h}, y^2, w)}^{-1}(\mathfrak{b} + \mathfrak{u}t)), \quad \Xi_2^2 \circ \Xi_3(t, \mathfrak{z}) = y^2, \quad \mathfrak{z} = {}^t(\mathfrak{h}, y^2, y^3, \mathfrak{u}, \mathfrak{a}, \mathfrak{b}).$$

In view of (2.80), the four remaining components \mathfrak{h} , y^2 , \mathfrak{u} and \mathfrak{b} are stationary, so that:

$$\tilde{\mathfrak{J}}_0^1(\tau, t) = \tilde{\mathfrak{J}}_0^1(0, t) = f_{\mathfrak{z}_{i0}^2}^{-1}(\mathfrak{z}_{i0}^1 \cos \mathfrak{K}_{(\mathfrak{z}_{i0}^1, \mathfrak{z}_{i0}^2, \mathfrak{z}_{i0}^4)}^{-1}(\mathfrak{z}_{i0}^6 + \mathfrak{z}_{i0}^4 t)), \quad \tilde{\mathfrak{J}}_0^2(\tau, t) = \tilde{\mathfrak{J}}_0^2(0, t) = \mathfrak{z}_{i0}^2.$$

Finally, noting that $\mathfrak{z}_{i0} = \bar{\mathfrak{Z}}_{i0}^{s2}$ with $\bar{\mathfrak{Z}}_{i0}^{s2}$ as in (2.64), we can directly recover (2.104), with T_e and ν_e given by (2.71a) and (2.71b), and:

$$(2.105) \quad \Gamma_e(\mathfrak{h}, y^2, w, \tilde{\theta}) := b_e(f_{y^2}^{-1}(\mathfrak{h} \cos \mathfrak{K}_{(\mathfrak{h}, y^2, w)}^{-1}(\tilde{\theta})), y^2).$$

\square

As a by-product of Proposition 2.1, Lemma 2.9 (applied here with $F = b_e \circ \Xi_1 \equiv b_e$) and Lemma 2.14, with $\check{\mathfrak{z}} = (\check{\mathfrak{z}}^1, \check{\mathfrak{z}}^2, \check{\mathfrak{z}}^4) = (\mathfrak{h}, y^2, w)$, we find that:

$$(2.106) \quad \langle b_e \circ \bar{Y}_0(\tau, \cdot) \rangle = \langle \Gamma_\varepsilon \rangle(\mathfrak{h}, y^2, w) = \frac{-4}{\pi k(\mathfrak{h}, y^2, w)} \int_0^{\mathfrak{h}} \frac{r b_e(f_{y^2}^{-1}(\sqrt{\mathfrak{h}^2 - r^2}), y^2) dr}{\mathcal{C}(f_{y^2}^{-1}(\sqrt{\mathfrak{h}^2 - r^2}), y^2) \cos g_w^{-1}(r)}. \quad \circ$$

2.3.6. *Proof of Theorem 1.* Theorem 1 is a compilation of results obtained in Section 2. It makes reference to the domains $\mathcal{B}\bar{e}lt$ and $\mathcal{E}\mathcal{B}\bar{e}lt$, which come from interpretations of (2.10) and (2.55) in terms of the variables (x, v) . More precisely:

$$(2.107) \quad \mathcal{B}\bar{e}lt := \{x; y(x) \in \mathcal{B}elt\}, \quad \mathcal{E}\mathcal{B}\bar{e}lt := \{(x, v); (y(x), \varsigma, |v|, \mathbf{a}) \in \mathcal{E}\mathcal{B}elt\}.$$

The Assertion 1) was discussed at the very end of Paragraph 2.2.4, see the Mathematical interpretation 2. How the set $\mathcal{B}\bar{e}lt$ can be adjusted to fit with concrete observations is explained in Paragraph 2.1.1, just before Lemma 2.1.

The Assertion 2) brings information about the compactness and the stability of $\mathcal{E}\mathcal{B}\bar{e}lt$. Since the application $y(\cdot)$ is a diffeomorphism and $O \circ y(x)$ is an orthogonal matrix, it suffices to argue at the level of $\mathcal{E}\mathcal{B}elt$. As stated in Lemma 2.6, the set $\mathcal{E}\mathcal{B}elt$ is preserved for all intermediate times $t \in \mathbb{R}$ by the flow associated to the first modulation equation (2.27), that is under the action of $\Xi(t, \cdot)$ for all $t \in \mathbb{R}$. In view of (2.80), it is also invariant under the flow induced by the second modulation equation (2.78), that is under the action of $\langle \bar{\mathfrak{Z}}_0 \rangle(\tau, \cdot)$ for all $\tau \in [0, \mathcal{T}]$. Considering (2.79), nothing is changing at this level.

To reach the exact flow $(x_\varepsilon, v_\varepsilon)(\tau, \cdot)$, see (2.72), the extra terms of the expansion (2.73) and the remainder $\varepsilon^N r_\varepsilon^3$ must be incorporated. On the one hand, the profiles $\bar{\mathfrak{Z}}_j$ and $\bar{\mathfrak{Z}}_j^*$ with $j \in \mathbb{N}^*$ are bounded functions (with at least ε in factor). On the one hand r_ε^3 is controlled as indicated in (2.91). This means that $(x_\varepsilon, v_\varepsilon)$ remains at a distance $\sim \varepsilon$ of $\mathcal{E}\mathcal{B}\bar{e}lt$.

To justify the part 3) of Theorem 1, the starting point is Lemma 2.13. The matter is to follow how the asymptotic expansion (2.100) is transformed when coming back to the original variables (x, v) . This can be done first by passing from \mathfrak{z}_ε to $z_\varepsilon = {}^t(y_\varepsilon, u_\varepsilon)$ through (2.25), and then by using the inverse transformation:

$$(2.108) \quad (x_\varepsilon, v_\varepsilon)(\tau, x, v) = (y^{-1}(y_\varepsilon), {}^tO(y_\varepsilon) e^{\varphi_\varepsilon \Lambda / \varepsilon} u_\varepsilon)(\tau, y(x), O \circ y(x) v).$$

The application $\Xi(t, \cdot) + \varepsilon \tilde{\Xi}(t, \theta, \cdot)$ is smooth, with uniform estimates in $(t, \theta) \in \mathbb{T}_T \times \mathbb{T}$. On the other hand, since the matrix Λ is skew-symmetric, the action of $e^{\varphi_\varepsilon \Lambda / \varepsilon}$ is unitary. It is not at these levels that difficulties appear. The problem comes when replacing φ_ε at the level of (2.108) by the expression of φ_ε obtained line (2.99), in order to make $\psi_{p\varepsilon}$ appear. This introduces a loss of precision by a factor ε^{-2} . Still, by choosing N large enough, the approximation (1.7) becomes valid for arbitrary powers of ε . The nonlinear geometric optics is therefore completely justified in L^∞ on the long time interval $[0, \mathcal{T}]$.

The last point to highlight is how to keep track of the initial data ${}^t(x, v)$ in (1.5). As already mentioned, there is a smooth dependence of \mathfrak{z}_ε on $\mathfrak{z}_{i\varepsilon}$, when solving (2.22). The aim of (2.81) was to express $\mathfrak{z}_{i\varepsilon}$ in function of ${}^t(y, w)$, resulting in Definitions 2.6 and 2.7. Now, to recover (1.7), it suffices to take:

$$(2.109) \quad \psi_{p\varepsilon}^\varepsilon(\tau, x, v) := \psi_{p\varepsilon}(\tau, y(x), O \circ y(x) v), \quad \psi_s^\varepsilon(\tau, x, v) := \psi_s(\tau, y(x), O \circ y(x) v).$$

3. PROPAGATION OF SINGULARITIES FOR THE MAXWELL PART

Our aim in this Part 3 is to get a qualitative mechanism to explain what contains the spectrograms, and especially the aspects underlined in line (C3) of Subsection 1.2, see the [picture](#) displayed in [1] and the [figure 1](#), p. 623 of [20]. From this point of view, the subsequent analysis is convincing, although it is still at a preliminary stage. It must be seen as a critical stepping stone towards more refined studies. To begin with, it is important to further discuss the model (1.1).

3.1. From Vlasov-Maxwell equations to oscillatory integrals. The first thing to do is to detail the origin of (1.1), starting from the standard formulation.

3.1.1. *The model.* Relativistic Vlasov-Maxwell equations can be written [18, 26]:

$$(3.1) \quad \begin{cases} \partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + q_e (\mathbf{E} + \mathbf{v} \times \mathbf{B} + \mathbf{v} \times \tilde{\mathbf{B}}_e) \cdot \nabla_{\mathbf{p}} f = 0, \\ \epsilon_r \partial_t \mathbf{E} - c^2 \operatorname{curl} \mathbf{B} = -\frac{q_e}{\epsilon_0} \int_{\mathbb{R}^3} f \mathbf{v} \, dp, & \operatorname{div}(\epsilon_r \mathbf{E}) = \frac{q_e}{\epsilon_0} \int_{\mathbb{R}^3} f \, dp, \\ \partial_t \mathbf{B} + \operatorname{curl} \mathbf{E} = 0, & \operatorname{div} \mathbf{B} = 0, \end{cases}$$

where $q_e \simeq 1,6 \times 10^{-19} C$ is the charge of the electron, and $\epsilon_0 \simeq 8,8 \times 10^{-12} F m^{-1}$ is the vacuum permittivity. The positive definite matrix ϵ_r is the relative permittivity. The letter $\tilde{\mathbf{B}}_e$ stands for the dipole model of the Earth's magnetic field, described in Application 1. With $\tilde{B}_e(\cdot)$ as in (2.7), it can be approximated by:

$$(3.2) \quad \tilde{\mathbf{B}}_e(\mathbf{x}) = \tilde{b}_0 \tilde{B}_e(\mathbf{R}^{-1} \mathbf{x}), \quad \tilde{b}_0 \simeq 3,1 \times 10^{-5} T, \quad \mathbf{R} \simeq 6,37 \times 10^6 m.$$

The relativistic framework comes from the presence of the energetic electrons contained in the [outer radiation belt](#). The velocity \mathbf{v} is limited by the speed of light $c \simeq 3 \times 10^8 ms^{-1}$, and it is linked to the relativistic momentum \mathbf{p} through the relations:

$$(3.3) \quad \mathbf{v}(\mathbf{p}) = \frac{\mathbf{p}}{m_e} \left(1 + \frac{|\mathbf{p}|^2}{m_e^2 c^2}\right)^{-1/2}, \quad \mathbf{p}(\mathbf{v}) = m_e \mathbf{v} \left(1 - \frac{|\mathbf{v}|^2}{c^2}\right)^{-1/2},$$

where $m_e \simeq 9,1 \times 10^{-31} kg$ is the mass of the electron. It follows that:

$$\nabla_{\mathbf{p}} = \frac{1}{m_e} \left(1 - \frac{|\mathbf{v}|^2}{c^2}\right)^{1/2} \left[\nabla_{\mathbf{v}} - \nabla_{\mathbf{v}} \left(\frac{\mathbf{v}}{c} \otimes \frac{\mathbf{v}}{c}\right) \right], \quad |D_{\mathbf{v}} \mathbf{p}(\mathbf{v})| = m_e^3 \left(1 - \frac{|\mathbf{v}|^2}{c^2}\right)^{-5/2}.$$

The solution of (3.1) can be expressed in the form:

$$(3.4) \quad \begin{pmatrix} f \\ \mathbf{E} \\ \mathbf{B} \end{pmatrix}(\mathbf{t}, \mathbf{x}, \mathbf{v}) = \begin{pmatrix} a_f f \\ \tilde{b}_0 a_m E \\ c^{-1} \tilde{b}_0 a_m B \end{pmatrix} \left(\frac{\mathbf{t}}{\mathbf{T}}, \frac{\mathbf{x}}{\mathbf{R}}, \frac{\mathbf{v}}{c} \right), \quad t := \frac{\mathbf{t}}{\mathbf{T}}, \quad x := \frac{\mathbf{x}}{\mathbf{R}}, \quad v := \frac{\mathbf{v}}{c}.$$

In (3.4), the multiplication by c^{-1} is aimed to recover the classical symmetric form of the Maxwell equations. The factor \tilde{b}_0 is introduced to make clear, through the multiplication by the dimensionless coefficient a_m (with "m" for magnetic), the comparison of the external magnetic field $\tilde{\mathbf{B}}_e$ with the self-consistent electromagnetic field (E, B) . Fix the observation time \mathbf{T} such that $\mathbf{T} := c^{-1} \mathbf{R} \simeq 10^{-2} s$.

and (in case **0b**) a bimodal Maxwellian distribution peaked around opposite speeds:

$$(3.9) \quad f_1(\mathbf{v}) := \frac{1}{2\pi^{3/2}\alpha^3} \left[\exp\left(-\frac{(\mathbf{v}_{\parallel} - \beta)^2}{\alpha^2}\right) + \exp\left(-\frac{(\mathbf{v}_{\parallel} + \beta)^2}{\alpha^2}\right) \right] \exp\left(-\frac{|\mathbf{v}_{\perp}|^2}{\alpha^2}\right).$$

The solutions $f(\nu, t, \mathbf{x}, \mathbf{v})$ of (3.1), which are issued from initial data like $(f_0 + \nu f_1)(\mathbf{v})$ can always be put in the form $f(\nu, t, \mathbf{x}, \mathbf{v}) = f_0(\mathbf{v}) + \nu f_1^l(t, \mathbf{x}, \mathbf{v}) + O(\nu^2)$ with $f(0, t, \mathbf{x}, \mathbf{v}) = f_0(\mathbf{v})$ where f_1^l is the solution to the equations (3.1) linearized around $(f_0, 0, 0)$. Numerical studies [18] indicate that the time behaviour of f_1^l might undergo *weibel* instabilities (in case **0a**) or *two-stream* instabilities (in case **0b**). These two types of instabilities are among the factors which are put forward to account for the sudden onset of whistler wave phenomena. However, there are some potential drawbacks to this approach:

- 1a. Contrary to the observations (presence of well located Van Allen belts), objects like f_0 and most importantly f_1 , which are derived from Maxwellian distributions, are not well localized in space;
- 1b. Neither f_0 nor f_1 take into account the oscillations induced by the Vlasov part, and especially the occurrence of several different scales, revealed at the level of (1.1) by the nondimensionalization of the equations;
- 1c. In connection with the items a) and b), the linearization procedure of (3.1) along f_0 emphasises the instabilities induced by the term $(\dot{\mathbf{E}} + \mathbf{v} \times \dot{\mathbf{B}}) \cdot \nabla_{\mathbf{v}} f_0$. The corresponding effects, although interesting, are not sure to prevail.
- 1d. This perturbation theory fails to really grasp the influence of the external field \tilde{B}_e , which is in fact at the core of the subject.

3.1.3. *A new approach to whistler waves.* Other aspects deserve to be developed in addition to the foregoing. Let us move on to another interpretation of what occurs. Instead of working at the level of (3.1), with a ground state $f_0(\mathbf{v})$ and its perturbations, it seems more realistic to consider the system (1.1) in order to construct solutions, or approximate solutions, which remain conveniently localized over a long period of time, till $\tau \sim 1$. The spatial projection of the support of $f(t, \cdot)$ should be contained in a compact neighbourhood of the Van Allen belts. The analysis of Part 2 allows to go in that direction.

When discussing localization properties of the density f , the first stage is to describe what happens at the time $t = 0$. In other words, clarifications are needed as to the choice of the initial data f in (1.4). At this level, physical aspects must be taken into account. As a by-product of a solar flare, less than half of the earth is lit by the sun, say at time $t = 0$. In these circumstances, the support of f in the direction y^3 is expected to be contained in some closed interval J of length less than π .

The Van Allen belt occupies a non-zero finished volume in the magnetosphere. This can be achieved only by allowing variations of all variables y^1 , y^2 and y^3 . As already explained in Application 4, from a dynamical point of view, the propagation is organized by energy levels. Now, at a constant energy level \mathfrak{h} , variations in y^1 are linked with variations in ς . And finally, it is more convenient to argue with the variables \mathfrak{h} , y^2 and y^3 , which lead to an adequate localization near the extended Van Allen belt.

During a whistler event, the kinetic energy E_c of the fast-moving particles is presumably concentrated near a specific value. Since E_c depends on the amplitude \mathbf{u} of the velocity, this furnishes a physical argument to work near a particular value w of \mathbf{u} . The values which are put forward concerning E_c are typically of the order of $\simeq 1$ Mev. To provide with some orders of magnitude, $\mathbf{u} \simeq 0,95c$ is equivalent to $E_c \simeq 1,69$ Mev and $c(\mathbf{u}) \simeq 0,33$. There is also a mathematical reason for selecting a special value w of \mathbf{u} . Indeed, the amplitude \mathbf{u} of the velocity is a conserved quantity as long as $E \equiv 0$. And this allows technically to examine separately what happens in the case of different values of \mathbf{u} . Consider the system (1.1) written in the variables $(y, \varsigma, \mathbf{u}, \mathbf{a})$. The above points can be taken into account by adjusting the initial data f in the form:

$$(3.10) \quad f(y, \varsigma, \mathbf{u}, \mathbf{a}) = \tilde{f}(\sqrt{\mathcal{H}_e}(y^1, y^2, \varsigma, \mathbf{u}), y^2, y^3) \delta_w(\mathbf{u}), \quad \tilde{f} \in \mathcal{C}_0^\infty(\mathbb{R}_+ \times \mathbb{R}_-^* \times \mathbb{R}; \mathbb{R}).$$

The control on \mathcal{H}_e imposed through (3.10) gives rise to a bound on y^1 . Thus, the support of $f(\cdot)$ is compact. To be more precise, it is convenient to select the values (\mathfrak{h}, y^2) coming from the identification of the extended Van Allen belt. To this end, it suffices to impose:

$$(3.11) \quad \text{supp } \tilde{f}(\cdot) \subset K_{\tilde{f}} := \{(\mathfrak{h}, y^2, y^3); 0 \leq \mathfrak{h} \leq f_{ey^2}(h_e(y^2)), y^2 \in \text{supp } h_e(\cdot), y^3 \in J\}.$$

Apply Theorem 1 in the case $(E, B) \equiv (0, 0)$. Then, the formula (1.4) with f as in (3.10) provides access to some solution f_ε of the Vlasov part of (1.1). Moreover, in view of Application 6, the two quantities \mathcal{H}_e and y^2 are preserved, whereas y^3 undergoes some rotation. This is coherent with the description of a ring current inside the Van Allen belt. The use of this expression f_ε instead of $f_0(v)$ has the following advantages:

- 2a. The expression f_ε is well localized in space and also in velocity;
- 2b. The expression f_ε inherits all the characteristics forced by the fast dynamics of the Vlasov part, including above each position $x \in \mathbb{R}^3$ (with varying parameters along the field lines), some thermal anisotropy and two-stream effects;
- 2c. The geometric features of \tilde{B} run through all the analysis of Part 2, leading finally to the refined notions of principal and secondary phases.

That being said, working with f_ε raises many new difficulties. A first problem stems from the fact that $(f_\varepsilon, 0, 0)$ is not at all a solution to (1.1). The situation is even worse, since the function f_ε is not intended to be the leading part of an approximate solution to the system (1.1). Still, the oscillations involved in f_ε are pertinent. This is a rather subtle aspect that we need to address below.

The first instabilities to appear are not linked with the emergence of (E, B) . Instead, they are due to the singular factors ε^{-2} and ε^{-1} at the level of (1.1). It follows that the solution operator associated with (1.1) is very sensitive to changes introduced in the initial data. Such considerations are usual in non linear geometric optics. The reader is referred to [19] for a general discussion and concrete examples. To get around this, the filtering methods of Part 2 separate the scales. They bring to light modulation equations, see Paragraph 2.3.2, which are well-posed equations providing information on the intermediate and long time dynamics. The fact that ψ_s does not depend on (E, B) , and that the components ψ_{p0} and ψ_{p1} of $\psi_{p\varepsilon}$ depend on (E, B) in a smooth way, is of particular interest.

The structures of the phases and of the profiles are quite stable under the influence of (E, B) . This is a basic point because this justifies why the detected phases have a full meaning. In this article, our aim is to give a straightforward access to the major aspects involved during the creation of whistler waves. From this point of view, it just happens that the knowledge of ψ_s and $\psi_{p\varepsilon}$ is a great help.

The Part 2 reveals that the instabilities come from the substitution of the fast variables θ and ϑ for the principal and secondary phases. A small modification can indeed be amplified by a factor ε^{-2} or ε^{-1} . This is why the solutions to the Vlasov part, obtained after substitution, are difficult to control. After the present study, the question will remain of finding a complete WKB expansion. Beyond, a whole range of stability issues arises. The considerations of Paragraph 3.1.2 should come up again, in a different context. Also, in connection with the problem of artificial whistler emissions, a study of resonant interaction of oscillations should be required, in the spirit of what has been done in [17]. All these items will be addressed in forthcoming publications. Before going that far in the analysis, there is an essential step. It is to consider that f_ε is a given function and from this, to examine the electromagnetic wave $(E_\varepsilon, B_\varepsilon)$ generated by the linear Cauchy problem:

$$(3.12) \quad \begin{cases} \partial_\tau(\epsilon_r E_\varepsilon) - \varepsilon^{-1} \operatorname{curl} B_\varepsilon = -\varepsilon^{-1} \kappa j_\varepsilon, & \operatorname{div}(\epsilon_r E_\varepsilon) = \kappa \rho_\varepsilon, & E_\varepsilon(0, \cdot) = 0, \\ \partial_\tau B_\varepsilon + \varepsilon^{-1} \operatorname{curl} E_\varepsilon = 0, & \operatorname{div} B_\varepsilon = 0, & B_\varepsilon(0, \cdot) = 0, \end{cases}$$

where j_ε and ρ_ε are given by (1.3), with f_ε as in (1.4). As usual when dealing with hyperbolic systems [16], the solution $(E_\varepsilon, B_\varepsilon)$ to (3.12) can be represented as a finite sum, implying spectral projectors $\Pi_i(\xi)$ with corresponding eigenvalues $\lambda_i(\xi)$. The mappings $\Pi_i(\cdot)$ and $\lambda_i(\cdot)$ are \mathcal{C}^∞ on $\mathbb{R}^3 \setminus \{0\}$ and positively homogeneous of degree zero and one respectively:

$$(3.13) \quad \begin{pmatrix} E_\varepsilon \\ B_\varepsilon \end{pmatrix}(t, x) = \sum_i \int_0^\tau \int_{\mathbb{R}^3} e^{i(\tau-s)\omega_i(\xi)/\varepsilon} e^{i(x-y)\cdot\xi} \Pi_i(\xi) j_\varepsilon(s, y) ds dy d\xi.$$

The aim of what follows is to describe as precisely as possible, for data microlocalized at physically pertinent frequencies, expressions $(E_\varepsilon, B_\varepsilon)$ given by integrals like in (3.13). Particular care will be paid to the dependence on f_ε via j_ε .

3.1.4. *Dispersion relation for whistler waves.* It is time to clarify the choice and the role of the positive definite matrix ϵ_r . Recall that the approximation $\mu \simeq \mu_0$ (with μ_0 being the permeability of the free space) is correct for space plasmas. In vacuum, the permittivity is constant $\epsilon \equiv \epsilon_0$, so that $\epsilon_r \equiv Id$. Then, the system (3.12) reduces to a wave equation associated with the (rescaled) dispersion relation $\omega(\xi)$ indicated below:

$$\partial_{\tau r}^2 E - \varepsilon^{-2} \Delta E = \mathcal{F}(\varepsilon, f_\varepsilon), \quad \omega(\xi) = \pm|\xi|.$$

The situation under consideration is in fact more complicated because the magnetosphere is a non-homogeneous magnetized plasma. The substitution of μ for μ_0 can be viewed as correct, but the approximation $\epsilon \simeq \epsilon_0$ is certainly not. The presence of some non trivial dielectric tensor $\epsilon_r \not\equiv Id$ must be taken into account. The determination of the resulting dispersion relation $\omega(\cdot)$ is usually achieved in the context of the Appleton-Hartree equation, see [15]. For right-handed waves propagating parallel to the magnetic field, the relevant dispersion curves are drawn in Figure 10 of [3].

This can be specified for wavenumbers k in a range corresponding to whistler modes, that is $k \sim 10^{-5} m^{-1}$ when dealing with the typical frequency 1 – 10 kHz. Then, the following **model law** (see 5.3.3 or [15]) for $\omega(k)$ is usually proposed by physicists (below $\tilde{c} \in]0, 1/2]$):

$$(3.14) \quad \frac{c^2 |k|^2}{\omega(k)^2} = 1 + \frac{\varepsilon \omega_{pe}^2}{\omega(k) (1 - \varepsilon \omega(k))} \sim \frac{\varepsilon \omega_{pe}^2}{\omega(k) (1 - \varepsilon \omega(k))}, \quad \tilde{c} \varepsilon^{-1} \leq \omega(k) \leq \varepsilon^{-1}.$$

The speed of propagation is 1 at the level of (3.5). Thus, in dimensionless variables, the frequency $c|k| \sim 10^4$ must be replaced by $|\xi| \sim \varepsilon^{-1}$. This corresponds to a semiclassical regime. In (3.14), the symbol ω_{pe} represents the **electron plasma frequency**, see (3.6). The ratio $\mu := (\varepsilon \omega_{pe})^{-1}$ between the electron frequency and the plasma frequency is a small parameter, typically of the order 10^{-1} . Please refer to [21]- p. 51 for a representative sample of experimental values. During geomagnetic disturbances, the number density of electrons n_e and therefore ω_{pe} (which depends linearly on $\sqrt{n_e}$) increases inside whistler ducts. As a result, the **overdense condition**, which means that μ is small, can be reinforced up to the value $\mu \sim 10^{-2}$ or a little less, but still with $\varepsilon \ll \mu$. In view of this mechanism [15], the equality in (3.14) can be replaced by the equivalence \sim on the right side of (3.14), leading to a simplified relation:

$$(3.15) \quad \omega(\xi) = \pm \frac{1}{\varepsilon} \lambda(\varepsilon \mu |\xi|), \quad \lambda(z) = \lambda_M \frac{z^2}{1 + z^2} \leq 1.$$

The factor λ_M will be called the *threshold limit* (of the dispersion law). It appears when the frequency $\omega(k)$ of (3.14), expressed in seconds, is changed to fit with the scale of t , meaning that $\lambda_M \simeq T \in \mathbb{R}_+^*$. Recall (2.8) and that $c(w) \simeq 3^{-1}$. The following quantities:

$$(3.16) \quad \lambda_M \simeq 10^{-2}, \quad c(w) b_m \simeq 3^{-7}, \quad c(w) b_M \simeq 2 \times 3^{-4},$$

are clearly comparable, with $c(w) b_m < \lambda_M$. For purposes of the subsequent analysis, what counts most is the scalings of (3.15), together with the majoration of $\lambda(\cdot)$. The precise structure of λ is of course important to capture the geometrical aspects of the propagation, but it does not change the basic effects highlighted in Theorem 2.

At this stage, we can consider that whistler waves correspond to data involving frequencies of size $\mu |\xi| \sim \varepsilon^{-1}$. Moreover, their propagation is achieved through a polarization according to a spectral projector $\Pi_i(\xi)$ with corresponding dispersion relation $\omega_i(\xi)$ satisfying (3.15). This can be summed up as follows.

Assumption 3.1. *[dispersion relation] The propagation of whistler waves is governed by a law of the form (3.15), implying a real analytic strictly increasing function $\lambda(\cdot)$ satisfying:*

$$(3.17a) \quad \lambda : \mathbb{R}_+^* \longrightarrow]0, \lambda_M], \quad \lambda' : \mathbb{R}_+^* \longrightarrow \mathbb{R}_+^*,$$

$$(3.17b) \quad \lambda(0) = 0 < c(w) b_m < \lambda_M := \lim_{z \rightarrow +\infty} \lambda(z).$$

In what follows, because this is more practical, we will work with (y, w) , instead of the original variables (x, v) . The change of variables performed in Paragraph 2.1.2 is orthogonal, whereas $\omega(\cdot)$ depends only on the norm $|\xi|$ of ξ . Thus, the dispersion relations (3.15) is not modified by such a transformation.

3.1.5. *A class of oscillatory integrals.* At the level of (3.12), the constraint on the divergence is propagated. It can be forgotten. What matters is the source term j , which is now given by an integration with respect to w . Note $(y_\varepsilon, \varsigma_\varepsilon, \mathbf{u}_\varepsilon, \mathbf{a}_\varepsilon)(\tau, \cdot)$ the flow induced by (1.5) in the case $(E, B) = (0, 0)$, interpreted after the transformation Ξ_1 of Paragraph 2.2.3. Since $\mathbf{u}_\varepsilon(\tau, \cdot) \equiv w$ for all $\tau \in [0, \mathcal{T}]$, the localization induced by the Dirac mass $\delta_w(\mathbf{u})$ is maintained at the initial position $\mathbf{u} = w$. Now, the amplitude w of the velocity being fixed, the vector w becomes a function $w(\varsigma, \mathbf{a})$ of (ς, \mathbf{a}) , as indicated in (2.31), and sometimes simply denoted by w in what follows. The formula (1.3) gives rise to:

$$(3.18) \quad j_\varepsilon(\tau, y) = \int_{\mathbb{T}^2} \tilde{f}_\varepsilon(\tau, y, \varsigma, \mathbf{a}) \, d\varsigma \, d\mathbf{a},$$

with by construction:

$$\tilde{f}_\varepsilon(\tau, y, \varsigma, \mathbf{a}) = f((y_\varepsilon, \varsigma_\varepsilon, w, \mathbf{a}_\varepsilon)(-\tau, y, w)) (1 - w^2)^{-5/2} w^2 \sin \varsigma \times {}^t O(y) w(\varsigma, \mathbf{a}).$$

The oscillating structure of $(y_\varepsilon, \varsigma_\varepsilon, \mathbf{a}_\varepsilon)$ which has been exhibited in Part 2 is preserved under nonlinear composition. Thereby:

$$(3.19) \quad \tilde{f}_\varepsilon(\tau, y, \varsigma, \mathbf{a}) = \tilde{F}\left(\varepsilon, \tau, y, \varsigma, \mathbf{a}, \frac{\psi_s(\tau, y, w)}{\varepsilon}, \frac{\psi_{p\varepsilon}(\tau, y, w)}{\varepsilon^2}\right) + O(\varepsilon^N).$$

The profile $\tilde{F} \in \mathcal{C}^\infty([0, 1] \times [0, \mathcal{T}] \times \mathbb{R}^3 \times \mathbb{T}^4; \mathbb{R})$ can be expanded in a Fourier series to get:

$$(3.20) \quad \tilde{F} = \sum_{(p,q) \in \mathbb{Z}^2} a^{pq}(\varepsilon, \tau, y, \varsigma, \mathbf{a}) e^{i(p\vartheta + q\theta)}, \quad a^{pq} \in \mathcal{C}^\infty([0, 1] \times [0, \mathcal{T}] \times \mathbb{R}^3 \times \mathbb{T}^2; \mathbb{R}),$$

with:

$$a^{pq}(\varepsilon, \tau, y, \varsigma, \mathbf{a}) = \sum_{j=0}^N \varepsilon^j a_j^{pq}(\tau, y, \varsigma, \mathbf{a}) + O(\varepsilon^{N+1}), \quad \forall (p, q, N) \in \mathbb{Z}^2 \times \mathbb{N}.$$

After the separation of scales, the long time speed of propagation is finite. It follows that the function $\tilde{F}(\varepsilon, \tau, \cdot, \varsigma, \mathbf{a})$ remains compactly supported in y as long as τ is bounded. This property is transferred to the coefficients a_j^{pq} . Fix $\mathcal{T} \in \mathbb{R}_+^*$ as in Paragraph 2.3.2. There is a compact set $K \subset \mathbb{R} \times \mathbb{R}_-^* \times \mathbb{R}$ such that:

$$(3.21) \quad \text{supp } a_j^{pq}(\tau, \cdot, \varsigma, \mathbf{a}) \subset K, \quad \forall (\tau, \varsigma, \mathbf{a}, p, q, j) \in [0, \mathcal{T}] \times \mathbb{T}^2 \times \mathbb{Z}^2 \times \mathbb{N}.$$

The frequencies implied in VLF signals are of the order of 1 kHz (in a range from 100 Hz to 10 kHz). In terms of the present scalings, this corresponds to a localization of ξ in an annulus of size $\mu |\xi| \sim \varepsilon^{-1}$. This can be achieved by implementing a cutoff function:

$$(3.22) \quad \chi \in \mathcal{C}_0^\infty(\mathbb{R}_+^*; \mathbb{R}_+), \quad \chi \not\equiv 0, \quad \text{supp } \chi \subset \mathcal{A} := \{\xi; 0 < r \leq |\xi| \leq R < +\infty\}.$$

To be sure that $\mu |\xi| \sim \varepsilon^{-1}$, we multiply by $\chi(\varepsilon \mu |\xi|)$ on the Fourier side. The equation (3.12) on $(E_\varepsilon, B_\varepsilon)$ is linear. It follows that the contributions brought by the different projections Π_i and the various terms in the expansion (3.20) can be studied separately. On this basis, the solution $(E_\varepsilon, B_\varepsilon)$ to (3.12) can be viewed as a linear combination:

$$(3.23) \quad \begin{pmatrix} E_\varepsilon \\ B_\varepsilon \end{pmatrix}(\varepsilon t, x) = \sum_i \sum_{\pm} \sum_{p \in \mathbb{Z}} \sum_{q \in \mathbb{Z}} \sum_j \mathcal{I}(i, \pm, p, q, j; \varepsilon, t, x),$$

of oscillatory integrals having the following form:

$$(3.24) \quad \mathcal{I}(\varepsilon, t, x) := \frac{-\kappa}{\varepsilon^{1-j}} \int_0^{\varepsilon t} \int_{K \times \mathbb{T}^2 \times \mathbb{R}^3} e^{i \tilde{\Psi}(\varepsilon, \varepsilon t, x; s, y, \varsigma, \mathbf{a}, \xi) / \varepsilon^2} \tilde{a}(s, y, \varsigma, \mathbf{a}, \xi) \, dV,$$

where the notation dV is for the volume element $dV := ds dy d\zeta d\mathbf{a} d\xi$. For the indices i giving rise to (3.15), as explained in Paragraph 3.1.4, we find that (after specifying again the various parameters):

$$\begin{aligned}\tilde{\Psi}(\pm, p, q; \varepsilon, \tau, x; s, y, \varsigma, \mathbf{a}, \xi) &:= \pm(\tau - s)\lambda(\varepsilon\mu|\xi|) + (x - y) \cdot (\varepsilon^2 \xi), \\ &\quad + \varepsilon p \psi_s(s, y, w(\varsigma, \mathbf{a})) + q \psi_{p\varepsilon}(s, y, w(\varsigma, \mathbf{a})), \\ \tilde{a}(p, q, j; s, y, \varsigma, \mathbf{a}, \xi) &:= \chi(\varepsilon\mu|\xi|) \Pi_i(\xi) a_j^{pq}(s, y, \varsigma, \mathbf{a}).\end{aligned}$$

At the level of (3.24), perform the change of variables $s = \varepsilon \tilde{s}$ and $\xi = \varepsilon^{-1} \mu^{-1} \tilde{\xi}$. Rename \tilde{s} and $\tilde{\xi}$ as respectively s and ξ to find:

$$(3.25) \quad \mathcal{I}(\varepsilon, t, x) := \frac{-\kappa}{\varepsilon^{3-j}} \int_0^t \int_{K \times \mathbb{T}^2 \times \mathbb{R}^3} e^{i\tilde{\Psi}(\varepsilon, \varepsilon t, x; \varepsilon s, y, \varsigma, \mathbf{a}, (\varepsilon\mu)^{-1}\xi)/\varepsilon^2} \tilde{a}(\varepsilon s, y, \varsigma, \mathbf{a}, (\varepsilon\mu)^{-1}\xi) dV.$$

Look at the part $\psi_{p\varepsilon}(\varepsilon s, y, w)$ inside the function $\tilde{\Psi}(\cdot)$ of (3.25). Recall (2.103) and (2.104). Since w is fixed, and because y^3 and \mathbf{a} do not appear at the level of $b_e \circ \bar{Y}_0$ (Lemma 2.14), we can skip the dependence on (y^3, w, \mathbf{a}) at the level of (2.103) which becomes:

$$(3.26) \quad \psi_{p\varepsilon}(\varepsilon s, y, w) = \varepsilon c(w) \int_0^s b_e \circ \bar{Y}_0(0, y^1, y^2, \varsigma, r) dr + O(\varepsilon^2).$$

It follows that the function $\tilde{\Psi}(\cdot)$ of (3.25) is of size $O(\varepsilon)$. Isolating the main term, that is the one with ε in factor, makes a new phase $\Psi(\pm, q; t, x; s, y, \varsigma, \xi)$ appear, namely:

$$(3.27) \quad \begin{aligned}\Psi \equiv \Psi(\pm, q; \cdot) &:= \pm(t - s)\lambda(|\xi|) + \mu^{-1}(x - y) \cdot \xi \\ &\quad + q c(w) \int_0^s b_e \circ \bar{Y}_0(0, y^1, y^2, \varsigma, r) dr.\end{aligned}$$

Since Ψ does not depend on \mathbf{a} , we can apply Fubini to put apart the integration with respect to \mathbf{a} , in order to exhibit:

$$(3.28) \quad \mathcal{I}(\varepsilon, t, x) = \int_0^t \int_{K \times \mathbb{T} \times \mathbb{R}^3} e^{i\Psi(t, x; s, y, \varsigma, \xi)/\varepsilon} a(\varepsilon, s, y, \varsigma, \xi) ds dy d\zeta d\xi,$$

where $a(p, q, j; \cdot) \in \mathcal{C}^\infty(]0, 1] \times [0, t] \times \mathbb{R} \times \mathbb{R}_+^* \times \mathbb{R} \times \mathbb{T} \times \mathbb{R}^3)$ is given by:

$$a(\varepsilon, s, y, \varsigma, \xi) := \frac{-\kappa}{\varepsilon^{3-j}} \chi(|\xi|) \Pi_i(\xi) \left(\int_{\mathbb{T}} e^{\frac{2i\pi ps}{T_e(y, w)}} e^{i\check{\Phi}(\varepsilon, s, y, w)} a_j^{pq}(\varepsilon s, y, \zeta, \mathbf{a}) d\mathbf{a} \right),$$

with $w \equiv w(\varsigma, \mathbf{a})$ and, for the sake of completeness:

$$\check{\Phi}(\varepsilon, s, y, w) := q c(w) \int_0^s \left\{ \partial_\tau \bar{\Phi}_0(\varepsilon r, \cdot) + \varepsilon^{-1} [b \circ \bar{Y}_0(\varepsilon r, \cdot) - b \circ \bar{Y}_0(0, \cdot)] \right\}_{|\cdot| \equiv (y, w, r)} dr.$$

From (3.21) and (3.22), it follows that $a(\cdot)$ is uniformly compactly supported in the sense that for all $(p, q, j) \in \mathbb{Z}^2 \times \mathbb{N}$, we have:

$$(3.29) \quad \text{supp } a(\varepsilon, s, \cdot) \subset K := K \times \mathbb{T} \times \mathcal{A}, \quad \forall (\varepsilon, s) \in]0, 1] \times \mathbb{R}.$$

We can recognize $\mathcal{I}(\varepsilon, \cdot)$ as an oscillatory integral implying the phase Ψ and the profile a .

3.2. Stationary phase expansions. The purpose of this Section 3.2 is to examine the asymptotic behaviour of $\mathcal{I}(\varepsilon, t, x)$ when $\varepsilon \in]0, 1]$ goes to zero. In this perspective, the influence of (p, q, j) inside $a(\cdot)$ plays no role. Moreover, we can normalize a by putting the multiplication by ε^{j-3} apart. In fact, what counts is only the smoothness of $a(\cdot)$ and to impose (3.29) on $a(\cdot)$. None the less, it is important to take into account the sign \pm and the choice of $q \in \mathbb{Z}$, which are involved in the phase Ψ . For this reason, when needed, we will mark the dependence on (\pm, q) . To describe $\mathcal{I} \equiv \mathcal{I}(\pm, q; \varepsilon, t, x)$, we will adopt the perspective developed in [16], with adaptations.

Fix a position (t, x) . A preliminary step is to locate the critical points $\mathfrak{C} \equiv \mathfrak{C}(\pm, q; t, x)$ of the phase $\Psi \equiv \Psi(\pm, q; t, x)$. When doing this, particular care must be given to the initial time $s = 0$ and to the final time $s = t$, which are boundary points. The set \mathfrak{C} can be separated into $\mathfrak{C} = \mathfrak{C}_i \cup \mathfrak{C}_b$, where \mathfrak{C}_i collects the interior critical points:

$$\mathfrak{C}_i \equiv \mathfrak{C}_i(\pm, q; t, x) := \{ (s, y, \varsigma, \xi) \in]0, t[\times \mathbb{K}; \nabla_{s, y, \varsigma, \xi} \Psi(t, x; s, y, \varsigma, \xi) = 0 \},$$

whereas $\mathfrak{C}_b = \mathfrak{C}_b^0 \cup \mathfrak{C}_b^t$ includes the boundary critical points of type $r \in \{0, t\}$, that is:

$$\mathfrak{C}_b^r \equiv \mathfrak{C}_b^r(\pm, q; t, x) := \{ (y, \varsigma, \xi) \in \mathbb{K}; \nabla_{y, \varsigma, \xi} \Psi(t, x; r, y, \varsigma, \xi) = 0 \}, \quad r \in \{0, t\}.$$

The non-stationary phase principle says that, when computing \mathcal{I} , the contribution brought by integrating outside any fixed neighbourhood of \mathfrak{C} is of size $O(\varepsilon^\infty)$. This allows to restrict the discussion to points located in \mathfrak{C} . Now, the differential conditions involved in the definition of \mathfrak{C} can be decomposed into:

$$(3.30a) \quad \partial_s \Psi = 0 \quad \Longleftrightarrow \quad 0 = \mp \lambda(|\xi|) + q c(w) b_e \circ \bar{Y}_0(0, y^1, y^2, \varsigma, s),$$

$$(3.30b) \quad \nabla_y \Psi = 0 \quad \Longleftrightarrow \quad 0 = -\xi + \mu q c(w) \int_0^s \nabla_y (b_e \circ \bar{Y}_0)(0, y^1, y^2, \varsigma, r) dr,$$

$$(3.30c) \quad \partial_\varsigma \Psi = 0 \quad \Longleftrightarrow \quad 0 = \int_0^s \partial_\varsigma (b_e \circ \bar{Y}_0)(0, y^1, y^2, \varsigma, r) dr,$$

$$(3.30d) \quad \nabla_\xi \Psi = 0 \quad \Longleftrightarrow \quad 0 = \pm \mu (t - s) \lambda'(|\xi|) |\xi|^{-1} \xi + x - y.$$

To move forward, it is necessary to gather more information on the function $b_e \circ \bar{Y}_0$, which represents how the amplitude of the geomagnetic field fluctuates along the field lines. This is done in Paragraph 3.2.1. The study of \mathfrak{C}_b is based on the relations (3.30b) and (3.30c). In Paragraph 3.2.2, we show that $\mathfrak{C}_b = \emptyset$ for t large enough. In Paragraph 3.2.3, we explain why all harmonics are not relevant (due to a dispersive effect), and why only mirror points can bring non trivial contributions. Then, it remains to examine \mathfrak{C}_i . Note that the three conditions (3.30a), (3.30b) and (3.30c) do not depend on the variables (t, x) , except through the condition $s \leq t$. They select, independently of the choice of (t, x) , the emission points (s, y, ς, ξ) , that is the positions from which the signals are leaving. Their description is achieved in Paragraphs 3.2.5 and 3.2.6. Two situations can be differentiated according to the respective positions of $c(w) b_e \circ \bar{Y}_0(\cdot)$ and λ_M . The case where $c(w) b_e \circ \bar{Y}_0(\cdot)$ is everywhere below the threshold limit λ_M is first examined, in Paragraph 3.2.5. The case of overlapping values is discussed next, in Paragraph 3.2.6. Finally, the propagation aspects are encoded inside (3.30d). They are briefly alluded in Paragraph 3.2.7.

3.2.1. *Preliminary estimates.* Recall the formulas (2.71b) and (2.104) which indicate that the functions ν_e and Γ_e , and therefore $b_e \circ \bar{Y}_0$ as well as derived quantities like $(b_e \circ \bar{Y}_0)^*$, \dots , depend on ς only through \mathfrak{h} . The substitution of \mathfrak{h} is the one of (2.104). To avoid multiple notations, we will not change notations when such factorizations are implied. With this convention, without leading to confusion, we can for instance write:

$$(3.31) \quad b_e \circ \bar{Y}_0(0, y^1, y^2, \varsigma, r) = b_e \circ \bar{Y}_0(\sqrt{\mathcal{H}_e}(y^1, y^2, \varsigma), y^1, y^2, r).$$

More specifically, the expression $\langle b_e \circ \bar{Y}_0 \rangle(0, \cdot) \equiv \Gamma_e$ can be viewed as a function of (\mathfrak{h}, y^2) only, as revealed by (2.106). Fix $(y^1, y^2, \varsigma) \in \mathbb{R} \times \mathbb{R}_+^* \times \mathbb{T}$. The function $b_e \circ \bar{Y}_0(0, y^1, y^2, \varsigma, \cdot)$ is periodic. It undergoes variations between a minimal value $b_m(y^1, y^2, \varsigma)$ and a maximal value $b_M(y^1, y^2, \varsigma)$. This means that (2.8) can be precised into :

$$(3.32) \quad b_m \leq b_m(y^1, y^2, \varsigma) \leq b_e \circ \bar{Y}_0(0, y^1, y^2, \varsigma, s) \leq b_M(y^1, y^2, \varsigma) \leq b_M, \quad \forall s \in \mathbb{R}_+,$$

with for example, in view of (2.104):

$$0 < b_m(y^1, y^2, \varsigma) \equiv b_m(\mathfrak{h}, y^2) := \inf \Gamma_e(\mathfrak{h}, y^2, w, \cdot), \quad \mathfrak{h} = \sqrt{\mathcal{H}_e}(y^1, y^2, \varsigma).$$

Again, due to the periodicity, there are optimal controls $C_j(y^1, y^2, \varsigma) \in \mathbb{R}_+^*$ with $j \in \{1, 2, 3\}$ such that, for all $s \in \mathbb{R}_+$:

$$(3.33a) \quad \|\nabla_y(b_e \circ \bar{Y}_0)(0, y^1, y^2, \varsigma, s)\| \leq C_1(y^1, y^2, \varsigma),$$

$$(3.33b) \quad \left\| \int_0^s \nabla_y(b_e \circ \bar{Y}_0)^*(0, y^1, y^2, \varsigma, r) dr \right\| \leq C_2(y^1, y^2, \varsigma).$$

$$(3.33c) \quad \left\| \int_0^s \partial_{\mathfrak{h}}(b_e \circ \bar{Y}_0)^* dr \right\| \leq C_3(y^1, y^2, \varsigma).$$

Then, compactness arguments allow to derive uniform bounds:

$$(3.34) \quad C_j(y^1, y^2, \varsigma) \leq C_j < +\infty, \quad \forall (y^1, y^2, \varsigma) \in K \times \mathbb{T}, \quad \forall j \in \{1, 2, 3\}.$$

Additional information may be obtained by applying to the expression $b_e \circ \bar{Y}_0(0, y^1, y^2, \varsigma, \cdot)$ the decomposition of Lemma 2.2. With \mathfrak{h} and $\langle \Gamma_e \rangle$ as in (2.104) and (2.106), this leads to:

$$(3.35) \quad \int_0^s (b_e \circ \bar{Y}_0)(0, y^1, y^2, \varsigma, r) dr = \langle \Gamma_e \rangle(\mathfrak{h}, y^2) s + \int_0^s (b_e \circ \bar{Y}_0)^*(0, y^1, y^2, \varsigma, r) dr.$$

In view of conditions (3.10) and (3.11), the part $\langle \Gamma_e \rangle(\mathfrak{h}, y^2)$ of (3.35) only involves specific values of \mathfrak{h} , those which are related to the identification of the extended Van Allen belt. Since the function $\langle \Gamma_e \rangle(\cdot)$ does not depend on y^3 , it suffices to consider the projection of the compact $K_{\bar{f}}$, given by (3.11), onto the two first variables \mathfrak{h} and y^2 , that is:

$$(3.36) \quad Pr(K_{\bar{f}}) := \{(\mathfrak{h}, y^2); 0 \leq \mathfrak{h} \leq f_{ey^2}(h_e(y^2)), y^2 \in \text{supp } h_e(\cdot)\}.$$

Assumption 3.2. [the mean value $\langle \Gamma_e \rangle \equiv \mathcal{A}(\Gamma_e)(\mathfrak{h}, y^2)$ of Γ_e is without critical points] There are positive constants $c_i \in \mathbb{R}_+^*$ and $C_i \in \mathbb{R}_+^*$ with $j \in \{4, 5\}$ such that:

$$(3.37a) \quad 0 < c_4 \leq |\partial_{\mathfrak{h}} \langle \Gamma_e \rangle(\mathfrak{h}, y^2)| \leq C_4, \quad \forall (\mathfrak{h}, y^2) \in Pr(K_{\bar{f}}),$$

$$(3.37b) \quad 0 < c_5 \leq \partial_{y^2} \langle \Gamma_e \rangle(\mathfrak{h}, y^2) \leq C_5, \quad \forall (\mathfrak{h}, y^2) \in Pr(K_{\bar{f}}).$$

Assumption 3.2 is realistic. Indeed:

Lemma 3.1. *[about the relevance of Assumption 3.2] For energy levels $\mathfrak{h} \in \mathbb{R}_+^*$ which are small enough, Assumption 3.2 is sure to be verified.*

Proof. Since w is fixed, we can omit to mention w when it is not necessary. In the integral defining $\langle \Gamma_e \rangle$ at the level of line (2.106), change r into $\mathfrak{h}^{-1}r$ to get:

$$(3.38) \quad \langle \Gamma_e \rangle(\mathfrak{h}, y^2) = \frac{-4}{\pi k_{(\mathfrak{h}, y^2, w)}} \int_0^1 \frac{\mathfrak{h}^2 r b_e(f_{y^2}^{-1}(\mathfrak{h}\sqrt{1-r^2}), y^2) dr}{\mathcal{C}(f_{y^2}^{-1}(\mathfrak{h}\sqrt{1-r^2}), y^2) \cos g_w^{-1}(\mathfrak{h}r)}.$$

There are two ways of looking at (3.38). The first is to exploit the information (2.43) together with the definitions of $f_{y^2}(\cdot)$ and $g_w(\cdot)$ to derive:

$$\begin{aligned} \mathcal{C}(f_{y^2}^{-1}(\mathfrak{h}\sqrt{1-r^2}), y^2) &= (\partial_{y^1} \mathcal{C})(0, y^2) (f_{y^2}^{-1})'(0) \mathfrak{h} \sqrt{1-r^2} + O(\mathfrak{h}^2(1-r^2)), \\ \cos g_w^{-1}(\mathfrak{h}r) &= -(g_w^{-1})'(0) \mathfrak{h}r + O(\mathfrak{h}^2 r^2). \end{aligned}$$

This allows to justify the convergence of the integral (3.38), and to see that $\langle \Gamma_e \rangle(0, y^2) \neq 0$. The other view is to recall Proposition 2.1, where it was noted that the operation $\mathcal{A}(\cdot) \equiv \langle \cdot \rangle$ amounts to integrate with respect to a probability density, so that:

$$\langle \Gamma_e \rangle(\mathfrak{h}, y^2) = b_e(0, y^2) + \frac{2 (\partial_{y^1 y^1}^2 b_e)(0, y^2) (f_{y^2}^{-1})'(0)}{\pi k_{(0, y^2, w)} (\partial_{y^1} \mathcal{C})(0, y^2) (g_w^{-1})'(0)} \left(\int_0^1 \sqrt{1-r^2} dr \right) \mathfrak{h}^2 + O(\mathfrak{h}^4).$$

The reason explaining the $O(\mathfrak{h}^4)$ instead of a $O(\mathfrak{h}^3)$ is that $\langle \Gamma_e \rangle(\cdot, y^2)$ is an even function. From this and (2.50), we can deduce that:

$$\partial_{y^2} \langle \Gamma_e \rangle(\mathfrak{h}, y^2) = \partial_{y^2} b_e(0, y^2) + O(\mathfrak{h}^2), \quad \partial_{y^2} b_e(0, y^2) < 0.$$

In view of the construction leading to f and g , see for instance (2.50), we have:

$$(f_{y^2}^{-1})'(0) = f'_{y^2}(0)^{-1} > 0, \quad (g_w^{-1})'(0) = g'_w(0)^{-1} = \sqrt{2c(w)} > 0.$$

This line, completed with (2.11c) and (2.45), gives rise to:

$$\mathfrak{h}^{-1} \partial_{\mathfrak{h}} \langle \Gamma_e \rangle(\mathfrak{h}, y^2) = c + O(\mathfrak{h}^2), \quad c > 0.$$

It suffices to select $\mathfrak{h} \in \mathbb{R}_+^*$ small enough to recover (3.37). It is even possible, with the above informations, to quantify a maximal size of \mathfrak{h} allowing to get the minorations of (3.37). Since the remainders are $O(\mathfrak{h}^2)$ in place of $O(\mathfrak{h})$, the effect is enhanced. \square

Lemma 3.2. *[extra properties related to $\langle \Gamma_e \rangle$ after substitution of \mathfrak{h} for $\sqrt{\mathcal{H}_e}(y^1, y^2, \varsigma, w)$] There are positive constants c_6 and C_6 such that, for all $(y^1, y^2, \varsigma) \in K \times \mathbb{T}$, we have:*

$$(3.39) \quad 0 < c_6 \leq \| \nabla_y [\langle \Gamma_e \rangle(\sqrt{\mathcal{H}_e}(y^1, y^2, \varsigma, w), y^2)] \| \leq C_6 < +\infty.$$

Proof. The domain $K \times \mathbb{T}$ is compact. Thus, it suffices to show that the function under consideration is nonzero. First, remark that:

$$\partial_{y^1} [\langle \Gamma_e \rangle(\mathfrak{h}, y^2)] = \frac{1}{2} \partial_{\mathfrak{h}} \langle \Gamma_e \rangle(\mathfrak{h}, y^2) \mathcal{H}_e(y^1, y^2, \varsigma)^{-1/2} d^1(y^1, y^2)^{-1} \mathcal{C}(y^1, y^2).$$

Taking into account Assumptions 2.2 and 3.2, together with condition (3.37), this expression can be zero if and only if $y^1 = 0$. But then the derivative with respect to y^2 cannot also be zero. Indeed, the identity $\mathcal{H}_e(0, y^2, \varsigma) = \mathcal{P}_2(\varsigma, w)$ and (3.37b) give rise to:

$$\partial_{y^2} \mathcal{H}_e(0, y^2, \varsigma) = 0, \quad \partial_{y^2} [\langle \Gamma_e \rangle(\mathcal{H}_e(0, y^2, \varsigma), y^2)] = \partial_{y^2} \langle \Gamma_e \rangle(\mathfrak{h}, y^2) \neq 0. \quad \square$$

3.2.2. *Absence of boundary critical points for sufficiently large times.* From now on, we suppose that Assumption 3.2 is verified. Knowing that $\xi \in \mathcal{A}$, the relation (3.30b) written at time $s = 0$ or for $q = 0$ yields:

$$\| \nabla_y \Psi(\pm, q; t, x; 0, y, \varsigma, \xi) \| = \| \nabla_y \Psi(\pm, 0; t, x; s, y, \varsigma, \xi) \| = \| \xi \| \geq r > 0.$$

Therefore $\mathfrak{C}_b^0(\pm, q; t, x) = \emptyset$ and $\mathfrak{C}(\pm, 0; t, x) = \emptyset$ for all (t, x) . This preliminary result can be further elaborated.

Lemma 3.3. *[exclusion of boundary critical points] Given $q \in \mathbb{Z}^*$, define:*

$$(3.40) \quad t_m(q, \mu) := \mu^{-1} |q|^{-1} C_1^{-1} r, \quad t_M(q, \mu) := c_6^{-1} ((\mu |q| c(w))^{-1} R + C_2).$$

For all $t > t_M(q, \mu)$, we find that $\mathfrak{C}_b(\pm, q; t, x) = \emptyset$ and that:

$$(3.41a) \quad \mathfrak{C}_i(\pm, q; t, x) \cap \{(s, y, \varsigma, \xi); s \in [0, t_m(q, \mu)[\} = \emptyset,$$

$$(3.41b) \quad \mathfrak{C}_i(\pm, q; t, x) \cap \{(s, y, \varsigma, \xi); s \in]t_M(q, \mu), t[\} = \emptyset.$$

Proof. The part (3.41a) is a simple consequence of:

$$\| \nabla_y \Psi(\pm, q; t, x; s, y, \varsigma, \xi) \| \geq r - \mu |q| C_1 s > 0, \quad \forall s \in [0, t_m(q, \mu)[.$$

To extract (3.41b), exploit first (3.35) to interpret (3.30b) according to:

$$\nabla_y [\langle \Gamma_e \rangle(\mathfrak{h}, y^2)] s = (\mu q c(w))^{-1} \xi - \int_0^s \nabla_y (b_e \circ \bar{Y}_0)^*(0, y^1, y^2, \varsigma, r) dr, \quad q \neq 0.$$

Then, use a combination of (3.22), (3.33b) and (3.39). \square

For $t > t_M(q, \mu)$, the study of $\mathfrak{C}(\pm, q; t, x)$ is reduced to the one of $\mathfrak{C}_i(\pm, q; t, x)$. This information will be crucial in the next paragraph 3.2.3 since it allows to integrate by parts, separately with respect to s .

3.2.3. *Sorting of the harmonics.* As usual when dealing with nonlinear dispersive equations, not all harmonics q are relevant.

Lemma 3.4. *[selection of the propagated harmonics] Fix any time t with $t > t_M(1, \mu)$. The condition $\mathfrak{C}(\pm, q; t, x) \neq \emptyset$ can be realized only if:*

$$(3.42) \quad q \in \mathcal{H} := \{q \in \mathbb{Z}^*; |q| \leq \lambda_M b_m^{-1} c(w)^{-1}\}, \quad \text{Card } \mathcal{H} < +\infty.$$

Proof. Fix any $q \in \mathbb{Z}$. Remark that $t_M(1, \mu) \geq t_M(q, \mu)$ for all $q \in \mathbb{Z}^*$. From Lemma 3.3, for $t > t_M(1, \mu)$, we have that:

$$\mathfrak{C}_b^t(\pm, q; t, x) = \emptyset, \quad \mathfrak{C}(\pm, q; t, x) \equiv \mathfrak{C}_i(\pm, q; t, x), \quad \forall q \in \mathbb{Z}.$$

In this context, the function $a(\cdot)$ may as well be assumed to be compactly supported in the interval $]0, t[$. Then, to make a preliminary selection of the critical points, we can look independently at (3.30a). Exploit (3.32) to get:

$$|\partial_s \Psi(\pm, q; t, x; s, y^1, y^2, \varsigma, \xi)| \geq \begin{cases} |q| c(w) b_m(y^1, y^2, \varsigma) + \inf_{r \leq |\xi| \leq R} \lambda(|\xi|) & \text{if } \pm q \in \mathbb{Z}_-, \\ |q| c(w) b_m(y^1, y^2, \varsigma) - \lambda_M & \text{if } \pm q \in \mathbb{N}^*. \end{cases}$$

The second minoration invites to compare $c(w) b_m$ with λ_M . Interesting situations are sure to appear because $1 < \lambda_M b_m^{-1} c(w)^{-1} \simeq 22$, see (3.16).

It should be emphasized here that the comparison of the two quantities $c(w) b_m$ and λ_M makes sense from a physical point of view. On the one hand, the nondimensionalization of Paragraph 3.1.1 yields a function $b_e \circ \bar{Y}_0(0, y^1, y^2, \varsigma, \cdot)$ which is of size $\simeq 1$, with minimum and maximum values estimated in optimal manner as in (3.32). On the other hand, the threshold limit λ_M , which was introduced in Assumption 3.1, is an experimental data that is also of size $\simeq 1$.

The phase $\Psi(\pm, q)$ for $\pm q \in \mathbb{Z}_-$, or for $\pm q \in \mathbb{N}^*$ with $|q| > \lambda_M b_m^{-1} c(w)^{-1}$, has no stationary points, whereas the domain of integration is compact. As a result of nonstationary phase theorem, we find $\mathcal{I}(\pm, q) = O(\varepsilon^\infty)$ in such cases. The quantity $\mathcal{I}(\pm, q)$ may be non-negligible modulo ε^∞ only if $\pm q \in \mathbb{N}^*$ and $|q| \leq \lambda_M b_m^{-1} c(w)^{-1}$. \square

Obviously, there is only a finite number of harmonics q which can satisfy (3.42). Moreover, where such q exist, the values $q = \pm 1$ are necessarily included.

3.2.4. Reduction to a spatial phase. From the perspective of the Maxwell equations (3.12), the variables x and v do not intervene in the same way. The velocity component v appears when integrating to identify j . On the basis of Paragraph 3.1.5, it is involved only through the component ς of w .

Lemma 3.5. *[the signals are necessarily emanating from the two mirror points] There is $\mu_0 \in \mathbb{R}_+^*$ such that, for all $\mu \in]0, \mu_0[$ and $t > t_M(1, \mu)$, the condition $(s, y, \varsigma, \xi) \in \mathfrak{C}(\pm, q; t, x)$ is equivalent to $\varsigma = \pi/2$.*

Proof. From Lemma 3.3, the discussion is reduced to times s such that $s \geq t_m(q, \mu)$. Now, look at (3.30c). The formulas (2.71b) and (2.104) indicate that $b_e \circ \bar{Y}_0$ depends on ς only through \mathfrak{h} . Recall the convention (3.31). We have:

$$(3.43) \quad \partial_\varsigma \Psi = - \left[\partial_{\mathfrak{h}} \langle \Gamma_e \rangle s + \int_0^s \partial_{\mathfrak{h}} (b_e \circ \bar{Y}_0)^* dr \right] \mathcal{H}_e^{-1/2} c(w)^{-1} 2^{-1} \cotan \varsigma .$$

The condition $\partial_\varsigma \Psi = 0$ is sure to be satisfied when $\varsigma = \pi/2$. The value $\varsigma = \pi/2$ corresponds to positions where $w^\parallel = 0$, that is to mirror points. In this sense, when $\varsigma = \pi/2$, we may say that the signals are generated from the two mirror points.

It remains to show that the condition $\varsigma = \pi/2$ is also necessary to obtain $\partial_\varsigma \Psi = 0$. To this end, at the level of (3.43), consider the expression which is inside the brackets. It is the sum of two functions. Due to (3.37a), the first function is linear in s with a non zero slope. In view of (3.33c), the second function is bounded for all $s \in \mathbb{R}_+$. Concretely:

$$(3.44) \quad \left| (\partial_{\mathfrak{h}} \langle \Gamma_e \rangle s + \int_0^s \partial_{\mathfrak{h}} (b_e \circ \bar{Y}_0)^* dr \right| \geq c_4 t_m(q, \mu) - C_3, \quad \forall s > t_m(q, \mu) .$$

For $\mu \in]0, \mu_0[$ with:

$$\mu_0 = r c(w) b_m C_4 C_1^{-1} C_3^{-1} \lambda_M^{-1},$$

the right-hand term of (3.44), the minimizer, is sure to be not zero. \square

The level energy $\mathfrak{h} \in \mathbb{R}_+^*$ and the altitude $y^2 \in \mathbb{R}_+^*$ being fixed, the periodic function $b_e \circ \bar{Y}_0(0, y^1, y^2, \varsigma, \cdot)$ oscillates between a minimal value $b_m(\mathfrak{h}, y^2)$ and a maximal value $b_M(\mathfrak{h}, y^2)$. The following analysis depends on how $c(w) b_m(\mathfrak{h}, y^2)$ is positioned with respect to λ_M . Given q with $|q| c(w) b_m(\mathfrak{h}, y^2) < \lambda_M$, two situations may arise.

As we will see in the next Paragraphs 3.2.5 and 3.2.6, two cases can be distinguished on the basis of the following chief characteristics:

- 3a. The case $|q| c(w) b_M(\mathfrak{h}, y^2) < \lambda_M$. The whistler episode is spread over a finite period of time t , still large for small values of μ . Namely, it is of size $c \mu^{-1}$ with a fixed constant c not depending on R ;
- 3b. The case $|q| c(w) b_M(\mathfrak{h}, y^2) > \lambda_M$. The whistler episode is of size $c(R) \mu^{-1}$ with a constant $c(R)$ growing to infinity when the control R , on the norms $|\xi|$ of frequencies, tends to $+\infty$. From this perspective, it can be extended beyond any time $t \in \mathbb{R}_+^*$. But, in practice, the size of $|\xi|$ is limited.

3.2.5. *Harmonics q for which the variations of $|q| c(w) b_e \circ \bar{Y}_0$ are below the threshold limit.* This is the case 3a, that is when the domain of definition of the function $\lambda^{-1}(\cdot)$ contains the interval $[|q| c(w) b_m, |q| c(w) b_M]$. Select $t > t_M(1, \mu)$. Combining Lemmas 3.3 and 3.5, the condition $(s, y, \varsigma, \xi) \in \mathfrak{C}(\pm, q)$ appears to be equivalent to $\varsigma = \pi/2$, together with (3.30a) and (3.30b). On the one hand, from (3.30a), we can deduce that $|\xi| < \lambda^{-1}(|q| c(w) b_M)$. On the other hand, from the equation (3.30b), together with (3.33b), (3.35) and (3.39), it is easy to infer that $|\xi| \geq \mu |q| c(w) (c_6 s - C_2)$. It follows that:

$$(s, y, \varsigma, \xi) \notin \mathfrak{C}(\pm, q), \quad \forall s > \tilde{t}_M(q, \mu) := c_6^{-1} [(\mu |q| c(w))^{-1} \lambda^{-1}(|q| c(w) b_M) + C_2].$$

For $s \leq \tilde{t}_M(q, \mu)$, the relation (3.30b) and the majoration (3.33a) give rise to:

$$|\xi| \leq R_M(q, \mu) := c_6^{-1} C_1 \lambda^{-1}(|q| c(w) b_M) + \mu c_6^{-1} C_1 C_2 |q| c(w).$$

Since $q \in \mathcal{H}$ and $\mu \leq \mu_0$, a uniform bound on $|\xi|$ is available:

$$|\xi| \leq R_M := c_6^{-1} C_1 [\lambda^{-1}(\lambda_M b_m^{-1} b_M) + \mu_0 C_2 \lambda_M b_m^{-1}].$$

From now on, we fix R with $R > R_M$. This condition implies that there is no restriction coming from (3.30b) when working with (3.30a) on the time interval $[0, \tilde{t}_M(q, \mu)]$. Indeed, since the control on $|\xi|$ is not too restrictive, the count of all possible values of ξ shall be exhaustive. The direction ξ can be freely adjusted in order to recover (3.30b). Then, we can substitute ξ in (3.30a) with the value extracted from (3.30b), and concentrate on the identity thus obtained.

The situation 3a, if it happens, includes at least the value $q = 1$ with the sign $+$. Now, the discussion being similar for all other related values of q , it can be illustrated by investigating only the case of $\mathcal{I} \equiv \mathcal{I}_+^1$ with $\Psi \equiv \Psi_+^1$. Fix $(y^1, y^2) \in \mathbb{R} \times \mathbb{R}_+^*$. Combining (3.30a) and (3.30b), it is easy to check that $(s, y, \frac{\pi}{2}, \xi) \in \mathfrak{C}(+, 1)$ if and only if s is adjusted so that:

$$(3.45) \quad Osc(s) = \mu c(w) \left\| \int_0^s \nabla_y (b_e \circ \bar{Y}_0)(0, y^1, y^2, \pi/2, r) dr \right\|$$

where we have introduced the function:

$$Osc : \mathbb{R}_+ \longrightarrow [\lambda^{-1}(c(w) b_m), \lambda^{-1}(c(w) b_M)] \subset \mathbb{R}_+^*,$$

given by:

$$Osc(s) := \lambda^{-1}(c(w) b_e \circ \bar{Y}_0(0, y^1, y^2, \pi/2, s)), \quad \forall s \in \mathbb{R}_+.$$

We call \mathbb{S}_μ^y the set of all solutions to (3.45). On the other hand, the cardinal of \mathbb{S}_μ^y is denoted by $N_\mu^y := \text{Card}(\mathbb{S}_\mu^y)$. Recall that the overdense condition means that $\mu \in]0, \mu_0]$ with μ_0 sufficiently small (to be fixed later).

Proposition 3.1. *[asymptotic time distribution of the emission points under the overdense condition; case where $c(w)b_M$ is below the threshold limit] There is a constant $C_8 \in \mathbb{R}_+^*$ such that the number N_μ^y of solutions to (3.45) is finite. It is equivalent to $N_\mu^y \sim C_8 \mu^{-1}$ when μ goes to zero. The elements s_μ^j of \mathbb{S}_μ^y can always be ranked in order of increase, with $1 \leq j \leq N_\mu^y$. There exists $\mu_0 \in \mathbb{R}_+^*$ such that, for all $C \in]0, C_8[$, there is a constant $M(C) \in \mathbb{R}_+^*$ such that, for all $\mu \in]0, \mu_0]$, there is a subset $\tilde{\mathbb{S}}_\mu^y \subset \mathbb{S}_\mu^y$ satisfying:*

$$(3.46) \quad C \mu^{-1} \leq \text{Card}(\tilde{\mathbb{S}}_\mu^y), \quad M(C) \leq s_\mu^{j+1} - s_\mu^j \leq T(y^1, y^2, \pi/2, w)/4, \quad \forall s_\mu^j \in \tilde{\mathbb{S}}_\mu^y.$$

Proof. Fix any $y \in K$. By construction:

$$\mathcal{H}_e(y, \pi/2, w, \mathbf{a}) = \mathcal{P}_1(y) - c(w) \ln(\sin \frac{\pi}{2}) = \mathcal{P}_1(y).$$

This means that $\bar{Y}_0(0, y, \pi/2, \cdot)$ starts at initial time $s = 0$ from an extremal position $\pm y_m^1$. Lemma 2.10 can therefore be applied. Since $b_e(\cdot, y^2)$ is an even function, it follows that the function $Osc(\cdot)$ is periodic of period $T/2$. More precisely, it is strictly decreasing on the interval $[0, T/4]$ from a maximum value $O_M \leq \lambda^{-1}(c(w)b_m)$ when $s = 0$, to a minimum positive value $O_m \geq \lambda^{-1}(c(w)b_m)$ that is reached when $s = T/4$. It is strictly increasing on the interval $[T/4, T/2]$ from O_m to O_M , and so on. Briefly, we can assert that:

$$(3.47a) \quad Osc(0) = O_M, \quad Osc(T/4) = O_m, \quad Osc(T/2) = O_M,$$

$$(3.47b) \quad Osc'(0) = 0, \quad Osc'(T/4) = 0, \quad Osc'(T/2) = 0,$$

$$(3.47c) \quad Osc''(0) < 0, \quad Osc''(T/4) > 0, \quad Osc''(T/2) < 0,$$

$$(3.47d) \quad Osc'(s) < 0, \quad \forall s \in]0, T/4[, \quad Osc'(s) > 0, \quad \forall s \in]T/4, T/2[.$$

First of all, there is a need to explain why N_μ^y is finite. This is because all expressions involved in (3.45) are real analytic: the functions b_e and λ by definition, and the application \bar{Y}_0 by virtue of the ODE version of Cauchy-Kovalevskaya Theorem. Thus, we can interpret the condition (3.45) in the form:

$$\mu^{-1} = \mathcal{A}(s), \quad \mathcal{A}(s + T/2) = \mathcal{A}(s) + cT/2, \quad \forall s \in \mathbb{R}, \quad c \in \mathbb{R},$$

where $\mathcal{A}(\cdot)$ is real analytic. Therefore:

$$\text{Card}\{s \in [0, T/2]; \mu^{-1} = \mathcal{A}(s)\} \leq P := \text{Card}\{s \in [0, T/2]; \mathcal{A}'(s) = 0\} < +\infty.$$

Since $\mathcal{A}'(\cdot)$ is periodic with period $T/2$, we can also deduce that:

$$(3.48) \quad \text{Card}\{s \in [nT/2, (n+1)T/2] \cap \mathbb{S}_\mu^y\} \leq P, \quad \forall (\mu, n) \in]0, \mu_0] \times \mathbb{N}.$$

To have a more accurate count of the elements contained in \mathbb{S}_μ^y , it is convenient to transcribe the relation (3.45) in another equivalent form. It turns out that $s \in \mathbb{S}_\mu^y$ if and only if:

$$(3.49) \quad \mathcal{F}(s, \mu) = 0, \quad \mathcal{F}(s, \mu) := Osc(s) - \mu \eta s - \mu Rsc(s),$$

where, taking into account (3.39), we have to deal with the positive number:

$$\eta := c(w) \|\nabla_y \langle b_e \circ \bar{Y}_0 \rangle(0, y^1, y^2, \pi/2)\| = c(w) \|\nabla_y [\langle \Gamma_e \rangle(\mathfrak{h}, y^2)]\| \in \mathbb{R}_+^*,$$

and the auxilliary function $Rsc(s) := Rsc1(s) + Rsc2(s)$ with:

$$\begin{aligned} Rsc1(s) &:= +c(w) \|\nabla_y [\langle \Gamma_e \rangle(\mathfrak{h}, y^2)]\| s + \int_0^s \nabla_y (b_e \circ \bar{Y}_0)^*(0, y, \pi/2, r) dr \|\cdot\|, \\ Rsc2(s) &:= -c(w) \|\nabla_y [\langle \Gamma_e \rangle(\mathfrak{h}, y^2)]\| s. \end{aligned}$$

To help the understanding of the proof, please refer to the following picture 1.

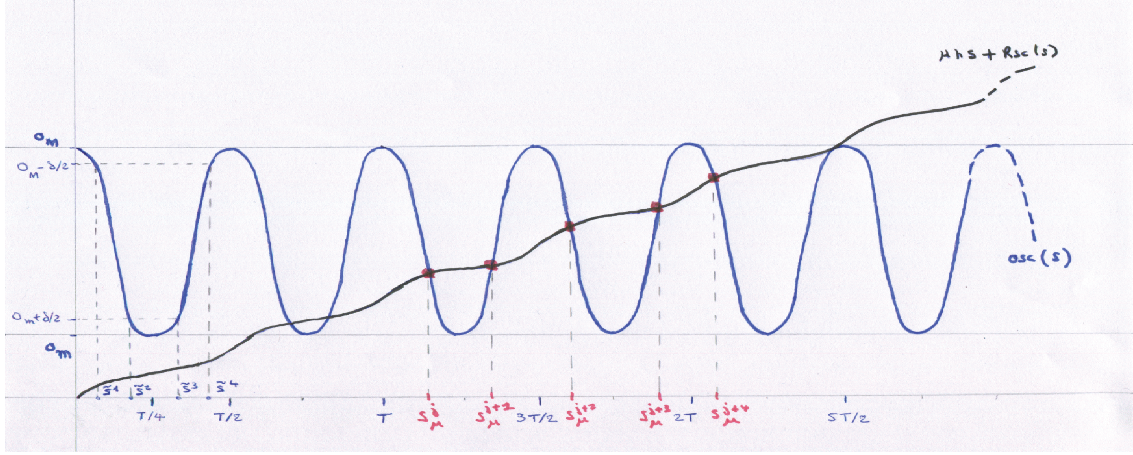


Figure 1. Below the threshold limit.
 Display of the solutions to $\mathcal{F}(s, \mu) = 0$.

The relation (3.45) can be realized only if $s \geq s_m := O_m C_1^{-1} \mu^{-1} c(w)^{-1}$. Define:

$$\mu_0 := \min(r c(w) b_m C_4 C_1^{-1} C_3^{-1} \lambda_M^{-1}; c_6 O_m c(w)^{-1} C_1^{-1} C_2^{-1}).$$

For $\mu \leq \mu_0$ and s as above, the part $Rsc1(\cdot)$ of $Rsc(\cdot)$ is sure to be non zero. It follows that the function $Rsc1(\cdot)$, and therefore $Rsc(\cdot)$, is smooth for $s \geq s_m$. The application $Rsc1(\cdot)$ has a particular form. It is the norm of a linear function plus a periodic function. Due to this specificity, it is bounded, with a bounded derivative. The same is true concerning $Rsc(\cdot)$. For instance, we have:

$$|Rsc(s)| \leq \int_0^{T/4} \|\nabla_y (b_e \circ \bar{Y}_0)^*(0, y, \pi/2, r)\| dr \leq C_2(\mathfrak{h}, y^2), \quad \forall s \in \mathbb{R}_+.$$

We can introduce:

$$R_M^0 := \sup \{|Rsc(s)|; s \in \mathbb{R}_+\} < +\infty, \quad R_M^1 := \sup \{|Rsc'(s)|; s \geq s_m\} < +\infty.$$

A preliminary remark is that the condition $s \in \mathbb{S}_\mu^y \cap [nT/2, (n+1)T/2]$ can be realized only if the integer n is such that:

$$(3.50) \quad \frac{2}{\eta T} (\mu^{-1} O_m - R_M^0) - 1 \leq n \leq \frac{2}{\eta T} (\mu^{-1} O_M + R_M^0).$$

In view of (3.47), $Osc'(s) = 0$ if and only if $Osc(s) = O_m$ or $Osc(s) = O_M$. After shifting the value of $Osc(s)$ a little away from O_m and O_M , we can get a quantified version of this property. In fact, for $C_9 \in \mathbb{R}_+^*$ and $\mu_0 \in \mathbb{R}_+^*$ respectively sufficiently large and small, we have for all $\mu \in]0, \mu_0]$ that:

$$(3.51) \quad 2 \mu (\eta + R_M^1) \leq |Osc'(s)|, \quad \forall s := Osc^{-1}([O_m + C_9 \mu, O_M - C_9 \mu]).$$

Given $n \in \mathbb{N}$, consider the condition:

$$(3.52) \quad \frac{2}{\eta T} (\mu^{-1} O_m + C_9 + R_M^0) + 1 \leq n \leq \frac{2}{\eta T} (\mu^{-1} O_M - C_9 - R_M^0) - 1.$$

The set \mathbb{S}_μ^y can be decomposed into the disjoint union $\mathbb{S}_\mu^y = \tilde{\mathbb{S}}_\mu^y \cup (\tilde{\mathbb{S}}_\mu^y)^c$, where:

$$\tilde{\mathbb{S}}_\mu^y := \{s \in \mathbb{S}_\mu^y; s \in [nT/2, (n+1)T/2], n \text{ is as in (3.52)}\}.$$

Combining (3.48), (3.50) and (3.52), it is easy to infer that:

$$(3.53) \quad \text{Card}(\tilde{\mathbb{S}}_\mu^y)^c \leq 4 (C_9 + 2 R_M^0 + 2) P \eta^{-1} T^{-1}.$$

Define the times t_μ^j with $j \in \{1, 2, 3, 4\}$ in such a way that:

$$(3.54a) \quad (t_\mu^1, t_\mu^2) \in]0, T/4[^2, \quad \text{Osc}(t_\mu^1) = O_M - C_9 \mu, \quad \text{Osc}(t_\mu^2) = O_m + C_9 \mu,$$

$$(3.54b) \quad (t_\mu^3, t_\mu^4) \in]T/4, T/2[^2, \quad \text{Osc}(t_\mu^3) = O_m + C_9 \mu, \quad \text{Osc}(t_\mu^4) = O_M - C_9 \mu.$$

Consider also the translated versions of these times:

$$(3.55) \quad t_\mu^j(n) := t_\mu^j + nT/2, \quad (j, n) \in \{1, 2, 3, 4\} \times \mathbb{N}.$$

Select n as in (3.52). By construction, we have:

$$(3.56) \quad O_m + \mu C_9 \leq \text{Osc}(s) \leq O_M - \mu C_9, \quad \forall s \in \mathbb{S}_\mu^y \cap [nT/2, (n+1)T/2],$$

implying that:

$$\mathbb{S}_\mu^y \cap [nT/2, (n+1)T/2] \subset I_\mu(n) := [t_\mu^1(n), t_\mu^2(n)] \cup [t_\mu^3(n), t_\mu^4(n)].$$

For such n , observe that:

$$(3.57) \quad \mathcal{F}(t_\mu^j(n), \mu) > 0, \quad \forall j \in \{1, 4\}, \quad \mathcal{F}(t_\mu^j(n), \mu) < 0, \quad \forall j \in \{2, 3\}.$$

Secondly, deduce from (3.51) and (3.56) that:

$$(3.58) \quad |\partial_s \mathcal{F}(s, \mu)| \geq \mu (\eta + R_M^1) > 0, \quad \forall s \in I_\mu(n).$$

On one side, we have (3.53) with a bound not depending on $\mu \in]0, \mu_0]$. On the other side, from (3.57) and (3.58), it follows that, for all integer n satisfying (3.52), we have:

$$\text{Card} \mathbb{S}_\mu^y \cap [t_\mu^1(n), t_\mu^2(n)] = 1, \quad \text{Card} \mathbb{S}_\mu^y \cap [t_\mu^3(n), t_\mu^4(n)] = 1, \quad \text{Card} \mathbb{S}_\mu^y \cap I_\mu(n) = 2.$$

This explains why $N_\mu^y(\mu) \sim C_8 \mu^{-1}$ with $C_8 := 4(O_M - O_m)/(\eta T)$.

Let us now take a look at the second part of (3.46). Fix any constant $C \in]0, C_8[$. Define $\delta := (C_8 - C) \eta T/8 \in \mathbb{R}_+^*$. If necessary, diminish μ_0 to be sure that:

$$0 < \mu_0 \leq \delta (2 C_9 + 2 R_M^0 + \eta T)^{-1}.$$

Replace (3.52) with the more restrictive condition:

$$(3.59) \quad 2 (O_m + \delta) (\mu \eta T)^{-1} \leq n \leq 2 (O_M - \delta) (\mu \eta T)^{-1}.$$

Define the times \tilde{s}^j with $j \in \{1, 2, 3, 4\}$ in such a way that:

$$(3.60a) \quad (\tilde{s}^1, \tilde{s}^2) \in]0, T/4[^2, \quad \text{Osc}(\tilde{s}^1) = O_M - (\delta/2), \quad \text{Osc}(\tilde{s}^2) = O_m + (\delta/2),$$

$$(3.60b) \quad (\tilde{s}^3, \tilde{s}^4) \in]T/4, T/2[^2, \quad \text{Osc}(\tilde{s}^3) = O_m + (\delta/2), \quad \text{Osc}(\tilde{s}^4) = O_M - (\delta/2).$$

The set of integers n satisfying (3.59) yields at least $C\mu^{-1}$ solutions in \mathbb{S}_μ^y . Moreover, any of these solutions s_μ^j is such that:

$$O_m + (\delta/2) \leq \text{Osc}(s_\mu^j) \leq O_M - (\delta/2), \quad s_\mu^j \in [nT/2, (n+1)T/2],$$

and therefore:

$$0 < M(C) \leq s_{j+1}(\mu) - s_j(\mu), \quad M(C) := \min(\tilde{s}^3 - \tilde{s}^2; \tilde{s}^1 + (T/2) - \tilde{s}^4). \quad \square$$

3.2.6. Emission points in the case of overlapping values. This is the case **3b**, as is mentioned at the end of Subsection 3.2.4. Again, the related discussion is similar for all corresponding values of q . As before, it can be illustrated by investigating the case of $\mathcal{I} \equiv \mathcal{I}_+^1$, with $\Psi \equiv \Psi_+^1$ and $c(\mathbf{w})b_m(\mathbf{h}, y^2) < \lambda_M < c(\mathbf{w})b_M(\mathbf{h}, y^2)$. Fix \mathbf{h} and y . Given $\delta \in [0, \lambda_M - c(\mathbf{w})b_m[$, define the times \check{s}^j with $j \in \{1, 2, 3, 4\}$ so that:

$$(3.61a) \quad \check{s}^1 \in]0, T/4[, \quad c(\mathbf{w})b_e \circ \bar{Y}_0(\mathbf{h}, y^1, y^2, \check{s}_\delta^1) = \lambda_M,$$

$$(3.61b) \quad \check{s}^2 \in]0, T/4[, \quad c(\mathbf{w})b_e \circ \bar{Y}_0(\mathbf{h}, y^1, y^2, \check{s}_\delta^2) = \lambda_M - \delta,$$

$$(3.61c) \quad \check{s}^3 \in]T/4, T/2[, \quad c(\mathbf{w})b_e \circ \bar{Y}_0(\mathbf{h}, y^1, y^2, \check{s}_\delta^3) = \lambda_M - \delta,$$

$$(3.61d) \quad \check{s}^4 \in]T/4, T/2[, \quad c(\mathbf{w})b_e \circ \bar{Y}_0(\mathbf{h}, y^1, y^2, \check{s}_\delta^4) = \lambda_M.$$

By restricting μ_0 if necessary, we can always impose:

$$(3.62) \quad 0 < 2\mu_0(\eta + R_M^1) \leq |\text{Osc}'(s)|, \quad \forall s \in [\check{s}^1, \check{s}^2] \cup [\check{s}^3, \check{s}^4].$$

Proposition 3.2. *[asymptotic time distribution of the emission points under the overdense condition; case of overlapping values] There exist two constants $\mu_0 \in]0, 1]$ and $c_{10} \in \mathbb{R}_+^*$ such that, for all radius $R > C_1 c_{10}$, $\mu \in]0, \mu_0]$ and for all $n \in [c_{10}\mu^{-1}, C_1^{-1}R\mu^{-1}] \cap \mathbb{N}$, there are exactly two emission points $(s_n^1, y, \varsigma, \xi)$ with $s_n^1 \in]nT/2, (2n+1)T/4[$ and $(s_n^4, y, \varsigma, \xi)$ with $s_n^4 \in](2n+1)T/4, nT/2[$, which are such that (when $n \rightarrow +\infty$ or $\mu \rightarrow 0$):*

$$(3.63) \quad s_n^1 = nT/2 + \check{s}_0^1 + o(1), \quad s_n^4 = nT/2 + \check{s}^4 + o(1).$$

As a consequence of (3.63), two successive emission points are asymptotically separated by a uniform gap, in the sense that:

$$(3.64) \quad s_n^4 - s_n^1 = \check{s}^4 - \check{s}^1 + o(1), \quad s_{n+1}^1 - s_n^4 = T/2 + \check{s}^1 - \check{s}^4 + o(1).$$

Proof. The application $\lambda^{-1}(\cdot)$ is not defined on the whole image of $c(\mathbf{w})b_e \circ \bar{Y}_0(\mathbf{h}, y^1, y^2, \cdot)$. This is why the formulation (3.49) cannot be used here. Still, the system (3.30a)-(3.30b) is equivalent to:

$$(3.65) \quad \widetilde{\mathcal{F}}(s, \mu) = 0, \quad \widetilde{\mathcal{F}}(s, \mu) := c(\mathbf{w})b_e \circ \bar{Y}_0(\mathbf{h}, y^1, y^2, s) - \lambda(\mu\eta s + \mu Rsc(s)),$$

Choose $\mu_0 \in \mathbb{R}_+^*$ to recover all the aforementioned conditions, including (3.62) as well as $\mu_0 \leq (R_M^0)^{-1}\lambda^{-1}(\lambda_M - \delta)$. Let (s, ξ) be a position that is adjusted as indicated in (3.30b). Define $c_{10} := 2c_6^{-1}\lambda^{-1}(\lambda_M - \delta)$. We find that:

$$(3.66) \quad \lambda^{-1}(\lambda_M - \delta) \leq \mu\eta s + \mu Rsc(s) = |\xi| \leq R, \quad \forall s \in [c_{10}\mu^{-1}, C_1^{-1}R\mu^{-1}].$$

Select n as in Proposition 3.2. By construction:

$$\widetilde{\mathcal{F}}(nT/2, \mu) \geq c(\mathbf{w})b_M - \lambda_M > 0, \quad \widetilde{\mathcal{F}}((2n+1)T/4, \mu) < c(\mathbf{w})b_m - \lambda_M + \delta < 0.$$

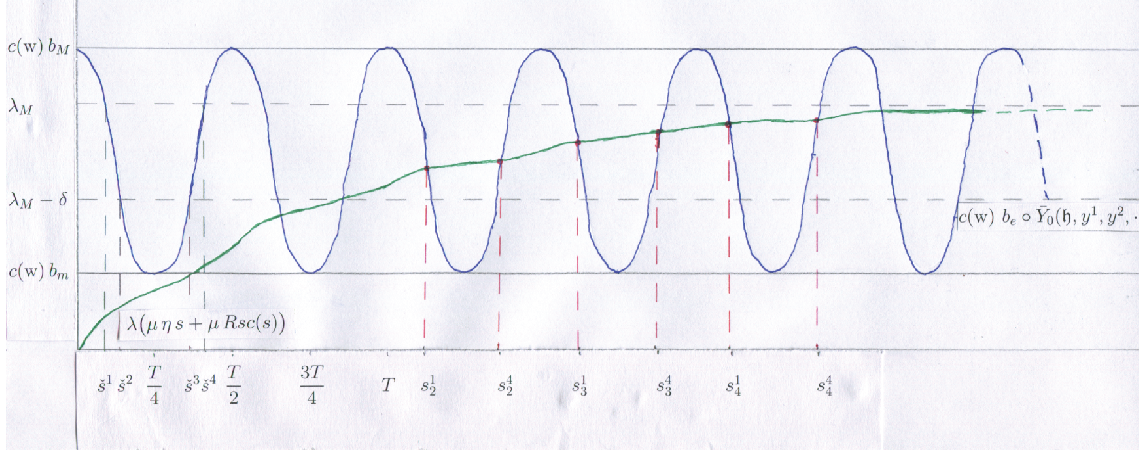


Figure 2. Case of overlapping values.

Display of the solutions to $\widetilde{\mathcal{F}}(s, \mu) = 0$.

Thus, there is at least one solution s_n^1 to (3.65) in the interval $[nT/2, (2n+1)T/4]$, and at least one solution s_n^4 to (3.65) in the interval $[(2n+1)T/4, (n+1)T/2]$. Due to (3.66), these solutions must satisfy:

$$\lambda_M - \delta \leq \lambda(\mu \eta s_n^1 + \mu Rsc(s_n^1)) = c(w) b_e \circ \bar{Y}_0(\mathfrak{h}, y^1, y^2, s_n^1),$$

$$\lambda_M - \delta \leq \lambda(\mu \eta s_n^4 + \mu Rsc(s_n^4)) = c(w) b_e \circ \bar{Y}_0(\mathfrak{h}, y^1, y^2, s_n^4).$$

This implies that:

$$s_n^1 \in]nT/2 + \check{s}^1, nT/2 + \check{s}^2[, \quad s_n^4 \in]nT/2 + \check{s}^3, nT/2 + \check{s}^4[.$$

From Assumption 3.1 and the condition (3.62), we can infer that:

$$(3.67a) \quad \begin{aligned} \partial_s \widetilde{\mathcal{F}}(s_n^1, \mu) &= \lambda' \circ Osc(s_n^1) [Osc'(s_n^1) - \mu \eta - \mu Rsc'(s_n^1)] \\ &\leq -\lambda' \circ Osc(s_n^1) \mu_0 (\eta + R_M^1) < 0, \end{aligned}$$

$$(3.67b) \quad \begin{aligned} \partial_s \widetilde{\mathcal{F}}(s_n^4, \mu) &= \lambda' \circ Osc(s_n^4) [Osc'(s_n^4) - \mu \eta - \mu Rsc'(s_n^4)] \\ &\geq \lambda' \circ Osc(s_n^4) \mu_0 (\eta + R_M^1) > 0. \end{aligned}$$

It follows that (3.65) has at most one solution s_n^1 in the interval $]nT/2, (2n+1)T/4[$, and at most one solution s_n^4 in the interval $](2n+1)T/4, (n+1)T/2[$.

By definition of the threshold limit λ_M , we have:

$$\lambda_M = \lim_{n \rightarrow +\infty} \lambda(\mu \eta s_n^j + \mu Rsc(s_n^j)) = \lim_{n \rightarrow +\infty} c(w) b_e \circ \bar{Y}_0(\mathfrak{h}, y^1, y^2, s_n^j), \quad \forall j \in \{1, 4\}.$$

By construction, we have:

$$\frac{d}{ds}(b_e \circ \bar{Y}_0)(\mathfrak{h}, y^1, y^2, \check{s}^1) < 0, \quad \frac{d}{ds}(b_e \circ \bar{Y}_0)(\mathfrak{h}, y^1, y^2, \check{s}^4) > 0.$$

The function $c(w) b_e \circ \bar{Y}_0(\mathfrak{h}, y^1, y^2, \cdot)$ is invertible in a neighbourhood of \check{s}^1 , with a continuous inverse. It takes the value λ_M at \check{s}^1 . Because it is periodic with period $T/2$, we have $c(w) b_e \circ \bar{Y}_0(\mathfrak{h}, y^1, y^2, s_n^1 - nT/2) = \lambda_M + o(1)$. Since $s_n^1 - nT/2 \in]0, T/4[$, by using the inverse alluded above, we can recover (3.63). A similar reasoning applies for s_n^4 . \square

3.2.7. *Quick comments about the propagation of singularities.* The Lagrangian manifold associated to the phase $\Psi(\pm, q; \cdot)$ is the set of points:

$$\mathfrak{L}(\pm, q; \cdot) := \left\{ (t, x, \nabla_{t,x} \Psi(\pm, q; t, x; s, y, \varsigma, \xi)); (s, y, \varsigma, \xi) \in \mathfrak{C}(\pm, q; t, x) \right\}.$$

As a by-product of (3.58) or (3.67), using the implicit theorem, it can be stated that the times $s(\pm, q; y)$ giving rise to emission points by applying Propositions 3.1 and 3.2 are smooth functions of (y^1, y^2) , not depending on y^3 . From (3.30b), extract:

$$\xi(\pm, q; y) := \mu q \int_0^{s(\pm, q; y)} \nabla_y (b_e \circ \bar{Y}_0)(0, y^1, y^2, \frac{\pi}{2}, r) dr.$$

The knowledge of $s(\pm, q; y)$ and $\xi(\pm, q; y)$ furnishes global parametrizations of $\mathfrak{L}(\pm, q; \cdot)$, by implementing the following mapping $\iota(\pm, q)$ where, for simplicity, the reference to (\pm, q) is skipped on the right-hand side:

$$\iota(\pm, q) : (t, y) \mapsto \begin{pmatrix} t \\ x \\ \partial_t \Psi \\ \nabla_x \Psi \end{pmatrix} = \begin{pmatrix} t \\ y \mp \mu (t - s(y)) \lambda'(|\xi(y)|) |\xi(y)|^{-1} \xi(y) \\ \pm \lambda(|\xi(y)|) \\ \mu^{-1} \xi(y) \end{pmatrix}.$$

The space-time projections of the bicharacteristic curves $t \mapsto \iota(\pm, q; t, y)$ give rise to the rays of geometric optics, which are straight lines travelling at the group velocity. To apply the stationary phase theorem to the integral $\mathcal{I}(\pm, q; \varepsilon, t, x)$ near some interior critical point, say $(s, y, \varsigma, \xi) \in \mathfrak{C}_i(\pm, q; t, x)$, the computation of the second derivative of $\Psi(\pm, q; t, x, \cdot)$ with respect to (s, y, ς, ξ) is needed, in order to avoid the caustic set, that is degenerate situations. The study of this aspect, together with the determination of a more accurate model for the dispersion relation $\lambda(\cdot)$, will be investigated in other contributions.

Acknowledgments. This work received the support of the *Agence Nationale de la Recherche* projet NOSEVOL (ANR 2011 BS01019 01) and ANR blanc DYFICOLTI.

REFERENCES

- [1] [http://en.wikipedia.org/wiki/Whistler_\(radio\)](http://en.wikipedia.org/wiki/Whistler_(radio)).
- [2] <http://alteaspace.wordpress.com/tag/van-allen-belts/>.
- [3] <http://farside.ph.utexas.edu/teaching/plasma/Plasmahtml/node51.html>.
- [4] Martin Braun. Mathematical remarks on the Van Allen radiation belt: a survey of old and new results. *SIAM Rev.*, 23(1):61–93, 1981.
- [5] A. J. Brizard and T. S. Hahm. Foundations of nonlinear gyrokinetic theory. *Rev. Modern Phys.*, 79(2):421–468, 2007.
- [6] J.-Y. Chemin, B. Desjardins, I. Gallagher, and E. Grenier. *Mathematical geophysics*, volume 32 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press Oxford University Press, Oxford, 2006. An introduction to rotating fluids and the Navier-Stokes equations.
- [7] Christophe Cheverry. Cascade of phases in turbulent flows. *Bull. Soc. Math. France*, 134(1):33–82, 2006.
- [8] Christophe Cheverry, Isabelle Gallagher, Thierry Paul, and Laure Saint-Raymond. Semiclassical and spectral analysis of oceanic waves. *Duke Math. J.*, 161(5):845–892, 2012.
- [9] Rui Dilão and Rui Alves-Pires. Chaos in the Störmer problem. In *Differential equations, chaos and variational problems*, volume 75 of *Progr. Nonlinear Differential Equations Appl.*, pages 175–194. Birkhäuser, Basel, 2008.

- [10] Dungey. Loss of van allen electrons due to whistlers. *Planet Space Sci*, 11, 1963.
- [11] E. Frénod and E. Sonnendrücker. Long time behavior of the two-dimensional Vlasov equation with a strong external magnetic field. *Math. Models Methods Appl. Sci.*, 10(4):539–553, 2000.
- [12] Isabelle Gallagher and Laure Saint-Raymond. On pressureless gases driven by a strong inhomogeneous magnetic field. *SIAM J. Math. Anal.*, 36(4):1159–1176 (electronic), 2005.
- [13] R. T. Glassey and J. W. Schaeffer. Global existence for the relativistic Vlasov-Maxwell system with nearly neutral initial data. *Comm. Math. Phys.*, 119(3):353–384, 1988.
- [14] E. Grenier. Pseudo-differential energy estimates of singular perturbations. *Comm. Pure Appl. Math.*, 50(9):821–865, 1997.
- [15] Helliwell. *Whistlers and Related Ionospheric Phenomena*. Stanford University Press, 1965.
- [16] Jean-Luc Joly, Guy Métivier, and Jeffrey Rauch. Nonlinear oscillations beyond caustics. *Comm. Pure Appl. Math.*, 49(5):443–527, 1996.
- [17] Jean-Luc Joly, Guy Metivier, and Jeffrey Rauch. Transparent nonlinear geometric optics and Maxwell-Bloch equations. *J. Differential Equations*, 166(1):175–250, 2000.
- [18] S. Le Bourdier. *Méthodes déterministes de résolution des équations de Vlasov-Maxwell relativistes en vue du calcul de la dynamique des ceintures de Van Allen*. PhD thesis, Ecole Centrale Paris, 2007.
- [19] G. Métivier and S. Schochet. The incompressible limit of the non-isentropic Euler equations. *Arch. Ration. Mech. Anal.*, 158(1):61–90, 2001.
- [20] Santolik O. New results of investigations of whistler-mode chorus emissions. *Nonlin. Processes Geophys.*, 15:621–630, 2008.
- [21] P. Pfannmöller. *Whistler Wave Propagation in Inhomogeneous Plasmas*. PhD thesis, Ernst-Moritz-Arndt-Universität Greifswald, 2011.
- [22] Pandey R. S., Srivastava U. C., Srivastava A. K., Kumar S., and Singh D. K. Pitch angle loss-cone anisotropic magneto plasma in presence of parallel electric a.c. field. *Archives of Physics Research*, 1:126–136, 2010.
- [23] Steven Schochet. Fast singular limits of hyperbolic PDEs. *J. Differential Equations*, 114(2):476–512, 1994.
- [24] Serra. Nonlinear shift of wave parameters. *Planet Space -Sci*, 32, 1984.
- [25] L. R. O. Storey. An investigation of whistling atmospherics. *Phil. Trans. Roy. Soc.*, 246, 1953.
- [26] A. Tenerani. *Dynamics of ion-scale coherent magnetic structures and coupling with whistler waves during substorms*. PhD thesis, Université Pierre et Marie Curie, 2012.
- [27] C. E. Weatherburn. On Lamé families of surfaces. *Ann. of Math. (2)*, 28(1-4):301–308, 1926/27.

(Ch. Cheverry) INSTITUT MATHÉMATIQUE DE RENNES, CAMPUS DE BEAULIEU, 263 AVENUE DU GÉNÉRAL LECLERC CS 74205 35042 RENNES CEDEX, FRANCE

E-mail address: christophe.cheverry@univ-rennes1.fr