THE RELATIVISTIC VLASOV MAXWELL EQUATIONS FOR STRONGLY MAGNETIZED PLASMAS

CHRISTOPHE CHEVERRY AND SLIM IBRAHIM

Abstract. An important challenge in plasma physics is to determine whether ionized gases can be confined by strong magnetic fields. After properly formulating the model, this question leads to a penalized version of the Relativistic Vlasov Maxwell system, marked by the role of a singular factor ε^{-1} corresponding to the inverse of a cyclotron frequency. In this paper, we prove in this context the existence of classical $\mathcal{C}^1$-solutions for a time independent of ε. We also investigate the stability of these smooth solutions.

Keywords. Kinetic equations ; Vlasov-Maxwell system ; Magnetized plasmas ; Lifespan of classical solutions ; Momentum support condition ; Energy estimates.

Contents

1. Introduction

Given a small parameter $\varepsilon > 0$, in this paper we analyze the well-posedness of the Magnetized Relativistic Vlasov-Maxwell (MRVM) system

$$\begin{aligned} \partial_t f + [\nu(\varepsilon\xi)\cdot\nabla_x] f - \frac{1}{\varepsilon^2}\,[\nu(\varepsilon\xi)\times \mathbf{B}_e(x)]\cdot\nabla_\xi f \\ = -M'(|\xi|)\,\frac{\xi\cdot\boldsymbol{E}}{|\xi|} + [\boldsymbol{E} + \nu(\varepsilon\xi)\times\boldsymbol{B})]\cdot\nabla_\xi f \end{aligned} \tag{1.1}$$

$$\nabla_x\cdot\boldsymbol{E} = -Q(f) \quad ; \quad \partial_t\boldsymbol{E} - \nabla\times\boldsymbol{B} = \boldsymbol{J}(f) \tag{1.2}$$

$$\nabla_x\cdot\boldsymbol{B} = 0 \quad ; \quad \partial_t\boldsymbol{B} + \nabla_x\times\boldsymbol{E} = 0. \tag{1.3}$$

Here, x and $\nu = \frac{\xi}{\sqrt{1+|\xi|^2}}$ are points in $\mathbf{R}^3$ representing position and velocity of charged particles (electrons), respectively. The unknown of system (1.1)-(1.3) are a density function $f(t,x,\xi)$

Date: today.

defined on $\mathbf{R}_t \times \mathbf{R}^3_x \times \mathbf{R}^3_\xi$, and a self generated electro-magnetic field $(\boldsymbol{E}, \boldsymbol{B})(t,x)$. Particles are out of a thermal equilibrium where velocity repartitions can be approximated by a Maxwell-Boltzmann distribution $M(\cdot)$ with small density (See Assumptions 2.2, and 2.3). The total charge Q and the current $\boldsymbol{J}$ are defined by

$$Q \equiv Q(f)(t,x) := \int f(t,x,\xi) d\xi \tag{1.4}$$

$$\boldsymbol{J} \equiv \boldsymbol{J}_\varepsilon(f)(t,x) := \int \nu(\varepsilon\xi) f(t,x,\xi) d\xi. \tag{1.5}$$

The system (1.1)-(1.3) is written after adimensionalization. All physical constants, except the small parameter $\varepsilon \ll 1$ which stands for the inverse of the electron cyclotron frequency, are normalized to one. We supplement (1.1)-(1.3) with initial conditions f^{in}, E^{in}, and B^{in}.

In statistical physics, the relativistic Vlasov-Maxwell system is a kinetic mean-field model for collisionless plasmas. It is commonly used in the context of planetary magnetospheres or fusion devices. In such applications, the plasmas are confined by a strong external magnetic field that is completely prescribed, and that is commonly represented by a vector valued spatial function of the form $x \mapsto \varepsilon^{-1}\mathbf{B_e}(x)$. The amplitude of the function $\mathbf{B_e}(\cdot)$ is of size one. The field $\mathbf{B_e}(\cdot)$ is usually represented by the dipole model when dealing with magnetospheres [6], and it can be derived from the knowledge of magnetic surfaces when studying tokamaks [7].

As it will be explained in Section 2, the study of the MRVM system (1.1)-(1.3) is a relevant way to describe phenomena occurring in magnetized, cold, dilute, neutral gases which are taken out of equilibrium. It allows to take into account many physical phenomena, especially in the framework of space plasmas. Our main goal here is to study the wellposedness and the stability of solutions to the Cauchy problem associated to (1.1)-(1.3). Since our problem depends on a small parameter, it is crucial to show the existence of solutions on a uniform time.

The Cauchy problem associated to the Relativistic Vlasov-Maxwell (RVM) system, which does not takes into account the influence of $\mathbf{B_e}(\cdot)$, has been extensively studied. A review is provided in the monograph [14]. Local existence and uniqueness of classical solutions for smooth, compactly supported data was established in [16]. Global existence of smooth solutions has been obtained for small data [17], for nearly neutral data [15] and in other different contexts [29]. But, in the case of large data, the global existence of solutions to the RVM system is still an unresolved problem. In addition, the Cauchy problem as well as the non-relativistic limit equation were studied in [1, 11, 30].

The above contributions related to global existence of classical solutions heavily rely on the spreading of the bicharacteristics (defined by (3.65, 3.66)) associated to the left part of (1.1), which is essential to induce a sort of decoupling between the density f and the electromagnetic field $(\boldsymbol{E}, \boldsymbol{B})$. However, in the presence of a strong magnetic field, such a spreading is not available. On the contrary, the bicharacteristics stay for a very long time in a compact set; they involve large amplitude oscillations [6, 7]; and, as a consequence, they enforce strong interactions between f and $(\boldsymbol{E}, \boldsymbol{B})$, which are the potential source of instabilities.

From a mathematical perspective, our problem is to study families of solutions to the RVM system that are generated by *large data*. This is reflected at the level of the MRVM system into the singular weight $\varepsilon^{-2}\nu(\varepsilon\xi) = \mathcal{O}(\varepsilon^{-1})$. A major difficulty arises because of this singular factor being placed in front of a differential operator with *variable coefficients* with respect to both variables x and ξ. This feature together with the large initial condition

$$\partial_t f|_{t=0} = \frac{1}{\varepsilon^2}\left[\nu(\varepsilon\xi) \times \mathbf{B}_e(x)\right] \cdot \nabla_\xi f^{in} + \mathcal{O}(1) = \mathcal{O}\Big(\frac{1}{\varepsilon}\Big), \quad 0 < \varepsilon \ll 1 \tag{1.6}$$

may compromise the existence of uniform Lipschitz estimates. To deal with (1.6), the initial data may be *prepared* (in the sense of Definition 4.2) to make the above first time-derivative uniformly bounded. Or, as expected in (1.6), the data may be *general* which clearly indicates the presence of large amplitude oscillations, and therefore the occurrence of large Lipschitz norms of both the density f, and the field $\boldsymbol{E}$ and $\boldsymbol{B}$. In the context of such large data, the existence of solutions to both RVM and MRVM systems on a uniform time interval $[0,T]$ is not at all evident. The main result of this paper is the following.

Theorem 1.1. *Assume that the background Boltzmann distribution M is in $\mathcal{C}_c^\infty(\mathbb{R}_+;\mathbb{R}_+)$, that the initial distribution $f^{in} \in \mathcal{C}^1(\mathbb{R}^3)$ is confined, and that $(E^{in},B^{in}) \in \mathcal{C}^1(\mathbb{R}^3)$ fits with f^{in} in the sense of (2.27). Then, the Cauchy problem for the MRVM system* (1.1)-(1.3) *is uniformly locally well-posed in the sense of Definition 4.4. Moreover, prepared data give rise to families of solutions which are uniformly bounded in the Lipschitz norm.*

The above theorem is part of a long tradition of works on the RVM system, going back to [15, 16]. It is also connected to problems arising in fast rotating fluids [2, 3, 5, 12, 19] or in nonlinear geometric optics [25, 26], which are scientific domains where questions about uniform estimates for large oscillating data have been and are commonly investigated.

The contributions related to [5, 25] deal with general hyperbolic nonlinear systems. Of course, the corresponding results could be applied to more specific situations, like the actual MRVM system. But they require a lot of prerequisites, among which more restricted prepared data and regularity assumptions which are going far beyond the actual $\mathcal{C}^1-$context; they do not take into account many peculiarities of the Vlasov and Maxwell equations, which will allow us to refine the standard statements; they do not care about the momentum support condition, which here plays a crucial part; and so on. In fact, there is much to do in this paper to adapt the approaches coming from [5, 25] to the framework inspired by [15, 16].

The paper is organized as follows. In section 2, a detailed derivation of (1.1)-(1.3) from the classical RVM system will be given. The proof of Theorem 1.1 hides a number of new difficulties which, after a work of preparation in Section 3, are solved in Section 4. Taking into account the material introduced in Subsection 4.1, we prove in Subsection 4.2 uniform L^∞-estimates on the family of solutions, from which the uniform lifespan (2.30) and the uniform confinement property (2.31) follow (Proposition 4.6). In Subsection 4.3, we control some weighted Lipschitz norm of the solutions (Proposition 4.9); then, we restrict our attention to the case of prepared data, and we get a uniform bound on the Lipschitz norm (Proposition 4.12).

Acknowledgement. A part of this work was done while C.C. was visiting the Department of Mathematics and Statistics of the University of Victoria, and S.I. was visiting IRMAR, the "Institut de Recherche Mathématique de Rennes". They both thank all members and staff at the two institutions for their warm hospitality. They also thank Dayton Preissl for interesting discussions about the text. Both C.C and S.I. were supported by France-Canada Research Fund. S.I. was supported by NSERC grant (371637-2014).

2. Modeling of collisionless magnetized plasmas

Subsection 2.1 is inspired by theoretical considerations [8] about magnetospheres [6], stars and fusion devices [7]. We show that the description of *real magnetized plasmas* forces to transform the classical Relativistic Vlasov Maxwell system (the so-called RVM system) into a Magnetized Relativistic Vlasov Maxwell system (the MRVM system), involving a large parameter ε^{-1}. In Subsection 2.2, this MRVM system is interpreted as a Valsov-Wave system (VW system).

2.1. **From the RVM system to the MRVM system.** The RVM system is built in coupling the Vlasov equation and the Maxwell's equations. It is applied here in a physical framework based on concrete considerations. This means to retain a number of specific assumptions, giving rise to special issues. These hypotheses are first and foremost related to the presence of a strong external magnetic field (Paragraph 2.1.1). They also imply a cold and small density assumption and some neutrality condition (Paragraph 2.1.2). At the end, this furnishes a formulation of the

RVM system, called the MRVM system, introduced in Paragraph 2.1.3. Open related questions are raised in Paragraph 2.1.4.

2.1.1. *The impact of a strong external magnetic field.* In view of a better understanding of what happens in magnetospheres [9] or fusion devices [10], it is important to consider the influence of a strong exterior inhomogeneous magnetic field, denoted by $\varepsilon^{-1}\,\mathbf{B}_e(x)$. "Strong" because the parameter ε is small; in practice, the dimensionless number ε stands for the inverse of the electron cyclotron frequency; it is often of size $\simeq 10^{-4}$. "Exterior" because the field is prescribed. "Inhomogeneous" because the function $\mathbf{B}_e(\cdot)$ does depend on the spatial variable $x \in \mathbb{R}^3$.

Assumption 2.1. [strong inhomogeneous magnetic field] *The function $\mathbf{B}_e(\cdot)$ is assumed to be smooth, with bounded derivatives, that is $\mathbf{B}_e \in \mathcal{C}_b^\infty(\mathbb{R}^3)$. It is of size one and does not vanish on all compact sets. More precisely, for all compact sets $K \subset \mathbb{R}^3$, there exists a positive constant $c \equiv c(K)$ such that*

$$\forall x \in K, \quad c(K) \le \mathrm{b}_e(x) \le c(K)^{-1} \quad ; \quad \mathrm{b}_e(x) := |\mathbf{B}_e(x)| \tag{2.1}$$

Moreover, it is divergence and curl free

$$\forall x \in \mathbb{R}^3, \quad \nabla_x \cdot \mathbf{B}_e(x) \equiv 0 \quad ; \quad \nabla_x \times \mathbf{B}_e(x) \equiv 0 \tag{2.2}$$

Note that the condition (2.2) is satisfied in the case of dipole models, like for the Earth's magnetic field [9]. We consider that there is only one species, say electrons in a background of stationary protons. These electrons are described by a scalar distribution function $\mathrm{f}(t,x,\xi)$ that gives at the time $t \in \mathbb{R}_+$ their probability density on the phase space $\mathbb{R}^3_x \times \mathbb{R}^3_\xi$. As usual, we denote $\nu(\xi)$ the velocity (with the speed of light normalized to one) and $\langle\xi\rangle$ the Lorentz factor

$$\forall \xi \in \mathbb{R}^3, \quad \nu(\xi) := \frac{\xi}{\langle\xi\rangle} \quad ; \quad 1 \le \langle\xi\rangle := \sqrt{1+|\xi|^2} \quad ; \quad |\nu(\xi)| < 1 \tag{2.3}$$

In this article, we will focus on the electron cyclotron regime, when $\varepsilon \ll 1$. Then, the motion of electrons is governed by the penalized Vlasov equation

$$\partial_t \mathrm{f} + [\nu(\xi)\cdot\nabla_x]\mathrm{f} = \big[\mathbf{E} + \nu(\xi)\times\big(\varepsilon^{-1}\mathbf{B}_e(x) + \mathbf{B}\big)\big]\cdot\nabla_\xi \mathrm{f} \tag{2.4}$$

The electromagnetic field $(\mathbf{E}, \varepsilon^{-1}\mathbf{B}_e + \mathbf{B})$ inside (2.4) depends only on (t,x), and it takes its values in $\mathbb{R}^3 \times \mathbb{R}^3$. It must satisfy Maxwell's equations. In view of (2.2), this means that the *self-consistent* electromagnetic field $(\mathbf{E},\mathbf{B})$ is satisfies

$$\partial_t \mathbf{E} - \nabla_x \times \big(\varepsilon^{-1}\mathbf{B}_e(x) + \mathbf{B}\big) = \partial_t \mathbf{E} - \nabla\times\mathbf{B} = \mathbf{J}(\mathrm{f}) \tag{2.5}$$

$$\partial_t\big(\varepsilon^{-1}\mathbf{B}_e(x) + \mathbf{B}\big) + \nabla_x \times \mathbf{E} = \partial_t \mathbf{B} + \nabla_x \times \mathbf{E} = 0 \tag{2.6}$$

and the compatibility conditions

$$\nabla_x \cdot \mathbf{E} = \rho_i - \rho(\mathrm{f}) \tag{2.7}$$

$$\nabla_x \cdot \big(\varepsilon^{-1}\mathbf{B}_e(x) + \mathbf{B}\big) = \nabla_x \cdot \mathbf{B} = 0 \tag{2.8}$$

In (2.7), the constant ρ_i represents the density of charge issued from ions. The expressions $\rho(\mathrm{f})$ and $\mathbf{J}(\mathrm{f})$ stand for the electron density of charge and the electric current, respectively. They can be computed according to

$$\rho(\mathrm{f})(t,x) = \int \mathrm{f}(t,x,\xi)\,d\xi \tag{2.9}$$

$$\mathbf{J}(\mathrm{f})(t,x) = \int \nu(\xi)\mathrm{f}(t,x,\xi)\,d\xi \tag{2.10}$$

We say that the vector valued function $\mathbf{U} := (\mathrm{f},\mathbf{E},\mathbf{B})$ is a solution to the RVM system if it satisfies the evolution equations (2.4,2.5,2.6) together with the compatibility conditions (2.7,2.8), where $\rho(\cdot)$ and $\mathbf{J}(\cdot)$ are as in (2.9,2.10). The RVM system is a well-established model for describing the time evolution in collisionless strongly magnetized plasmas.

2.1.2. *Stationary solutions.* We will consider a ionized gas that is a perturbation of a plasma at thermal equilibrium, characterized by $\mathbf{U} \equiv \mathbf{U}^s := (\mathrm{f}^s, \mathbf{E}^s, \mathbf{B}^s)$ with

$$\mathrm{f}^s(t,x,\xi) \equiv \mathrm{f}^s(\xi) := \varepsilon^{-2} M(\varepsilon^{-1}|\xi|), \qquad \mathbf{E}^s(t,x) := 0, \qquad \mathbf{B}^s(t,x) := 0 \tag{2.11}$$

Assumption 2.2. [cold and small density assumption] *The function $M(\cdot)$ is in $\mathcal{C}_c^\infty(\mathbb{R}_+;\mathbb{R}_+)$.*

In plasma physics, "cold" means that the velocities of most electrons are small in comparison to the speed of light or that the temperature of most electrons is a few electronvolts. The cold assumption is often used to model astrophysical plasmas or even plasmas located at the edge of fusion devices. As explained in [9, 10], this implies that the distribution function $\mathrm{f}(\cdot)$ is concentrated for velocities ξ such that $|\xi| \sim O(\varepsilon)$. This is reflected in Assumption 2.2 by the fact that the function $M(\cdot)$ is compactly supported (in ξ)

$$\exists R_M^{in} \in \mathbb{R}_+^*; \quad \sup M \subset [0, R_M^{in}] \tag{2.12}$$

From (2.12), it follows that

$$\| M \|_1 := \int M(|\xi|) d\xi < +\infty \quad ; \quad \| M' \|_1 := \int |M'(|\xi|)| d\xi < +\infty \tag{2.13}$$

Another ingredient of (2.11) is the size of the amplitude, which implies that the density of the plasma is "small". Thus, the plasma is dilute, in the sense that

$$\rho(\mathrm{f}^s)(t,x) = \int \mathrm{f}^s(t,x,\xi)\, d\xi = \varepsilon \| M \|_1 = O(\varepsilon) \tag{2.14}$$

The distribution function $\mathrm{f}^s(\cdot)$ depends only on $|\xi|$, and therefore we have (2.4); it is even in ξ while $\nu(\cdot)$ is odd, and thereby we have $\mathbf{J}(\mathrm{f}^s) \equiv 0$. Now, to obtain (2.7) with $\mathbf{E}^s \equiv 0$, the constant $\rho_i \equiv \rho_i(\varepsilon)$ must be adjusted accordingly.

Assumption 2.3. [neutrality assumption] *We impose :*

$$\rho_i = \rho(\mathrm{f}^s) = \varepsilon \| M \|_1 \tag{2.15}$$

Under (2.15), the expression $\mathbf{U}^s(\cdot)$ is a stationary solution to the RVM system. It can also be viewed as a solution to the RVM system associated with the initial data

$$(\mathrm{f}^s, \mathbf{E}^s, \mathbf{B}^s)_{|t=0} = \big(\varepsilon^{-2} M(\varepsilon^{-1}|\xi|), 0, 0\big) \tag{2.16}$$

Note that more general stationary solutions could be considered. In [10], a notion of shifted Maxwell-Boltzmann distribution is introduced. This allows to describe plasmas confined inside tokamaks. Then, the curl-free condition (2.2) is not required, but the density $\mathrm{f}^s(\cdot)$ turns to be more complicated than in (2.11), and the electromagnetic field $(\mathbf{E}^s, \mathbf{B}^s)$ is non zero. Here, for the sake of simplicity, we will stick to the choice (2.11).

2.1.3. *Perturbation theory.* Descriptions of cold plasmas through representations like (2.11) are rather restrictive. In reality, the observed self-consistent electromagnetic field $(\mathbf{E}, \mathbf{B})$ is non zero. Experimental measures indicate that $(\mathbf{E}, \mathbf{B}) \not\equiv (0,0)$, and it is clear that many important phenomena are linked to discrepancies from $(\mathrm{f}^s, 0, 0)$. Then, we can say that the plasma is out of equilibrium [8]. Since electrons are much lighter than ions, they move quicker. Thus, plasma phenomena out of equilibrium are mainly concerned with electrons moving in a (steady) background of ions. This allows for a focus on the time evolution of only one species of particles, namely electrons.

Away from thermal equilibrium, the probability density of electrons can differ from (2.11). Let $f(t,x,\varepsilon^{-1}\xi)$ be the distribution function which indicates at the time t in the phase space $\mathbb{R}_x^3 \times \mathbb{R}_\xi^3$

the divergence from (2.11). We look at solutions of (2.4,2.5,2.6) which represent fluctuations near the thermal equilibrium $(\mathrm{f}^s, 0, 0)$. Thus, at time $t = 0$, we impose

$$\mathrm{f}_{|t=0} = \mathrm{f}^{in} := \varepsilon^{-2}\, M(\varepsilon^{-1}\,|\mathrm{\xi}|) + \varepsilon^{-2}\ f^{in}(x, \varepsilon^{-1}\,\mathrm{\xi}) \tag{2.17}$$

$$\mathbf{E}_{|t=0} = \mathbf{E}^{in} := \varepsilon\, \boldsymbol{E}^{in} \tag{2.18}$$

$$\mathbf{B}_{|t=0} = \mathbf{B}^{in} := \varepsilon\, \boldsymbol{B}^{in} \tag{2.19}$$

and, consequently, we seek $\mathrm{f}(\cdot)$, $\mathbf{E}(\cdot)$ and $\mathbf{B}(\cdot)$ in the form

$$\mathrm{f}(t, x, \mathrm{\xi}) = \varepsilon^{-2}\, M(\varepsilon^{-1}\,|\mathrm{\xi}|) + \varepsilon^{-2}\ f(t, x, \varepsilon^{-1}\,\mathrm{\xi})\,, \qquad \mathbf{E} = \varepsilon\, \boldsymbol{E}\,, \qquad \mathbf{B} = \varepsilon\, \boldsymbol{B} \tag{2.20}$$

Expressed in terms of the new functions $f(t, x, \xi)$, $\boldsymbol{E}(t, x)$ and $\boldsymbol{B}(t, x)$, the system (2.4,2.5,2.6) can be decomposed into the transport equation

$$\begin{aligned} \partial_t f + [\nu(\varepsilon\xi) \cdot \nabla_x] f - \frac{1}{\varepsilon^2}\, [\nu(\varepsilon\xi) \times \mathbf{B}_e(x)] \cdot \nabla_\xi f \\ = M'(|\xi|)\, \frac{\xi \cdot \boldsymbol{E}}{|\xi|} + [\boldsymbol{E} + \nu(\varepsilon\xi) \times \boldsymbol{B})] \cdot \nabla_\xi f \end{aligned} \tag{2.21}$$

along with

$$\nabla_x \cdot \boldsymbol{E} = -Q(f) \quad ; \quad \partial_t \boldsymbol{E} - \nabla \times \boldsymbol{B} = \boldsymbol{J}(f) \tag{2.22}$$

$$\nabla_x \cdot \boldsymbol{B} = 0 \quad ; \quad \partial_t \boldsymbol{B} + \nabla_x \times \boldsymbol{E} = 0 \tag{2.23}$$

where

$$Q \equiv Q(f)(t, x) := \int f(t, x, \xi) d\xi \tag{2.24}$$

$$\boldsymbol{J} \equiv \boldsymbol{J}_\varepsilon(f)(t, x) := \int \nu(\varepsilon\xi) f(t, x, \xi) d\xi \tag{2.25}$$

We want to bring the reader's attention about the passage from the RVM system (2.4,2.5,2.6) to (2.21,2.22,2.23). There are changes taking place: first and foremost, the variable $\mathrm{\xi}$ is replaced by $\xi := \varepsilon^{-1}\mathrm{\xi}$ and the cold assumption becomes $|\xi| \leq R_M^{in}$; second, the singular factor $\varepsilon^{-1}\nu(\mathrm{\xi})$ is exchanged with $\varepsilon^{-2}\nu(\varepsilon\xi)$ which is still some $\mathcal{O}(\varepsilon^{-1})$; thirdly, there is the additional semilinear source term implying $\xi \cdot \boldsymbol{E}$ and coming from the perturbation procedure. To highlight these differences, the system built with (2.21,2.22,2.23) will be called the MRVM system, the first "M" being for *magnetized*. From now on, the unknown is $\boldsymbol{U} := (f, \boldsymbol{E}, \boldsymbol{B})$.

By construction, $\boldsymbol{U} \equiv 0$ is a special solution to the MRVM system with initial condition $\boldsymbol{U}^{in} \equiv 0$. Now, at time $t = 0$, we modify this initial data. In other words, we impose

$$\boldsymbol{U}_{|t=0} = \boldsymbol{U}^{in} \equiv (f^{in}, \boldsymbol{E}^{in}, \boldsymbol{B}^{in}) \in \mathcal{C}_c^1(\mathbb{R}^3 \times \mathbb{R}^3) \times \mathcal{C}_c^2(\mathbb{R}^3) \times \mathcal{C}_c^2(\mathbb{R}^3) \tag{2.26}$$

with $\boldsymbol{U}^{in} \not\equiv 0$. Of course, the expression $\boldsymbol{U}^{in}$ must be compatible, that is

$$\nabla_x \cdot \boldsymbol{E}^{in} = -\rho(f^{in}) = -\int f^{in}(x, \xi) d\xi \quad ; \quad \nabla_x \cdot \boldsymbol{B}^{in} = 0 \tag{2.27}$$

It is worth noting that $f^{in}(\cdot)$ and $f(\cdot)$ are real valued functions without sign condition. As a matter of fact, contrary to f, the expressions f^{in} and f do not represent (positive) densities but perturbations of densities. The two constraints inside (2.27) are propagated by the equations. In other words, assuming (2.27), the trace $\boldsymbol{U}(t, \cdot) = (f, \boldsymbol{E}, \boldsymbol{B})(t, \cdot)$ of a smooth solution will satisfy for all time $t \geq 0$ the condition

$$\nabla_x \cdot \boldsymbol{E} = -\rho(f) = -\int f(t, x, \xi) d\xi \quad ; \quad \nabla_x \cdot \boldsymbol{B} = 0. \tag{2.28}$$

As mentioned in the introduction for large initial data, the existence of solutions to both RVM and MRVM systems on a uniform time interval $[0, T]$ is not at all evident. In the next paragraph, we explain more precisely why is this.

2.1.4. *Open-ended questions about lifespan, confinement and stability, and their responses.* Real plasmas are contained in a finite volume and the cold assumption means that we focus on bounded velocities, with $|\xi| < +\infty$. Thus, at time $t = 0$, it is natural to impose the following.

Assumption 2.4. [confinement assumption] *The initial data $f^{in}(\cdot)$ is in $\mathcal{C}^1_c(\mathbb{R}^3 \times \mathbb{R}^3)$. It is compactly supported both in position $x \in \mathbb{R}^3$ and velocity $\xi \in \mathbb{R}^3$. More precisely*

$$\exists\, (\mathrm{R}^{in}, R^{in}) \in (\mathbb{R}^*_+)^2 \,; \qquad \sup f^{in}(\cdot) \subset \big\{ (x,\xi)\,;\, |x| \leq \mathrm{R}^{in}\,,\ |\xi| \leq R^{in} \big\} \tag{2.29}$$

As noticed in [6], the presence of a strong magnetic field in MRVM prevents bicharacteristics spreading. On the contrary, the bicharacteristics associated to the left part of (2.21), that is the bicharacteristics defined by (3.65, 3.66), stay for a very long time in a compact set; they involve large amplitude oscillations [6, 7]; and, by this way, they enforce strong interactions between f and $(\boldsymbol{E}, \boldsymbol{B})$, which are the potential source of instabilities.

In the case of large data, the global existence of solutions to the RVM system is a problem which is still unresolved. And therefore, the same applies to the MRVM system. Technically, the main difficulty arises through the singular factor ε^{-1} that appears inside (2.21) and (1.6). To our knowledge, current results furnish a finite lifespan T_ε, which can shrink to zero at the speed $T_\varepsilon \sim \varepsilon$. Not being able to prove that $T_\varepsilon = +\infty$, in view of applications, it would however be very interesting to know if T_ε can be uniformly bounded from below. This would be a rigorous intrusion in the domain of large amplitude oscillating $\mathcal{C}^1$-solutions to the RVM system, and this is our *first question*. Do we have

$$\exists T \in \mathbb{R}^*_+; \quad \forall \varepsilon \in]0,1), \quad 0 < T \leq T_\varepsilon \tag{2.30}$$

Another key result of Glassey-Strauss [16] shows that the solutions can be extended as long as the momentum support of $f \equiv f_\varepsilon$ remains bounded. Extensions of this criterion can be found in [24, 31]. Now, assuming (2.30), the *second question* which is related to (2.30) is about the existence of a uniform confinement. We would like to determine whether there exists bounded functions $\mathrm{R}(\cdot)$ and $R(\cdot)$ in $L^\infty([0,T])$ such that

$$\forall (\varepsilon, t) \in]0,1] \times [0,T], \quad \sup f_\varepsilon(t,\cdot) \subset \big\{ (x,\xi)\ ;\ |x| \leq \mathrm{R}(t)\,,\, |\xi| \leq R(t) \big\} \tag{2.31}$$

Denoting $\mathrm{R}^\infty \in \mathbb{R}^*_+$ and $R^\infty \in \mathbb{R}^*_+$ the sup norms of $\mathrm{R}(\cdot)$ and $R(\cdot)$, respectively this means to deal at any time $t \in [0,T]$ with the momentum support condition

$$\sup f(\cdot) \subset \big\{ (x,\xi)\ ;\ |x| \leq \mathrm{R}^\infty\,,\, |\xi| \leq R^\infty \big\} \tag{2.32}$$

The properties (2.28) and (2.32) are expected, and thereby we will work within the framework of classical compatible solutions, that is with

$$\begin{aligned} \mathcal{X} := \big\{ \boldsymbol{U} = (f, \boldsymbol{E}, \boldsymbol{B}) \in \mathcal{C}^1(\mathbb{R}^3 \times \mathbb{R}^3) \times \mathcal{C}^1(\mathbb{R}^3) \times \mathcal{C}^1(\mathbb{R}^3)\,; \\ \text{the two conditions (2.28) and (2.32) are verified for some } R^\infty \big\} \end{aligned} \tag{2.33}$$

A *third question* is related to the stability properties of the solutions thus exhibited. In the continuation of [28], we want to determine how the Lipschitz norm can deteriorate when ε goes to zero, and we want to measure how the difference (measured in relevant norms) between solutions can change over time. Our main result Theorem 1.1 gives answers to all of them.

Before getting into the substance of the text, preliminary steps are required. This starts in Subsection 2.2 with a reformulation of the MRVM system as a Vlasov Wave system (VW system).

2.2. **From the MRVM system to a VW system.** We adopt here the approach of [4, 27], with some necessary adaptations induced by the magnetized, small density and perturbative context. As in [4, 27], we seek in Paragraph 2.2.1 to write the electromagnetic field $(\boldsymbol{E}, \boldsymbol{B})$ in terms of a special electromagnetic four-potential $(\Phi, \boldsymbol{A})$, called the *Lienard-Wiechert potential*. As will be seen in Paragraph 2.2.2, this scalar potential Φ and this vector potential $\boldsymbol{A}$ are the solutions of a particular wave-type equation.

2.2.1. *Choice of the Lorenz Gauge.* In this paragraph, the discussion is completely general of solutions $(\boldsymbol{E},\boldsymbol{B})$ to Maxwell's equations (2.22,2.23) with charge and current densities Q and $\boldsymbol{J}$ as in (2.24,2.25). It does not explicitly involve the Vlasov equation (2.21). From the first condition inside (2.23), we have that $\nabla_x \cdot \boldsymbol{B} = 0$ so that $\boldsymbol{B} = \nabla_x \times \boldsymbol{A}$ for some vector field $\boldsymbol{A} : \mathbb{R} \times \mathbb{R}^3 \mapsto \mathbb{R}^3$ known as the *vector potential*. For the same reason, we can find a vector potential $\boldsymbol{A}^{in}$ such that

$$\boldsymbol{B}^{in} = \nabla_x \times \boldsymbol{A}^{in} \quad ; \quad \nabla_x \cdot \boldsymbol{A}^{in} = 0 \tag{2.34}$$

We can also rewrite the electric field in terms of a *scalar potential* $\varPhi : \mathbb{R} \times \mathbb{R}^3 \mapsto \mathbb{R}$ according to

$$\begin{aligned} \partial_t \boldsymbol{B} + \nabla_x \times \boldsymbol{E} = 0 \quad \Leftrightarrow \quad \partial_t(\nabla_x \times \boldsymbol{A}) + \nabla_x \times \boldsymbol{E} = 0 \quad &\Leftrightarrow \quad \nabla_x \times [\boldsymbol{E} + \partial_t \boldsymbol{A}] = 0 \\ &\Leftrightarrow \quad \boldsymbol{E} + \partial_t \boldsymbol{A} = -\nabla_x \varPhi \end{aligned}$$

Note that the negative sign in the last line is simply a convention, and hence

$$\boldsymbol{E} = -\nabla_x \varPhi - \partial_t \boldsymbol{A} \quad ; \quad \boldsymbol{B} = \nabla_x \times \boldsymbol{A} \tag{2.35}$$

It is important to note that these potentials are not uniquely defined in order to produce the same well defined vector field $(\boldsymbol{E},\boldsymbol{B})$. The following lemma explores this freedom.

Lemma 2.5. *Select $(\boldsymbol{E},\boldsymbol{B}) \in C^2(\mathbb{R}^4) \times C^2(\mathbb{R}^4)$ as in (2.35). Let $\boldsymbol{A}'$ and $\varPhi'$ be potentials which determine the same electromagnetic field $(\boldsymbol{E},\boldsymbol{B})$. Then, for some sufficiently smooth function $\lambda : \mathbb{R} \times \mathbb{R}^3 \mapsto \mathbb{R}$, we find*

$$\varPhi' := \varPhi - \partial_t \lambda \quad ; \quad \boldsymbol{A}' := \boldsymbol{A} + \nabla_x \lambda \tag{2.36}$$

Conversely, given any sufficiently smooth λ, we have that $\boldsymbol{A}'$ and $\varPhi'$ defined above will produce the same fields $(\boldsymbol{E},\boldsymbol{B})$.

Proof. Without loss of generality, let $\boldsymbol{\alpha} := \boldsymbol{A}' - \boldsymbol{A}$ and $\beta := \varPhi' - \varPhi$. Then, we have that

$$\nabla_x \times \boldsymbol{A}' = \boldsymbol{B} = \nabla_x \times \boldsymbol{A} = \nabla_x \times [\boldsymbol{A}' - \boldsymbol{\alpha}] \quad \Rightarrow \quad \nabla_x \times \boldsymbol{\alpha} = 0$$

Hence $\boldsymbol{\alpha} = \nabla_x \tilde{\lambda}$ for some scalar function $\tilde{\lambda}$. Similarly, we have that

$$\begin{aligned} -\nabla_x \varPhi' - \partial_t \boldsymbol{A}' = \boldsymbol{E} = -\nabla_x \varPhi - \partial_t \boldsymbol{A} &= -\nabla_x[\varPhi + \beta] - \partial_t[\boldsymbol{A} + \boldsymbol{\alpha}] \\ &\Rightarrow \nabla_x \beta + \partial_t \boldsymbol{\alpha} = 0 \end{aligned}$$

Plugging in $\boldsymbol{\alpha} = \nabla_x \tilde{\lambda}$, we obtain

$$\nabla_x[\beta + \partial_t \tilde{\lambda}] = 0 \quad \Rightarrow \quad \beta + \partial_t \tilde{\lambda} = k(t)$$

Define $\lambda := \tilde{\lambda} - \int_0^t k(t')dt'$. By construction, we have that $\beta = -\partial_t \lambda$ and $\boldsymbol{\alpha} = \nabla_x \tilde{\lambda} = \nabla_x \lambda$, which is the desired result. □

An interesting fact is that the correspondance that is pointed in Lemma 2.5 forms an equivalence relation $(\varPhi,\boldsymbol{A}) \sim (\varPhi',\boldsymbol{A}')$. As a matter of fact, choosing $\lambda = 0$ gives reflexivity; replacing λ by $-\lambda$ gives symmetry; and adding λ_1 and λ_2 according to $\lambda = \lambda_1 + \lambda_2$ gives transitivity.

Definition 2.6. *Define the choice of a Lorenz Gauge to be the selection of some electromagnetic four-potential $(\varPhi',\boldsymbol{A}') \sim (\varPhi,\boldsymbol{A})$ satisfying*

$$G := \partial_t \varPhi' + \nabla_x \cdot \boldsymbol{A}' = 0 \tag{2.37}$$

○

Start with any four-potential $(\varPhi,\boldsymbol{A})$. To show that it is possible to recover the Lorenz gauge for some well chosen $(\varPhi',\boldsymbol{A}')$, note that we can always adjust the scalar function λ in such a way that it is a solution of $\square_{x,t}\lambda = \partial_t \varPhi + \nabla_x \cdot \boldsymbol{A}$. Then

$$\partial_t^2 \lambda - \nabla_x^2 \lambda = \partial_t \varPhi + \nabla_x \cdot \boldsymbol{A} \quad \Leftrightarrow \quad \nabla_x \cdot \boldsymbol{A} + \nabla_x^2 \lambda = -\partial_t \varPhi + \partial_t^2 \lambda \quad \Leftrightarrow \quad \nabla_x \cdot \boldsymbol{A}' = -\partial_t \varPhi'$$

In contrast to Maxwell's equations, the equations on $\boldsymbol{A}$ deduced from (2.22,2.23) are not invariant under Gauge transformation [20]. The following is a nice consequence of the Lorenz Gauge.

Lemma 2.7. *Let* $(\boldsymbol{E}, \boldsymbol{B})$ *be* $C^2(\mathbb{R}^4)$ *fields determined by* $\boldsymbol{A}$ *and* Φ *in the Lorenz Gauge and solving Maxwell's equations. Then*

$$\Box_{t,x}\boldsymbol{A} = -\boldsymbol{J} \tag{2.38}$$

$$\Box_{t,x}\Phi = -Q \tag{2.39}$$

Conversely, given a $C^3(\mathbb{R}^4)\times C^3(\mathbb{R}^4)$ *electromagnetic four-potential* $(\Phi, \boldsymbol{A})$ *satisfying (2.38,2.39) together with the Lorenz Gauge condition (2.37), the electromagnetic field* $(\boldsymbol{E}, \boldsymbol{B})$ *defined by (2.35) is a* $C^2(\mathbb{R}^4)\times C^2(\mathbb{R}^4)$ *solution to Maxwell's equations (2.22,2.23).*

Proof. As already noted, the equations inside (2.23) are the same as (2.35). Knowing (2.35) and (2.37), we have

$$\begin{aligned}
\partial_t \boldsymbol{E} - \nabla_x \times \boldsymbol{B} = \boldsymbol{J} \quad &\Leftrightarrow \quad \partial_t[-\nabla_x \Phi - \partial_t \boldsymbol{A}] - \nabla_x \times (\nabla_x \times \boldsymbol{A}) = \boldsymbol{J} \\
&\Leftrightarrow \quad \nabla_x(\nabla_x \cdot \boldsymbol{A}) - \partial_t^2 \boldsymbol{A} - \nabla_x(\nabla_x \cdot \boldsymbol{A}) + \nabla_x^2 \boldsymbol{A} = \boldsymbol{J} \\
&\Leftrightarrow \quad \Box_{t,x}\boldsymbol{A} = -\boldsymbol{J}
\end{aligned}$$

As well as

$$\begin{aligned}
\nabla_x \cdot \boldsymbol{E} = \nabla_x \cdot [-\nabla_x \Phi - \partial_t \boldsymbol{A}] = -Q \quad &\Leftrightarrow \quad -\nabla_x^2 \Phi - \partial_t[\nabla_x \cdot \boldsymbol{A}] = -Q \\
&\Leftrightarrow \quad -\nabla_x^2 \Phi + \partial_t^2 \Phi = -Q \\
&\Leftrightarrow \quad \Box_{t,x}\Phi = -Q
\end{aligned}$$

Since all above lines are equivalences, we get the result. □

Keep in mind that (2.38,2.39) together with (2.37) is an overdetermined system. Indeed, this implies the compatibility condition

$$\partial_t Q + \nabla_x \cdot \boldsymbol{J} = 0 \tag{2.40}$$

which is actually the mass continuity equation in the case of (2.21).

2.2.2. *Lienard-Wiechert Potentials.* We now wish to write the fields of the MRVM system in terms of solutions to a wave equation. Let $u(t, x, \xi)$ be the scalar function (sometimes called the microscopic electromagnetic potential) which is a solution to the Cauchy problem built with

$$\Box_{t,x} u = -f \tag{2.41}$$

together with

$$u(0, x) = 0, \quad \partial_t u(t, x)|_{t=0} = 0 \tag{2.42}$$

With $\boldsymbol{E}^{in}$ as in (2.48) and $\boldsymbol{A}^{in}$ as in (2.34), let $\boldsymbol{A}^0(t, x)$ be the vector-valued function satisfying

$$\Box_{t,x}\boldsymbol{A}^0 = 0 \quad ; \quad \boldsymbol{A}^0_{|t=0} = \boldsymbol{A}^{in} \quad ; \quad \partial_t \boldsymbol{A}^0_{|t=0} = -\boldsymbol{E}^{in} \tag{2.43}$$

The Lienard-Wiechert potentials are correspondingly defined as

$$\Phi := \int u d\xi \quad ; \quad \boldsymbol{A} := \boldsymbol{A}^0 + \int u\nu(\varepsilon\xi) d\xi \tag{2.44}$$

In view of (2.35), this means that

$$\boldsymbol{E} = -\partial_t \boldsymbol{A}^0 - \int [\nu(\varepsilon\xi)\partial_t + \nabla_x] u d\xi \tag{2.45}$$

$$\boldsymbol{B} = \nabla_x \times \boldsymbol{A}^0 + \int \nabla_x \times [u\nu(\varepsilon\xi)] d\xi \tag{2.46}$$

The potential $\boldsymbol{A}^0$ is a fixed function determined by (2.41), independently of $\boldsymbol{U}$. The introduction of $\boldsymbol{A}^0$ allows to absorb the initial data $\boldsymbol{E}^{in}$ and $\boldsymbol{B}^{in}$. It induces a shift on the electromagnetic field, as indicated in (2.45,2.46). At the level of f, it generates a transport in the phase space.

But it is not involved in a coupling between f, $\boldsymbol{E}$ and $\boldsymbol{B}$. To avoid technicalities, from now on, we assume that $\boldsymbol{A}^0 \equiv 0$, or equivalently that

$$f_{|t=0} = f^{in} \not\equiv 0 \tag{2.47}$$

$$\boldsymbol{E}_{|t=0} = \boldsymbol{E}^{in} \equiv 0 \tag{2.48}$$

$$\boldsymbol{B}_{|t=0} = \boldsymbol{B}^{in} \equiv 0 \tag{2.49}$$

The condition of compatibility becomes

$$0 = -\rho(f^{in}) = -\int f^{in}(x,\xi) d\xi \quad ; \quad \nabla_x \cdot \boldsymbol{B}^{in} = 0 \tag{2.50}$$

Lemma 2.8. *The MRVM system (2.21,2.22,2.23) together with the coupling source terms of (2.24,2.25) and the initial data (2.47,2.48,2.49) is equivalent to the Vlasov-Wave system (2.21,2.41) closed by the relations (2.45,2.46) and the initial conditions (2.47,2.42).*

Proof. First, consider the initial data. The condition (2.47) is unchanged. On the other hand, the conditions (2.48,2.49) are a direct consequence of (2.42) together with (2.45,2.46).

From (2.41), with $\varPhi$ and $\boldsymbol{A}$ given by (2.44), we can easily deduce

$$\Box_{t,x} \varPhi := \int \Box_{t,x} u d\xi = -\int f d\xi \tag{2.51}$$

$$\Box_{t,x} \boldsymbol{A} := \int \Box_{t,x} u \nu(\varepsilon\xi) d\xi = -\int f \nu(\varepsilon\xi) d\xi \tag{2.52}$$

where, in the right hand side, we can recognize the operators Q and $\boldsymbol{J}$ of (2.24) and (2.25). Thus, we have (2.38,2.39) with the adequate definition of Q and $\boldsymbol{J}$. Now, applying Lemma 2.7, it suffices to check that the Lorenz Gauge condition (2.37) is indeed satisfied. We find

$$G(t,x) = \int \big[\partial_t u + \nabla_x \cdot \big(u \ \nu(\varepsilon\, \xi) \big) \big] \ d\xi \tag{2.53}$$

Exploiting (2.41), compute

$$\Box_{t,x} \, G = -\int \big\{ \partial_t f + \nu(\varepsilon\, \xi) \cdot \nabla_x f \big\} \ d\xi \tag{2.54}$$

According to (2.21), the above total derivative $\partial_t f + \nu(\varepsilon\,\xi) \cdot \nabla_x f$ can be replaced by

$$\begin{aligned} \partial_t f + \big[\nu(\varepsilon\,\xi) \cdot \nabla_x \big] f = {} & \mathrm{div}_\xi \left[\langle \varepsilon\, \xi \rangle^{-1} \, f \ \xi \wedge \mathbf{B}_e(x) \right] \\ & + M'(|\xi|) \, \frac{\xi \cdot E}{|\xi|} + \mathrm{div}_\xi \left[E + f \ \nu(\varepsilon\, \xi) \wedge B \right] \end{aligned} \tag{2.55}$$

After integration in ξ as required by (2.54), all terms implying div_ξ disappear. Besides, the term with $M'(\cdot)$ in factor does not contribute because it involves the integral of an odd function (in the variable ξ). There remains $\Box_{t,x} G = 0$. This is not sufficient to guarantee that $G \equiv 0$. Look at the initial data. It is clear that $G_{|t=0} \equiv 0$. On the other hand, we have

$$(\partial_t \, G)_{|t=0} = \int (\partial^2_{tt} u)_{|t=0} \ d\xi = -\int f^{in}(x,\xi) \ d\xi \, . \tag{2.56}$$

This is where the neutrality condition (2.27) plays a crucial role. It is necessary to guarantee that $\partial_t \, G_{|t=0} \equiv 0$, which in turn furnishes $G(t,\cdot) \equiv 0$ for all times $t \in \mathbb{R}_+$, that is (2.37). In other words, the constraint (2.27) appears as a compatibility condition allowing to solve the overdetermined system (2.37,2.38,2.39). $\square$

The system (2.21,2.41) with (2.45,2.46) is self-contained. This will be our starting point.

3. Preparatory work

This section collects identities that will be needed in the sequel. In Subsection 3.1, we remark that the solution $u(\cdot)$ of (2.41,2.42) can be determined through a convolution procedure implying some homogeneous distribution; we generalize such formulas, and we derive related estimates. In Subsection 3.2, we introduce commuting methods implying vector fields. In Subsection 3.3, we explain the content of a somewhat classical *division lemma*, already exploited in [4, 27]; this division lemma allows to replace (modulo error terms) the derivatives involved inside (2.45,2.46) by the total derivative $\partial_t + \nu(\varepsilon\,\xi)\cdot\nabla_x$ of (2.21). The final Subsection 3.4 is more original; it is specific to the present framework; it explains how to further convert, always in the context of (2.45,2.46) and again modulo error terms, the derivative $\partial_t + \nu(\varepsilon\,\xi)\cdot\nabla_x$ into a nonsingular derivative; this requires to deal with the penalized term that is implied at the level of (2.21); this means to extract uniform estimates (in ε) from the term which inside has ε^{-2} in factor.

3.1. **Convolution estimates.** The fundamental solution Y associated with $\Box Y = \delta(t,x)$ is

$$Y := \frac{1}{4\pi t}\mathbb{1}_{t>0}\delta(|x|-t) \tag{3.1}$$

Consequently, the solution $u(\cdot)$ of (2.41,2.42) is given by

$$u(t,x,\xi) = -Y * (f\,\mathbb{1}_{t>0}) \tag{3.2}$$

In (3.2), the symbol $*$ means a convolution with respect to the variables t and x (but not with respect to the variable ξ which can be forgotten here). More generally, we will have to consider expressions like

$$\mathrm{u}(t,x) = (pY) * (f\,\mathbb{1}_{t>0}) \tag{3.3}$$

where $p \in \mathcal{M}_m$, the space of $\mathcal{C}^\infty$ homogeneous functions on $\mathbb{R}^4 \setminus \{0\}$ of degree $m \in \mathbb{R}$. In other words, given $p \in \mathcal{M}_m$, we have

$$\forall \lambda \in \mathbb{R}_+^*, \quad \forall (t,x) \in \mathbb{R}^4 \setminus \{0\}, \quad p(\lambda t, \lambda x) = \lambda^m p(t,x) \tag{3.4}$$

We have $\mathcal{M}_m \subset M_m$, where M_m is the space of homogeneous distributions with domain $\mathbb{R}^4 \setminus \{0\}$, having degree m. For instance, we have $Y \in M_{-2}$. In Paragraph 3.1.1, we study (3.3) when $m \geq -1$. Then, in Paragraph 3.1.2, we investigate (3.3) in the critical case $m = -2$.

3.1.1. *Convolution estimates: the easy case.* This is when (3.3) is given by as a classical integral.

Lemma 3.1. *Let $p \in \mathcal{M}_m$ with $m \geq -1$. Select $f \in L^\infty(\mathbb{R}^4)$. The expression $\mathrm{u}(\cdot)$ given by (3.3) is well-defined as a usual integral with parameters. Moreover, we have*

$$|\mathrm{u}(t,x)| \leq \frac{t^{1+m}}{3} \parallel p(1,\cdot) \parallel_{L^\infty(\mathbb{S}^2)} \int_0^t \parallel f(s,\cdot) \parallel_{L^\infty(\mathbb{R}^3_x)} ds \tag{3.5}$$

where $\mathbb{S}^2$ is the unit sphere of $\mathbb{R}^3$.

We can apply (3.5) with $p \equiv 1$ and $m = 0$ to obtain

$$|u(t,x,\xi)| \leq \frac{t}{3} \int_0^t \parallel f(s,\cdot,\xi) \parallel_{L^\infty(\mathbb{R}^3_x)} ds \tag{3.6}$$

Proof. As explained in [18] (see Proposition 3.6.12), the homogeneous distributions $pY \in M_\beta$ with $\beta = m-2 > -4$ has a unique homogeneous extension in $\mathcal{D}'(\mathbb{R}^4)$. Thus, it can be applied to smooth test functions f. Now, another way to interpret (3.3) and to extend (3.3) in the case of more general functions f is to write $\mathrm{u}(\cdot)$ as an integral, and then to observe that the support of Y is the light cone

$$\sup Y \equiv \mathscr{LC} := \{|x| = t\} \subset \mathbb{R}^4 \tag{3.7}$$

With this in mind, we have (formally)

$$\begin{aligned}\mathrm{u}(t,x) &= \int_{\mathbb{R}^4} \frac{p(s,y)}{4\pi s} \mathbb{1}_{s>0} \delta(|y|-s) f(t-s,x-y) \mathbb{1}_{t-s>0}\, dsdy \\ &= \int_0^t \int_0^{+\infty} \int_{\mathbb{S}^2} \frac{p(s,r\omega)}{4\pi s} \delta(r-s) f(t-s,x-r\omega) r^2 dsdrd\sigma \\ &= \int_0^t \int_{\mathbb{S}^2} \frac{p(s,s\omega)}{4\pi} f(t-s,x-s\omega) sdsd\sigma \end{aligned} \tag{3.8}$$

where

$$\omega := \frac{x}{|x|} = \begin{bmatrix} \sin\phi\cos\theta \\ \sin\phi\sin\theta \\ \cos\phi \end{bmatrix} \in \mathbb{S}^2 := \{|x|=1\} \tag{3.9}$$

and where $d\sigma$ is the rotation-invariant surface element on $\mathbb{S}^2$. Because $p \in \mathcal{M}_m$ with $m \geq -1$, this can be viewed as the following (convergent) integral

$$\mathrm{u}(t,x) = \int_0^t \int_{\mathbb{S}^2} \frac{p(1,\omega)}{4\pi} f(t-s,x-s\omega)\, s^{1+m} dsd\sigma \tag{3.10}$$

from which we can easily deduce (3.5). $\square$

In view of the above proof, Lemma 3.1 can be improved in two directions. First, the result (3.5) does not change if p is multiplied by a smooth bounded function. Secondly, to obtain (3.5), it suffices to know that $p(\cdot)$ is smooth and well defined in a conic neighborhood $\mathcal{V}$ of $\{1\} \times \mathbb{S}^2$, where $\mathcal{V}$ is viewed as a subset of $(\mathbb{R} \times \mathbb{R}^3) \setminus \{0\}$.

3.1.2. *Convolution estimates: the critical case.* The case $p \in \mathcal{M}_{-2}$ is more difficult because all expressions $pY \in M_{-4}$ are not the restriction of some homogeneous element inside $\mathcal{D}'(\mathbb{R}^4)$.

Lemma 3.2. *Let $p \in \mathcal{M}_{-2}$. The distribution $pY \in M_{-4}$ can be extended as a homogeneous distribution on the whole time-space $\mathbb{R}^4$ if and only if*

$$\int_{\mathbb{S}^2} p(1,\omega) d\sigma = 0 \tag{3.11}$$

Now, assume (3.11). Then, given $f \in L^\infty(\mathbb{R}^4)$, the expression $\mathrm{u}(\cdot)$ of (3.3) is well-defined as a usual integral with parameters, and we have

$$|\mathrm{u}(t,x)| \leq \frac{t}{4\pi} \Big(\int_{\mathbb{S}^2} |p(1,\omega)| d\sigma \Big) \parallel \nabla_{t,x} f \parallel_{L^\infty} \tag{3.12}$$

Proof. Since $pY \in M_{-4}$, we have that

$$^t(t,x) pY \in M_{-3} \quad ; \quad \mathrm{div}_{t,x}\big(^t(t,x)(pY)\big) \in M_{-4} \tag{3.13}$$

Because $^t(t,x)p \in \mathcal{M}_{-1}$, from Proposition 3.6.12 of [18], we can assert that $^t(t,x)pY$ has a unique homogeneous extension in $\mathcal{D}'(\mathbb{R}^4)$. Moreover, from Euler relation, we know that

$$\mathrm{div}_{t,x}\big(^t(t,x)(pY)\big) \equiv 0 \text{ as an element of } M_{-4} \tag{3.14}$$

This implies that

$$\exists c \in \mathbb{R}; \quad \mathrm{div}_{t,x}\big(^t(t,x)(pY)\big) = c\delta \ \text{ in } \ \mathcal{D}'(\mathbb{R}^4) \tag{3.15}$$

where the constant c is called the residue of pY. As is well-known (Proposition 4.1.8 in [18]), the element $pY \in M_{-4}$ can be extended as a distribution in $\mathcal{D}'(\mathbb{R}^4)$ if and only if $c = 0$. Now, select a smooth function $\varphi(t,x)$ of the form $\varphi(t,x) = -\phi\big(|(t,x)|^2\big)$, where $\phi(\cdot) \in \mathcal{C}_c^\infty(\mathbb{R}_+)$ is such that

$$\phi(0) \neq 0, \quad \phi'(0) = 0, \quad \int_0^{+\infty} \frac{\phi'(t)}{t} dt \neq 0 \tag{3.16}$$

We can test (3.15) in the case of this special choice of $\varphi(\cdot)$ to obtain

$$\begin{aligned}
c\phi(0) &= \langle pY, 2|(t,x)|^2\phi'\big(|(t,x)|^2\big)\rangle \\
&= \int_0^\infty \int_{\mathbb{S}^2} p\Big(t, t\frac{x}{|x|}\Big)\frac{4t^2}{4\pi t}\phi'(2t^2)dtd\sigma \\
&= \int_0^\infty \frac{\phi'(2t^2)}{\pi t}dt \int_{\mathbb{S}^2} p(1,\omega)d\sigma = \frac{1}{2\pi}\int_0^\infty \frac{\phi'(t)}{t}dt\int_{\mathbb{S}^2} p(1,\omega)d\sigma
\end{aligned}$$

In view of (3.16), the condition $c=0$ is equivalent to (3.11). This furnishes the first part of Lemma 3.2. On the other hand, exploiting (3.11), we find

$$\mathrm{u}(t,x) = \frac{1}{4\pi}\int_0^t \int_{\mathbb{S}^2} p(1,\omega)\frac{f(t-s,x-sy)-f(t,x)}{s}dsd\sigma \tag{3.17}$$

which gives rise to (3.12). $\square$

3.2. **Commuting vector fields.** In this subsection, we exhibit vector fields that commute with the wave operator $\square_{t,x}$. With $v := \nu(\varepsilon\xi)$, define

$$T(v) := \partial_t + v\cdot\nabla_x \tag{3.18}$$

$$L_i := x_i\partial_t + t\partial_i, \quad i=1,2,3 \tag{3.19}$$

The existence of commuting vector fields associated with the operator $\square_{t,x}$ is well known property, see [21] or the survey article [22]. In particular, we have the following.

Lemma 3.3. *We have $[L_i,\square]=0$, $L_i\delta(t,x)=0$ and $L_iY=0$ as distributions.*

Proof. For the sake of completeness, we recall the proof. First

$$\begin{aligned}
\square L_i &= (\partial_t^2 - \Delta)(x_i\partial_t + t\partial_i) \\
&= [x_i\partial_t^3] + [2\partial_{ti}^2 + t\partial_{itt}^3] - [x_i(\partial_j^2+\partial_k^2)\partial_t + \partial_i^2(x_i\partial_t)] - t\Delta\partial_i \\
&= [x_i\partial_t^3] + [2\partial_{ti}^2 + t\partial_{itt}^3] - [x_i\Delta\partial_t + 2\partial_{it}^2] - [t\Delta\partial_i] \\
&= x_i\partial_t^3 + t\partial_{itt}^3 - x_i\Delta\partial_t - t\Delta\partial_i \\
&= (x_i\partial_t + t\partial_i)(\partial_t^2-\Delta) = L_i\square
\end{aligned}$$

Secondly, given $\phi(t,x)\in C_c^\infty(\mathbb{R}^4)$, we have

$$\begin{aligned}
\langle L_i\delta,\phi\rangle &= \int ([x_i\partial_t + t\partial_j]\delta(t,x))\phi(t,x)dxdt \\
&= -\int \delta(t,x)[x_i\partial_t+t\partial_i]\phi(t,x)dxdt \\
&= -(x_i\partial_t\phi(x,t) + t\partial_i\phi(x,t))|_{(x,t)=(0,0)} = 0
\end{aligned}$$

It follows that

$$0 = [\square, L_i]Y = \square L_iY - L_i\square Y = \square L_i Y - L_i\delta(t,x) = \square L_iY$$

We then have

$$\begin{aligned}
\langle L_iY,\phi\rangle = \langle (t\partial_i + x_i\partial_t)Y,\phi)\rangle = -\langle Y, t\partial_i\phi + x_i\partial_t\phi\rangle = -\left\langle \frac{\mathbb{1}_{t>0}\delta(|x|-t)}{4\pi}, \partial_i\phi + \frac{x_i}{t}\partial_t\phi\right\rangle \\
= -\int_{\mathbb{R}^3}[\partial_i\phi(x,|x|) + \frac{x_i}{|x|}\partial_t\phi(x,|x|)]dx = -\int_{\mathbb{R}^3}\partial_i[\phi(x,|x|)]dx = 0
\end{aligned}$$

This is the third relation. $\square$

3.3. **First transfer of derivatives.** The following Lemma is stated and proved in [4]. In order to introduce tools that will be useful, we repeat it in below with full details. Set $\partial_0 \equiv \partial_t$.

Lemma 3.4. *(Division Lemma) For any $v \in \mathbb{R}^3$, with $|v| < 1$ we have the following.*

(1) There exists $a_i^k(t,x) \in \mathcal{M}_{-k}$, with $i \in \{0,1,2,3\}$ and $k \in \{0,1\}$ such that

$$\partial_i Y = T(a_i^0 Y) + a_i^1 Y, \quad \forall i = 0,1,2,3 \tag{3.20}$$

(2) There exists $b_{ij}^k(t,x) \in \mathcal{M}_{-k}$ with $i,j \in \{0,1,2,3\}$ and $k \in \{0,1\}$ such that

$$\partial_{i,j}^2 Y = T^2(b_{ij}^0 Y) + T(b_{ij}^1 Y) + b_{ij}^2 Y \quad \forall i,j = 0,1,2,3 \tag{3.21}$$

(3) Moreover

$$\int_{\mathbb{S}^2} b_{ij}^2(1,\omega) d\sigma = 0 \tag{3.22}$$

Proof. Observe that

$$\begin{aligned}
\sum_{j=1}^3 v_j L_j &= \sum_{j=1}^3 x_j v_j \partial_t + v_j t \partial_j = x \cdot v \partial_t + tv \cdot \nabla_x \\
&= t(\partial_t + v \cdot \nabla_x) - t\partial_t + x \cdot v \partial_t \\
&= tT(v) + (x \cdot v - t)\partial_t = x \cdot v \partial_t + tv \cdot \nabla_x
\end{aligned} \tag{3.23}$$

Using (3.23), we get

$$\begin{aligned}
(t - x\cdot v)L_i + x_i \sum_{j=1}^3 v_j L_j &= (t - x \cdot v)(x_i \partial_t + t\partial_i) + x_i(x \cdot v\partial_t + tv \cdot \nabla_x) \\
&= tx_i\partial_t + t^2\partial_i - x_i x \cdot v \partial_t - tx \cdot v \partial_i + x_i x \cdot v \partial_t + tx_i v \cdot \nabla_x \\
&= t[(t - x \cdot v)\partial_i + x_i(\partial_t + v \cdot \nabla_x)] \\
&= t[(t - x \cdot v)\partial_i + x_i T(v)]
\end{aligned}$$

From $L_j Y = 0$ (Lemma 3.3) and (3.23), we have

$$(\sum_{j=1}^3 v_j L_j) Y = 0 = tT(v)Y + (x \cdot v - t)\partial_t Y$$

$$[(t - x \cdot v)L_i + x_i \sum_{j=1}^3 v_j L_j] Y = 0 = t[(t - x \cdot v)\partial_i + x_i T(v)]Y \tag{3.24}$$

We then define for $x \cdot v \neq t$ and $i \in \{1,2,3\}$

$$a_0(t,x) := \frac{t}{t - x \cdot v}, \quad a_i(t,x) := \frac{x_i}{x \cdot v - t} \tag{3.25}$$

Away from $x \cdot v = t \neq 0$, from (3.24), we can deduce that

$$\partial_i Y = a_i T(v) Y, \quad i = 1,2,3 \tag{3.26}$$

Let $x_0 = t$. Since $L_i Y = 0$, we have

$$-x \cdot v \partial_t Y = -v_i(L_i Y - t\partial_i Y) = tv \cdot \nabla_x Y \tag{3.27}$$

Adding $t\partial_t Y$ to (3.27), we obtain (3.26) for $i = 0$, that is

$$\partial_0 Y = a_0 T(v) Y \tag{3.28}$$

Looking at the definition (3.1), we have

$$\sup Y \subset \{(t,x) | 0 \leq |x| = t\} \tag{3.29}$$

Combining (3.26) and (3.28) (away from $x \cdot v = t \neq 0$) as well as (3.29), we can deduce that

$$\sup(\partial_i Y - a_i TY) \subset \{(t,x)|x \cdot v = t\} \cup \{(t,x)|t = 0\} \cap \{(t,x)|0 \leq |x| = t\} \tag{3.30}$$

But, for $0 < |x| = t$, since $|v| = |\nu(\xi)| < 1$, we have $t - x \cdot v \neq 0$. It follows that

$$\sup(\partial_i Y - a_i TY) \subset \{(0,0)\} \tag{3.31}$$

We have $a_i \in M_0$ so that $\partial_i Y - a_i TY \in M_{-3}$. As already seen, such homogeneous distributions on $\mathbb{R}^4 \setminus \{0\}$ of degree $\beta > -4$ have a unique homogeneous extension on $\mathbb{R}^4$. In view of (3.31), this means that

$$\partial_i Y - a_i TY = 0 \ \text{ in } \ \mathcal{D}'(\mathbb{R}^4), \quad i = 0,1,2,3 \tag{3.32}$$

Fix v with $|v| < 1$, and construct a function $\chi \equiv \chi_v \in C_c^\infty(\mathbb{R}^+)$ (depending on v) such that

$$0 \leq \chi \leq 1, \quad \chi_{[0,\frac{|v|+1}{2|v|}]} = 1, \quad \sup \chi \subset [0, \frac{1}{|v|}) \tag{3.33}$$

Introduce the auxilliary functions

$$a_i^0(t,x) := a_i(t,x)\chi\Big(\frac{|x|}{t}\Big) \in \mathcal{M}_0 \tag{3.34}$$

$$a_i^1(t,x) := -T(a_i^0) \in \mathcal{M}_{-1} \tag{3.35}$$

By construction, we have $\chi \equiv 1$ on a neighborhood of 1, and therefore $a_i^0 \equiv a_i$ on a neighborhood of $\sup Y$. We have $\partial_i Y = a_i TY = a_i^0 TY = T(a_i^0 Y) - T(a_i^0)Y$, and hence

$$\partial_i Y = T(a_i^0 Y) + a_i^1 Y, \quad i = 0,1,2,3 \tag{3.36}$$

This proves the subparagraph 1 of Lemma 3.4.

Now, let $m^k(t,x) \in \mathcal{M}_{-k}$, for $k \in \{0,1\}$. Then

$$\begin{aligned}
\partial_i(m^k Y) &= m^k \partial_i Y + Y \partial_i m^k \\
&= m^k[T(a_i^0 Y) - YT(a_i^0)] + Y\partial_i m^k \\
&= m^k T(a_i^0 Y) + a_i^0 YT(m^k) - a_i^0 YT(m^k) - m^k YT(a_i^0) + Y\partial_i m^k \\
&= T(m^k a_i^0 Y) - T(m^k a_i^0)Y + Y\partial_i m^k \\
&= T(m^k a_i^0 Y) + [\partial_i m^k - T(m^k a_i^0)]Y \qquad (3.37)
\end{aligned}$$

Coming back to (3.20), we have that

$$\partial_{ij}^2 Y = T\big(\partial_i(a_j^0 Y)\big) + \partial_i(a_j^1 Y) \tag{3.38}$$

(3.39)

Then, applying (3.37) with $m = a_j^0$ and $m = a_j^1$, we can obtain

$$\begin{aligned}
\partial_{ij}^2 Y &= \big\{T(T([a_j^0 a_i^0]Y) + [\partial_i a_j^0 - T(a_j^0 a_i^0)]Y)\big\} + \big\{T(a_j^1 a_i^0 Y) + [\partial_i a_j^1 - T(a_j^1 a_i^0)]Y\big\} \qquad (3.40)\\
&= T^2(a_j^0 a_i^0 Y) + T([\partial_i a_j^0 - T(a_j^0 a_i^0) + a_j^1 a_i^0]Y) + [\partial_i a_j^1 - T(a_j^1 a_i^0)]Y \qquad (3.41)
\end{aligned}$$

where we read off

$$b_{ij}^0 = a_j^0 a_i^0 \tag{3.42}$$

$$b_{ij}^1 = \partial_i a_j^0 - T(a_j^0 a_i^0) + a_j^1 a_i^0 \tag{3.43}$$

$$b_{ij}^2 = \partial_i a_j^1 - T(a_j^1 a_i^0) \tag{3.44}$$

This proves the subparagraph 2 of Lemma 3.4.

From (3.37), we have

$$M_{-4} \ni [\partial_i m^1 - T(m^1 a_i^0)]Y = \partial_i(m^1 Y) - T(m^1 a_i^0 Y) \tag{3.45}$$

Both m^1Y and $m^1a_i^0Y$ are in M_{-3}, and thus they have a unique homogenous extension to $\mathbb{R}^4$. It follows that the right hand side is well defined as some homogenous distribution in $\mathcal{D}'(\mathbb{R}^4)$, of degree -4. The same must apply to the left hand side. Now, it suffices to apply Lemma 3.2 with $p = [\partial_i m^1 - T(m^1a_i^0)] \in \mathcal{M}_{-2}$ to obtain

$$\int_{\mathbb{S}^2} \big(\partial_i m^1 - T(m^1a_i^0)\big)(1,\omega)d\sigma \tag{3.46}$$

This holds for any $m^1 \in \mathcal{M}_{-1}$. In particular, for $m^1 = a_j^1 = -T(a_j^0)$. This yields (3.22), proving the subparagraph 3 of Lemma 3.4. □

When dealing with L^∞-bounds extracted from (2.45,2.46), a key argument is to replace the derivatives $\nu(\varepsilon\,\xi)\,\partial_t + \nabla_x$ and ∇_x by the derivative $\partial_t + \nu(\varepsilon\,\xi)\cdot\nabla_x$ of (2.21). This can work because $|\nu(\xi)| < 1$, which implies that these two derivatives are transverse to the light cone $\mathscr{LC}$. This possibility of exchanging these derivatives can be viewed as a consequence of the preceding *division lemma* of [4, 27]. Define

$$p(t,x,\xi) := \frac{\nu(\xi)t - x}{\nu(\xi)\cdot x - t} \quad ; \quad q(t,x,\xi) := \frac{1}{\langle\xi\rangle^2}\frac{\nu(\xi)t - x}{\big[\nu(\xi)\cdot x - t\big]^2} \tag{3.47}$$

Remark that these two functions $p(\cdot)$ and $q(\cdot)$ are not defined on the whole time-space $\mathbb{R}^4$ but they are well defined away from $t = |x| = 0$, that is on a neighborhood of $\mathscr{LC}$.

Corollary 3.5. *[First transfer of derivatives] For all $\xi \in \mathbb{R}^3$, we have*

$$\big[\nu(\xi)\,\partial_t + \nabla_x\big]Y = -T(\xi)\,\big[p(t,x,\xi)Y\big] + q_{|\mathscr{LC}}(x,\xi)Y \tag{3.48}$$

Proof. We can define

$$a^0 := {}^t(a_1^0, a_2^0, a_3^0), \quad a^1 := {}^t(a_1^1, a_2^1, a_3^1) \tag{3.49}$$

Fix $\xi \in \mathbb{R}^3$. Then, with $v = \nu(\xi)$ and $\chi \equiv \chi_v \equiv \chi_{\nu(\xi)}$, we can consider

$$p^0(x,t,\xi) := -[va_0^0 + a^0] = \frac{\nu(\xi)t - x}{\nu(\xi)\cdot x - t}\chi\Big(\frac{|x|}{t}\Big), \quad p^0(\cdot,\xi) \in \mathcal{M}_0 \tag{3.50}$$

$$q^0(x,t,\xi) := Tp^0, \qquad q^0(\cdot,\xi) \in \mathcal{M}_{-1} \tag{3.51}$$

Using (3.36), this furnishes

$$[\nu(\xi)\partial_t + \nabla_x]Y = -T(p^0Y) + q^0Y \tag{3.52}$$

Because $\chi \equiv 1$ in a neighborhood of 1, on a suitable neighborhood of $\sup Y \equiv \mathscr{LC}$, we have $p^0 \equiv p$ and $q^0 \equiv q$, so that $p^0Y \equiv pY$ and $q^0Y \equiv qY$. Since the computation of qY involves a Dirac mass without implying derivatives, as indicated in (3.48), we have $qY \equiv q_{|\mathscr{LC}}Y$, with

$$p_{|\mathscr{LC}}(x,\xi) \equiv \frac{\nu(\xi)|x| - x}{\nu(\xi)\cdot x - |x|}, \qquad p(\cdot,\xi)_{|\mathscr{LC}} \in \mathcal{M}_0 \tag{3.53}$$

$$q_{|\mathscr{LC}}(x,\xi) \equiv \frac{1}{\langle\xi\rangle^2}\frac{\nu(\xi)|x| - x}{[\nu(\xi)\cdot x - |x|]^2}, \quad q(\cdot,\xi)_{|\mathscr{LC}} \in \mathcal{M}_{-1} \tag{3.54}$$

Since $\sup Y \equiv \mathscr{LC}$ intersects $\big\{\nu(\xi)\cdot x - t = 0\big\}$ only at the origin of $\mathbb{R}^4$, the two distributions pY and qY are respectively in M_{-2} and M_{-3}. Thus, they can be extended uniquely as elements of $\mathcal{D}'(\mathbb{R}^4)$. Now, the relation (3.52) with p^0 and q^0 replaced by p and q remains valid in the sense of $\mathcal{D}'(\mathbb{R}^4)$. This is exactly (3.48). □

In view of (2.45, 2.46), the direction ξ is aimed to be replaced by $\varepsilon\xi$. With this in mind, define

$$T_\varepsilon(\xi) := \partial_t + \nu(\varepsilon\xi)\cdot\nabla_x \tag{3.55}$$

as well as

$$p_\varepsilon(t,x,\xi) := p(t,x,\varepsilon\xi) \quad ; \quad q_\varepsilon(t,x,\xi) := q_{|\mathscr{L}\mathscr{C}}(x,\varepsilon\xi) \tag{3.56}$$

$$a^0_\varepsilon(t,x,\xi) := a^0(t,x,\varepsilon\xi) \quad ; \quad a^1_\varepsilon(t,x,\xi) := a^1(t,x,\varepsilon\xi) \tag{3.57}$$

Applying Corollary 3.5 with the parameter ξ replaced by $\varepsilon\xi$, the distribution

$$D \equiv D(\varepsilon,t,x,\xi) := \nu(\varepsilon\xi)\partial_t Y + \nabla_x Y \tag{3.58}$$

is transformed into

$$D = -T_\varepsilon(p_\varepsilon Y) + q_\varepsilon Y \tag{3.59}$$

Coming back to (2.45) with $\boldsymbol{A}^0 \equiv 0$ and using (3.2), it follows that

$$\boldsymbol{E} = -\int (p_\varepsilon Y) * T_\varepsilon(f\mathbb{1}_{t>0})d\xi + \int (q_\varepsilon Y) * (f\mathbb{1}_{t>0})d\xi \tag{3.60}$$

In the same way, exploiting (3.20), we find that

$$\boldsymbol{B} = -\int (a^0_\varepsilon Y) * T_\varepsilon(f\mathbb{1}_{t>0}) \times \nu(\varepsilon\xi)d\xi - \int (a^1_\varepsilon Y) * (f\mathbb{1}_{t>0}) \times \nu(\varepsilon\xi)d\xi \tag{3.61}$$

The right hand sides of both (3.60) and (3.61) involve only one differential action, namely $T_\varepsilon f$. Using (2.21), this becomes

$$T_\varepsilon f = M'(|\xi|)|\xi|^{-1}\xi \cdot \boldsymbol{E} + \nabla_\xi h \tag{3.62}$$

where

$$h := \frac{1}{\varepsilon^2}\big[\nu(\varepsilon\xi) \times \mathbf{B}_e(x)\big]f + \big[\boldsymbol{E} + \nu(\varepsilon\xi) \times \boldsymbol{B})\big]f \tag{3.63}$$

An integration by parts allows to shift the derivative ∇_ξ to the weights $p_\varepsilon(\cdot)$ or $a^0_\varepsilon(\cdot)$. This transfer is the key to L^∞-bounds because it removes one derivative from $f(\cdot)$. It also produces a gain of a small factor ε. By way of illustration, we consider below the case of $p_\varepsilon(\cdot)$.

Lemma 3.6. *[gain of a derivative and of a small factor ε]*

$$\Big|\int (p_\varepsilon Y) * \nabla_\xi(h\mathbb{1}_{t>0})d\xi\Big| \leq \frac{\varepsilon t}{3}\int_0^t \int \parallel \nabla_\xi p(1,\cdot,\varepsilon\xi) \parallel_{L^\infty(\mathbb{S}^2)} \parallel h(s,\cdot,\xi) \parallel_{L^\infty_x(\mathbb{R}^3)} ds d\xi \tag{3.64}$$

Proof. Just remark that

$$\int (p_\varepsilon Y) * \nabla_\xi(h\mathbb{1}_{t>0})d\xi = -\varepsilon \int \big(\nabla_\xi p(\cdot,\varepsilon\xi)Y\big) * (h\mathbb{1}_{t>0})d\xi$$

Like $p(\cdot,\varepsilon\xi)$, the function $\nabla_\xi p(\cdot,\varepsilon\xi)$ is in $\mathcal{M}_0$ near the cone $\mathscr{L}\mathscr{C}$. To recover (3.64), it suffices to apply Lemma 3.1. □

The factor ε appearing in (3.64) turns out to be crucial in several places. It can potentially absorb the singular factor ε^{-1} involved by $h(\cdot)$. Another way to proceed, which ultimately amounts to the same thing but which would appear more intrinsic, is to filter the Vlasov equation. This particular method is selected in the next subsection.

3.4. **Second transfer of derivatives.** Introduce the *approximated flow* that is the flow which is associated with the left part of the transport equation (2.21). Define

$$\dot{\mathrm{X}} = \nu(\varepsilon\,\Xi)\,, \qquad \mathrm{X}(0) = x \tag{3.65}$$

$$\dot{\Xi} = -\varepsilon^{-2}\nu(\varepsilon\,\Xi) \wedge \mathbf{B}_e(\mathrm{X})\,, \quad \Xi(0) = \xi \tag{3.66}$$

The functions $\mathrm{X}(\cdot)$ and $\Xi(\cdot)$ depend on the parameter $\varepsilon \in]0,1]$, on the time $t \in \mathbb{R}_+$, on the initial position $x \in \mathbb{R}^3$, and on the initial velocity $\xi \in \mathbb{R}^3$. They can be denoted by $\mathrm{X}(\varepsilon;t,x,\xi)$

and $\Xi(\varepsilon;t,x,\xi)$. Sometimes, as in the case of $f(\cdot)$, the dependence on ε will not be marked. Also, in many occasions, we will simply use $\mathrm{X}(t,\cdot)$ and $\Xi(t,\cdot)$. The flow

$$\mathrm{F} : (s,y,\eta) \longmapsto (t,x,\xi) := \big(s, \mathrm{F}_s(y,\eta)\big) \quad ; \quad \mathrm{F}_s(y,\eta) := \big(\mathrm{X}(s,y,\eta), \Xi(s,y,\eta)\big) \tag{3.67}$$

is a diffeomorphism from $\mathbb{R}\times\mathbb{R}^3\times\mathbb{R}^3$ onto itself. From (3.66), we easily get $|\Xi(\varepsilon;t,x,\xi)| = |\xi|$ for all $t\in\mathbb{R}$. Looking at (3.65,3.66), this means that the solution $(\mathrm{X},\Xi)(\cdot)$ remains in a compact set of $\mathbb{R}^3\times\mathbb{R}^3$, and therefore it is globally defined in time. For times $t\sim 1$, the flow $\mathrm{F}(\cdot)$ involves oscillations at the frequency ε^{-1}, and the main effect is a fast rotation (called gyration) around the field lines. We refer to the article [12] for a precise description of the flow $(\mathrm{X},\Xi)(\cdot)$ when the magnetic field $\mathbf{B}_e(\cdot)$ points in a fixed direction like in (4.62), and to [6, 7] for more general studies adapted to magnetospheres and tokamaks.

In what follows, our aim is to apply a sort of *filtering method* to get rid at the level of (2.21) or (3.66) of the singular factor $\varepsilon^{-2}\nu(\varepsilon\xi) = \varepsilon^{-1}\langle\varepsilon\xi\rangle^{-1}\xi = \mathcal{O}(\varepsilon^{-1})$, which can affect the local existence on a uniform time. In practice, this means to follow the particles along the oscillating trajectories associated to (3.65,3.66). This can be done by replacing $f(\cdot)$ into $g(\cdot)$ as indicated below

$$g(s,y,\eta) := f\circ \mathrm{F}(s,y,\eta) := f\big(s,\mathrm{X}(s,y,\eta),\Xi(s,y,\eta)\big) \tag{3.68}$$

$$f(t,x,\xi) := g\circ \mathrm{F}^{-1}(t,x,\xi) := g\big(t,\mathrm{X}(-t,x,\xi),\Xi(-t,x,\xi)\big) \tag{3.69}$$

Formulated in terms of $g(\cdot)$, the VW system (2.21,2.41) becomes (in conservative form)

$$\Box_{t,x}\, u(t,x,\xi) = -\ g\big(t,\mathrm{X}(-t,x,\xi),\Xi(-t,x,\xi)\big) \tag{3.70}$$

$$\begin{aligned} \partial_t g(t,x,\xi) = &+|\xi|^{-1}\, M'(|\xi|)\ \Xi(t,x,\xi)\cdot \boldsymbol{E}\big(t,\mathrm{X}(t,x,\xi)\big) \\ &+ \big\{\nabla_\xi\cdot\big[(\boldsymbol{E}+\nu\wedge\boldsymbol{B})f\big]\big\}\big(t,\mathrm{X}(t,x,\xi),\Xi(t,x,\xi)\big) \end{aligned} \tag{3.71}$$

The initial data $g_{|t=0}$ and $u_{|t=0}$ are as in (2.47,2.42). In particular, we still have $g_{|t=0}\equiv f^{in}$. On the other hand, the two identities (2.45) and (2.46) remain unchanged.

Lemma 3.7. *[Interpretation of $\boldsymbol{E}$ in terms of g] With D as in (3.58), the electric field $\boldsymbol{E}$ can be expressed according to*

$$\boldsymbol{E}(t,x) = \langle (D\otimes\mathbb{1}_\xi)\circ\tau_{t,x}\circ\mathcal{S}\circ\mathrm{F}\,,\,(g\mathbb{1}_{s>0})\rangle \tag{3.72}$$

where $\tau_{t,x}$ and $\mathcal{S}$ are the following time-space translation and symmetry

$$\tau_{t,x} : (s,y)\longmapsto (s+t,y+x) \quad ; \quad \mathcal{S} : (s,y)\longmapsto(-s,-y) \tag{3.73}$$

whereas the brackets $\langle\cdot,\cdot\rangle$ must be understood as an extension of the duality in $\mathcal{D}'(\mathbb{R}^4_{t,x}\times\mathbb{R}^3_\xi)$.

Proof. Introduce the abbreviated notation $\tilde f := f\mathbb{1}_{t>0}$. The pullback $\mathrm{F}^{-1}_*(\tilde f)$ of the generalized "test function" $\tilde f$ is as usual (see Definition 3.4.18 in [18]) given by

$$\mathrm{F}^{-1}_*(\tilde f) := \tilde f\circ\mathrm{F}\ \big|\det(D\mathrm{F}^{-1})\circ\mathrm{F}\big|^{-1} \tag{3.74}$$

Similarly, we can define $(\tau_{t,x})_*(\tilde f)$ and $\mathcal{S}_*(\tilde f)$. Taking into account (3.58), the formula (2.45) can be written as

$$\begin{aligned} \boldsymbol{E}(t,x) &= \int D*\tilde f d\xi = \int\langle D,\tilde f(t-s,x-y,\xi)\rangle d\xi \\ &= \big\langle D\otimes\mathbb{1}_\xi,\tilde f(t-\cdot,x-\cdot,\cdot)\big\rangle = \big\langle D\otimes\mathbb{1}_\xi,(\tau_{t,x})_*\circ\mathcal{S}_*(\tilde f)\big\rangle \\ &= \big\langle (D\otimes\mathbb{1}_\xi)\circ\tau_{t,x}\circ\mathcal{S},\tilde f\big\rangle \end{aligned} \tag{3.75}$$

The precise meaning of the above brackets $\langle\cdot,\cdot\rangle$ results from Subsections 3.1 and 3.3, see for instance (3.10). In fact, this can be viewed as some usual integral on $\mathbb{R}\times\mathbb{S}^2\times\mathbb{R}^3$. We have

$$\begin{aligned} \boldsymbol{E}(t,x) &= \langle (D\otimes\mathbb{1}_\xi)\circ\tau_{t,x}\circ\mathcal{S}\circ\mathrm{F}\circ\mathrm{F}^{-1},\tilde f\rangle \\ &= \langle (D\otimes\mathbb{1}_\xi)\circ\tau_{t,x}\circ\mathcal{S}\circ\mathrm{F},\mathrm{F}^{-1}_*(\tilde f)\rangle \end{aligned} \tag{3.76}$$

Recall that F_s is, for all s, a measure preserving $\mathcal{C}^1$-diffeomorphism on $\mathbb{R}^3 \times \mathbb{R}^3$. This remark is crucial. Combined with (3.68), it implies that $\mathrm{F}_*^{-1}(\tilde f) \equiv g \mathbb{1}_{s>0}$. The result (3.72) becomes a consequence of (3.76). $\square$

Introduce the auxiliary functions

$$K_\varepsilon^1(t,x,\xi) := \langle \varepsilon\xi \rangle^{-1} \big[\xi \wedge \mathbf{B}_e(x) \big] \tag{3.77}$$

$$K_\varepsilon^2(t,x,\xi) := \nabla_\xi p(t,x,\varepsilon\xi) \tag{3.78}$$

At the level of (3.72), replace D as indicated in (3.59). Now, the purpose is to pass from the derivative $T_\varepsilon(p_\varepsilon Y)$ to the derivative ∂_s of some expression plus some contribution of order zero. The interest of doing this is that ∂_s is on the left. This will allow to perform inside (3.72) a time integration by parts, while $\partial_s g$ is a "good" derivative since the right hand side of (3.71), in contrast with (2.21), does not contain the singular factor ε^{-2}.

Lemma 3.8. *[Second transfer of derivatives] With p_ε, q_ε, K_ε^1 and K_ε^2 as in (3.56), (3.77) and (3.78), we have*

$$[D \otimes \mathbb{1}_\xi] \circ \tau_{t,x} \circ \mathcal{S} \circ \mathrm{F} = D^1 + D^2 + \mathcal{D}^3 \tag{3.79}$$

with

$$D^1 := \partial_s \big[(p_\varepsilon Y \otimes \mathbb{1}_\xi) \circ \tau_{t,x} \circ \mathcal{S} \circ \mathrm{F} \big] \tag{3.80}$$

$$D^2 := (q_\varepsilon Y \otimes \mathbb{1}_\xi) \circ \tau_{t,x} \circ \mathcal{S} \circ \mathrm{F} \tag{3.81}$$

$$D^3 := (K_\varepsilon^1 \circ \mathrm{F}) \cdot \big[(Y K_\varepsilon^2 \otimes \mathbb{1}_\xi) \circ \tau_{t,x} \circ \mathcal{S} \circ \mathrm{F} \big] \tag{3.82}$$

Proof. By construction, given a locally integrable function $\psi(t,x,\xi)$, we have

$$\psi \circ \tau_{t,x} \circ \mathcal{S} \circ \mathrm{F}(s,y,\eta) = \psi\big(t-s, x - \mathrm{X}(s,y,\eta), \Xi(s,y,\eta)\big) \tag{3.83}$$

By testing (3.79) against a test function $g(\cdot)$ which is compactly supported in ξ, we can always work with ξ bounded. On the other hand, knowing that D is as in the right hand side of (3.59), the formula (3.79) is issued from the properties of the weight p_ε and of the flow F. It does not depend on the special structure of the distribution Y. It remains true for any smooth test function $\mathcal{Y}(t,x)$ whose support is conveniently localized to allow a multiplication by p_ε or by q_ε. In other words, the support of $\mathcal{Y}$ must be contained in a neighborhood of $\mathscr{LC}$, that is away from the origin and away from the singular set $\big\{(t,x); \nu(\varepsilon\xi)\cdot x - t = 0\big\}$. It suffices to show (3.79) in the case of such functions $\mathcal{Y}$. Then, a density argument gives (3.79) for distributions like Y. The smoothness of $\mathcal{Y}$ allows to exploit (3.83), and then to apply the chain rule as indicated below

$$\begin{aligned}
&\partial_s \big[p_\varepsilon\big(t-s, x-\mathrm{X}(s,y,\eta), \Xi(s)\big)\ \mathcal{Y}\big(t-s, x-\mathrm{X}(s,y,\eta)\big) \otimes \mathbb{1}_\eta \big] \\
&\quad = \partial_s \big[p_\varepsilon\big(t-s, x-\mathrm{X}(s,y,\eta), \Xi(s)\big)\big] \times \big[\mathcal{Y}\big(t-s, x-\mathrm{X}(s,y,\eta)\big) \otimes \mathbb{1}_\eta \big] \\
&\qquad - p_\varepsilon\big(t-s, x-\mathrm{X}(s,y,\eta), \Xi(s)\big) \times \big[T_\varepsilon\big(\Xi(s)\big) \mathcal{Y}\big(t-s, x-\mathrm{X}(s,y,\eta)\big) \otimes \mathbb{1}_\eta \big]
\end{aligned}$$

In the last line, we can commute the multiplication by p_ε with the derivative T_ε. The extra terms that are produced are compensated by terms coming from the second line. There remains

$$\begin{aligned}
&\partial_s \big[p_\varepsilon\big(t-s, x-\mathrm{X}(s,y,\eta), \Xi(s)\big)\ \mathcal{Y}\big(t-s, x-\mathrm{X}(s,y,\eta)\big) \otimes \mathbb{1}_\eta \big] \\
&\quad = \big[\varepsilon \dot\Xi \cdot \nabla_\xi p\big(t-s, x-\mathrm{X}(s,y,\eta), \varepsilon\Xi(s)\big)\big] \times \big[\mathcal{Y}\big(t-s, x-\mathrm{X}(s,y,\eta)\big) \otimes \mathbb{1}_\eta \big] \\
&\qquad - T_\varepsilon\big(\Xi(s)\big) \Big\{ p_\varepsilon\big(t-s, x-\mathrm{X}(s,y,\eta), \Xi(s)\big) \times \big[\mathcal{Y}\big(t-s, x-\mathrm{X}(s,y,\eta)\big) \otimes \mathbb{1}_\eta \big] \Big\}
\end{aligned}$$

With (2.3) and (3.66), we find $\varepsilon \dot\Xi = -K_\varepsilon^1 \circ \mathrm{F}$. Since $p(t,x,\cdot)$ is not only a function of $|\xi|$, we have $K_\varepsilon^1 \cdot K_\varepsilon^2 \not\equiv 0$. This means that the expression which inside (3.79) involves the functions K_ε^* does contribute. The weight $K_\varepsilon^1(\cdot)$ is, on the compact sets of $\mathbb{R}^3 \times \mathbb{R}^3$, uniformly bounded with respect to $\varepsilon \in (0,1]$. This is due to a compensation between the factor ε put in front of ξ inside p_ε (and issued from the cold assumption) and the singular factor ε^{-1} coming from (3.66). This would not be verified in the hot case, that is if $\varepsilon\,\xi$ would be replaced by ξ. Now, coming back to (3.59), we can deduce (3.79). $\square$

In view of Lemmas 3.7 and 3.8, we have

$$\boldsymbol{E} = \boldsymbol{E}^1 + \boldsymbol{E}^2 + \boldsymbol{E}^3 \quad ; \quad \boldsymbol{E}^i(t,x) = \big\langle D^i, (g\mathbb{1}_{s>0})\big\rangle \tag{3.84}$$

4. Well-posedness of the Cauchy problem

After a presentation in Subsection 4.1 of the functional framework, this section addresses the questions raised in Paragraph 2.1.4: uniform control in the sup norm in Subsection 4.2 and Lipschitz estimates in Subsection 4.3.

4.1. **The functional framework.** Fix an initial condition $\boldsymbol{U}^{in}(\cdot)$ satisfying (2.26). As was explained in Paragraph 2.43, the introduction of a potential $\boldsymbol{A}^0(\cdot)$ satisfying (2.43) allows to absorb $\boldsymbol{E}^{in}$ and $\boldsymbol{B}^{in}$. This is why, from now on, we will work with $\boldsymbol{E}^{in} \equiv 0$ and $\boldsymbol{B}^{in} \equiv 0$, while the initial condition f^{in} is aimed to vary.

With such $\boldsymbol{U}^{in} = (f^{in}, 0, 0)$, we can associate some initial data to the RVM system, as indicated in (2.17,2.18,2.19). Under Assumption 2.4, it is a well known fact [16, 28] that a classical solution $\mathbf{U} = (\mathrm{f}, \mathbf{E}, \mathbf{B})$ exists on a time interval $[0, \mathrm{T}_\varepsilon)$ with $\mathrm{T}_\varepsilon \in \mathbb{R}_+^*$. Let T_ε be the maximal time T_ε that can be obtained by this way. The maximum time T_ε is called the *lifespan* of the solution. As a consequence of [16, 28], the time T_ε can be bounded below by a constant $\delta_\varepsilon \in \mathbb{R}_+^*$ that depends only on the Lipschitz norm of f^{in}.

Interpreted according to (2.20) in terms of $\boldsymbol{U} = (f, \boldsymbol{E}, \boldsymbol{B})$, these results also furnish on $[0, T_\varepsilon)$ the local existence in time and the uniqueness of a classical solution to the MRVM system. For all $\varepsilon \in]0,1]$ and all time $t \in [0, T_\varepsilon)$, with $\mathcal{X}$ defined as in (2.33), there is a solution operator

$$\begin{aligned} \mathrm{S}_\varepsilon^t : \mathcal{X} &\longrightarrow \mathcal{X} \\ \boldsymbol{U}^{in} &\longmapsto \mathrm{S}_\varepsilon^t(\boldsymbol{U}^{in}) := \boldsymbol{U}(t,\cdot) = (f, \boldsymbol{E}, \boldsymbol{B})(t,\cdot) \end{aligned}$$

By this way, we recover families of solutions $\big(\mathrm{S}_\varepsilon(\boldsymbol{U}^{in})\big)_\varepsilon$ depending on the choice of ε and $\boldsymbol{U}^{in}$. In Paragraphs 4.1.1 and 4.1.2, we introduce definitions allowing to describe precisely what happens.

4.1.1. *Norms, bounded families and prepared data.* Different norms can be put on $\mathcal{X}$, like

$$\mathcal{N}(\boldsymbol{U}) := \parallel f \parallel_{L^\infty_{x,\xi}} + \parallel (\boldsymbol{E}, \boldsymbol{B}) \parallel_{L^\infty_x} \tag{4.1}$$

$$\mathcal{N}_1^1(\boldsymbol{U}) := \parallel f \parallel_{L^\infty_{x,\xi}} + \parallel (\boldsymbol{E}, \boldsymbol{B}) \parallel_{L^\infty_x} + \parallel \nabla_{x,\xi} f \parallel_{L^\infty_{x,\xi}} + \parallel \nabla_x (\boldsymbol{E}, \boldsymbol{B}) \parallel_{L^\infty_x} \tag{4.2}$$

The norm $\mathcal{N}$ is just the sup norm on L^∞; the norm $\mathcal{N}_1^1$ is the usual Lipschitz norm on $W^{1,\infty}$. Solving the MRVM system for all $\varepsilon \in]0,1]$ for a fixed initial condition $\boldsymbol{U}^{in}$ generates a family of solutions $(\boldsymbol{U}_\varepsilon)_\varepsilon$. Accordingly, we can introduce on $\mathcal{X}$ families of norms indexed by ε. Typically, we can consider

$$\begin{aligned} \mathcal{N}_\varepsilon^1(\boldsymbol{U}) := {} & \parallel f \parallel_{L^\infty_{x,\xi}} + \parallel (\boldsymbol{E}, \boldsymbol{B}) \parallel_{L^\infty_x} \\ & + \parallel \varepsilon \nabla_x f \parallel_{L^\infty_{x,\xi}} + \parallel \nabla_\xi f \parallel_{L^\infty_{x,\xi}} + \parallel \varepsilon \nabla_x (\boldsymbol{E}, \boldsymbol{B}) \parallel_{L^\infty_x} \end{aligned} \tag{4.3}$$

When computing $\mathcal{N}_\varepsilon^1(\boldsymbol{U})$, there is a difference of treatment between derivatives with respect to x and ξ. Precisely, the use of $\mathcal{N}_\varepsilon^1$ is a way to change how the functions are asymptotically evaluated when ε goes to zero. Obviously, we have

$$\forall \varepsilon \in]0,1], \quad \mathcal{N}(\boldsymbol{U}) \leq \mathcal{N}_\varepsilon^1(\boldsymbol{U}) \leq \mathcal{N}_1^1(\boldsymbol{U}) \quad ; \quad \mathcal{X}_1^1 \hookrightarrow \mathcal{X}_\varepsilon^1 \hookrightarrow \mathcal{X} \tag{4.4}$$

We can look at $\mathcal{X}$ as a normed space equipped with the sup norm. We can also define $\mathcal{X}_1^1$ and $\mathcal{X}_\varepsilon^1$ as the Banach spaces obtained by looking at $\mathcal{X}$ respectively with the norms $\mathcal{N}_1^1$ and $\mathcal{N}_\varepsilon^1$. We denote by $\mathcal{X}^\star$ with $\star \in \{\ , {}^1_1, {}^1_\varepsilon\}$ the functional space $\mathcal{X}$ equipped with the norm $\mathcal{N}^\star$. To study the MRVM system, the sole estimation of $\mathcal{N}^\star(\boldsymbol{U})$ does not suffice. It must be completed with a control on the momentum support. This motivates the following notion of bounded set.

Definition 4.1 (bounded set on $\mathcal{X}^\star$)**.** A family of subsets $(B_\varepsilon)_\varepsilon$ with $\varepsilon \in]0,1]$ and $B_\varepsilon \subset \mathcal{X}$ is said to be bounded according to $\mathcal{X}^\star$ with $^\star \in \{\ ,{}^1_1, {}^1_\varepsilon\}$ if:

(a) There exists a constant $C \in \mathbb{R}_+^*$ such that

$$\forall (\varepsilon, \boldsymbol{U}) \in]0,1] \times B_\varepsilon, \qquad \begin{cases} \mathcal{N}^\star(\boldsymbol{U}) \leq C & \text{if} \quad \star \in \{\ , {}^1_1\}, \\ \mathcal{N}^1_\varepsilon(\boldsymbol{U}) \leq C & \text{if} \quad \star \equiv {}^1_\varepsilon. \end{cases} \tag{4.5}$$

(b) There exists $(\mathrm{R}^{in}, R^{in}) \in (\mathbb{R}_+^*)^2$ such that (2.29) is verified for all $(\varepsilon, \boldsymbol{U}) \in]0,1] \times B_\varepsilon$. ∘

By extension, we say that B is bounded in $\mathcal{X}^\star$ if the stationary family $(B_\varepsilon)_\varepsilon$ with $B_\varepsilon = B$ is bounded according to $\mathcal{X}^\star$. Accordingly, a set B is bounded according to $\mathcal{X}^1_\varepsilon$ when all elements of B satisfy (2.29) for some $(\mathrm{R}^{in}, R^{in}) \in (\mathbb{R}_+^*)^2$, and when

$$\exists C \in \mathbb{R}_+^*; \quad \forall (\varepsilon, \boldsymbol{U}) \in]0,1] \times B, \quad \mathcal{N}^1_\varepsilon(\boldsymbol{U}) \leq C \tag{4.6}$$

In view of (4.4), this is equivalent to

$$\exists C \in \mathbb{R}_+^*; \quad \forall \boldsymbol{U} \in B, \quad \mathcal{N}^1_1(\boldsymbol{U}) \leq C \tag{4.7}$$

which means that B is bounded according to $\mathcal{X}^1_1$. Thus, for a fixed B, the notions of boundedness in $\mathcal{X}^1_\varepsilon$ and $\mathcal{X}^1_1$ coincide. But, when B does depend on ε, they can differ. An interesting situation is when B_ε is given by a singleton, typically when $B_\varepsilon \equiv \{\boldsymbol{U}_\varepsilon\}$ where $\boldsymbol{U}_\varepsilon$ is a solution to the MRVM system. In this case, we say that a family of functions $(\boldsymbol{U}_\varepsilon)_\varepsilon$ with $\varepsilon \in]0,1]$ and $\boldsymbol{U}_\varepsilon \in \mathcal{X}$ is bounded on $\mathcal{X}^\star$ when the family of unit sets $\big(\{\boldsymbol{U}_\varepsilon\}\big)_\varepsilon$ is bounded according to $\mathcal{X}^\star$. For the choice $^\star \equiv {}^1_\varepsilon$, this amounts to the same thing as

$$\exists C \in \mathbb{R}_+^*; \quad \forall \varepsilon \in]0,1], \quad \mathcal{N}^1_\varepsilon(\boldsymbol{U}_\varepsilon) \leq C \tag{4.8}$$

The time derivative is not estimated when computing the weighted Lipschitz norm $\mathcal{N}^1_\varepsilon(\boldsymbol{U}_\varepsilon)$ of a solution to the MRVM system. But, as this will be seen in Paragraph 4.3.1, it is deeply linked to spatial derivatives of $\boldsymbol{U}_\varepsilon$, and of the same size. Then, in view of (1.6), the following supplementary condition seems to be necessary to get families $(\boldsymbol{U}_\varepsilon)_\varepsilon$ of solutions to the MRVM system that could be bounded in $\mathcal{X}^1_1$.

Definition 4.2 (prepared data)**.** A family of subsets $(B_\varepsilon)_\varepsilon$ with $\varepsilon \in]0,1]$ and $B_\varepsilon \subset \mathcal{X}$ is said to be *prepared* if:

(a) The family $(B_\varepsilon)_\varepsilon$ is bounded according to $\mathcal{X}^1_1$.
(b) There exists a constant $C \in \mathbb{R}_+^*$ such that

$$\forall \varepsilon \in]0,1], \quad \forall U = (f, \boldsymbol{E}, \boldsymbol{B}) \in B_\varepsilon, \quad \| \, [\xi \times \mathbf{B}_e(x)] \cdot \nabla_\xi f \, \|_{L^\infty_{x,\xi}} \leq C\varepsilon \tag{4.9}$$

In particular, a family $(\boldsymbol{U}_\varepsilon)_\varepsilon$ with $\varepsilon \in]0,1]$ and $\boldsymbol{U}_\varepsilon = (f_\varepsilon, \boldsymbol{E}_\varepsilon, \boldsymbol{B}_\varepsilon) \in \mathcal{X}$ is said to be prepared if, viewed as the family of unit sets $(\{\boldsymbol{U}_\varepsilon\})_\varepsilon$, it is prepared. ∘

When $(f_\varepsilon)_\varepsilon$ is stationnary, with $f_\varepsilon \equiv f$ for all ε, the condition (4.9) is the same as

$$[\xi \times \mathbf{B}_e(x)] \cdot \nabla_\xi f = 0 \tag{4.10}$$

Given $\varepsilon_0 \in]0,1]$ and a family $(B_\varepsilon)_\varepsilon$ that is bounded in $\mathcal{X}^1_\varepsilon$, we can define the finite bound

$$\delta^\infty_1(\varepsilon_0) \equiv \delta^\infty_1\big((B_\varepsilon)_\varepsilon, \varepsilon_0\big) := \sup_{\varepsilon \in]0,\varepsilon_0]} \ \sup_{\boldsymbol{U} \in B_\varepsilon} \mathcal{N}^1_\varepsilon(\boldsymbol{U}) < +\infty \tag{4.11}$$

4.1.2. *Different notions of local well-posedness.* There exists many different ways of defining what is a well-posed Cauchy problem, see for instance [13]. In below, we introduce definitions that seem to be particularly adapted to the MRVM framework.

Definition 4.3 (conditional local well-posedness in $\mathcal{X}^\star$)**.** We say that the Cauchy problem for the MRVM system is *locally well-posed with uniform bounds in* $\mathcal{X}^\star$ if, for every family $(B_\varepsilon)_\varepsilon$ of bounded subsets in $\mathcal{X}^1_\varepsilon$, there exists a time $T \in \mathbb{R}_+^*$ such that:

(i) The family of mappings $(\mathrm{S}_\varepsilon)_\varepsilon$ is uniformly bounded. More precisely, we can find a modulus of continuity $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ and a constant $C \in \mathbb{R}_+^*$ such that

$$\forall (\varepsilon, \boldsymbol{U}^{in}, t) \in]0,1] \times B_\varepsilon \times \big[0, \min(T_\varepsilon, T)\big), \quad \mathcal{N}^\star\big(\mathrm{S}_\varepsilon^t(\boldsymbol{U}^{in})\big) \leq \omega\big(\mathcal{N}^\star(\boldsymbol{U}^{in})\big) \leq C \tag{4.12}$$

(ii) The continuation criterion on the momentum support is preserved in the following sense. We can find a bounded function $R(\cdot) \in L^\infty([0,T])$ such that, for all $\varepsilon \in]0,1]$ and for all initial data $\boldsymbol{U}^{in} \in B_\varepsilon$, the solution $\boldsymbol{U}_\varepsilon = (f_\varepsilon, \boldsymbol{E}_\varepsilon, \boldsymbol{B}_\varepsilon)$ to the MRVM system is satisfying

$$\forall t \in \big[0, \min(T_\varepsilon, T)\big), \quad \sup f_\varepsilon(t,\cdot) \subset \big\{(x,\xi)\,;\, |x| \leq \mathrm{R}^{in} + t\,,\ |\xi| \leq R(t)\big\} \tag{4.13}$$

where $R(\cdot)$ is some nondecreasing function on $\mathbb{R}_+$. ∘

In what follows, we will use (4.13) with $R(t) = R^{ini} + w(\delta_0^\infty)$ and δ_0^∞ given by (4.47). Applied in the case of $\mathcal{X}$, Definition 4.3 furnishes uniform bounds in the sup norm, while involving some regularity assumption. Indeed, the family $(B_\varepsilon)_\varepsilon$ is *a priori* assumed to be bounded in $\mathcal{X}_\varepsilon^1$. When dealing with $\mathcal{X}_\varepsilon^1$, keep in mind that $\mathcal{N}^\star$ must be replaced by $\mathcal{N}_\varepsilon^1$ at the level of (4.12). In both cases, Definition 4.3 imposes (L^∞ or Lipschitz) uniform bounds on the time interval $[0, T_\varepsilon)$ of existence. But there is no condition (especially no uniform minoration) on T_ε. This other aspect is taken into account below.

Definition 4.4 (uniform local well-posedness)**.** We say that the Cauchy problem for the MRVM system is *uniformly locally well-posed* if, for every family $(B_\varepsilon)_\varepsilon$ of bounded subsets in $\mathcal{X}_\varepsilon^1$, there exists a time $T \in \mathbb{R}_+^*$ such that:

(i) For all $\varepsilon \in]0,1]$ and for all initial condition $\boldsymbol{U}^{in} \in B_\varepsilon$, the MRVM system has a unique solution $\boldsymbol{U}_\varepsilon(\cdot)$ which is defined on $[0,T]$, and which satisfies

$$\boldsymbol{U}_\varepsilon(\cdot) \in \mathcal{C}\big([0,T];\mathcal{X}\big) \quad ; \quad \boldsymbol{U}_{\varepsilon|t=0} = \boldsymbol{U}^{in} \in B_\varepsilon \tag{4.14}$$

(ii) The Cauchy problem is locally well-posed with uniform bounds in $\mathcal{X}_\varepsilon^1$. ∘

Paragraph (i) of Definition 4.4 ensures the existence of some $T \in \mathbb{R}_+^*$ such that $0 < T \leq T_\varepsilon$ for all $\varepsilon \in]0,1]$. Then, Paragraph (ii) furnishes the validity of (4.12) and (4.13) on $[0,T]$, for a possibly smaller $T \in \mathbb{R}_+^*$. Now, it is expected that the life span T_ε becomes larger as the initial condition gets smaller. This prediction can be formalized as indicated below.

Definition 4.5 (uniform long time well-posedness for small data)**.** We say that the Cauchy problem for the MRVM system is *for small data uniformly well-posed for a long time* when, for all $T \in \mathbb{R}_+^*$, we can find $\varepsilon_0 \in]0,1]$ and $\delta_1^\infty \in]0,1]$ such that, for all family $(B_\varepsilon)_\varepsilon$ satisfying (4.11) with $\delta_1^\infty(\varepsilon_0) < \delta_1^\infty$, the following holds true:

(i) For all $\varepsilon \in]0,\varepsilon_0]$ and for all initial condition $\boldsymbol{U}^{in} \in B_\varepsilon$, the MRVM system has a unique solution $\boldsymbol{U}_\varepsilon(\cdot)$ which is defined on $[0,T]$ satisfying (4.14).

(ii) The family of mappings $\{\mathrm{S}_\varepsilon\}_\varepsilon$ is uniformly bounded. More precisely, we can find a modulus of continuity $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ and a constant $C \in \mathbb{R}_+^*$ such that

$$\forall (\varepsilon, \boldsymbol{U}^{in}, t) \in]0,\varepsilon_0] \times B_\varepsilon \times [0,T], \quad \mathcal{N}\big(\mathrm{S}_\varepsilon^t(\boldsymbol{U}^{in})\big) \leq \omega\big(\mathcal{N}(\boldsymbol{U}^{in})\big) \leq C \tag{4.15}$$

(iii) The continuation criterion on the momentum support is preserved on $[0,T]$. ∘

In Subsections 4.2 and 4.3, we will progressively consider situations where $\mathcal{X}^\star$ is equal to $\mathcal{X}$, $\mathcal{X}_\varepsilon^1$ and finally $\mathcal{X}_1^1$.

4.2. Uniform estimates in the sup norm.

We define on $L^\infty\big([0,t];\mathcal{X}\big)$ the following norms

$$\mathcal{N}_t(f) := \sup_{0 \leq s \leq t} ||f(s,\cdot,\cdot)||_{L^\infty_{x,\xi}} \quad ; \quad \mathcal{N}_t(\boldsymbol{E},\boldsymbol{B}) := \sup_{0 \leq s \leq t} ||(\boldsymbol{E},\boldsymbol{B})(s,\cdot)||_{L^\infty_x} \tag{4.16}$$

as well as

$$\mathcal{N}_t \equiv \mathcal{N}_t(\boldsymbol{U}) \equiv \mathcal{N}_t(f,\boldsymbol{E},\boldsymbol{B}) := \mathcal{N}_t(f) + \mathcal{N}_t(\boldsymbol{E},\boldsymbol{B}) \tag{4.17}$$

Norms related to $L^q_\xi(L^p_x)$ or $L^p_x(L^q_\xi)$ are commonly used in kinetic equations. As noted in [31], a control of the density $\mathrm{f}(\cdot)$ in $L^\infty_x(L^1_\xi)$ can serve as a substitute for the Glassey-Strauss criterion of explosion concerning the RVM system. The corresponding techniques can be exploited as long as there is no sign change at the level of $\mathrm{f}(\cdot)$.

But this approach is not at all adapted to the actual framework. As a matter of fact, when dealing with the MRVM system, there is no sign condition on $f(\cdot)$. This is why, as in [16], we will work with the $L^\infty_{x,\xi}$-norm and with the usual support momentum condition on $f(\cdot)$. A key statement is the following.

Proposition 4.6. *[local uniform bounds in the sup norm] The Cauchy problem for the MRVM system is locally well-posed with uniform bounds in $\mathcal{X}$ (in the sense of Definition 4.3).*

The proof will be achieved in three steps. In Subsection 4.2.1, we control f. In Subsection 4.2.2, we control $(\boldsymbol{E}, \boldsymbol{B})$. Then, in Subsection 4.2.3, we show Proposition 4.6.

4.2.1. *L^∞-bounds on the density.* First, observe that the amplitude of $f(\cdot)$ can indeed increase, due to the source term $\xi \cdot \boldsymbol{E}$ inside (2.21). But this remains under control.

Lemma 4.7 (control of f in sup norm)**.** *Select a function $f^{in}(\cdot)$ satisfying Assumption 2.4 and a bounded field $(\boldsymbol{E}, \boldsymbol{B})(\cdot) \in \mathcal{C}^1([0,T] \times \mathbb{R}^3 \times \mathbb{R}^3)$. Then, the transport equation (2.21) with initial condition $f^{in}(\cdot)$ has a $\mathcal{C}^1$-solution $f(\cdot)$ on $[0,T]$, which is subjected to*

$$\forall t \in [0,T], \quad \mathcal{N}_t(f) \leq \mathcal{N}_0(f) + ||M'||_{L^\infty_\xi} \int_0^t \mathcal{N}_s(\boldsymbol{E}, \boldsymbol{B}) ds \tag{4.18}$$

Proof. The *complete* characteristic curves (X, Ξ) associated with (2.21) can be obtained by integrating the following dynamical system

$$\dot{X} = \nu(\varepsilon \Xi), \qquad X(0,x,\xi) = x \tag{4.19}$$

$$\dot{\Xi} = -\varepsilon^{-2} \nu(\varepsilon \Xi) \times \mathbf{B}_e(X) - \boldsymbol{E}(t,X) - \nu(\varepsilon \Xi) \times \boldsymbol{B}(t,X), \qquad \Xi(0,x,\xi) = \xi \tag{4.20}$$

The $\mathcal{C}^1$-regularity hypothesis made on $\boldsymbol{E}$ and $\boldsymbol{B}$ in Lemma 4.7 guarantees the local existence of $\mathcal{C}^1$-solutions to (4.19,4.20), at least up to a stopping time $T^* \leq T$. Looking at (4.19,4.24), it is easy to infer that

$$|\dot{X}| \leq 1 \quad ; \quad |\Xi \cdot \dot{\Xi}| \leq |\boldsymbol{E} \cdot \Xi| \tag{4.21}$$

from which we can deduce that

$$|x| \leq \mathrm{R}^{in} \implies \forall t \in [0,T^*), \quad |X(t,x,\xi)| \leq \mathrm{R}^{in} + t \tag{4.22}$$

$$|\xi| \leq R^{in} \implies \forall t \in [0,T^*), \quad |\Xi(t,x,\xi)| \leq R^{in} + 2 \int_0^t \mathcal{N}_s(\boldsymbol{E}, \boldsymbol{B}) ds \tag{4.23}$$

Starting from $(x,\xi) \in \mathbb{R}^3 \times \mathbb{R}^3$, the solution $(X, \Xi)(\cdot, x, \xi)$ does not leave on $[0,T^*)$ some well chosen compact subset, so that $T^* = T$. There is therefore on $[0,T]$ an associated flow

$$\mathcal{F} : (s,y,\eta) \longmapsto (t,x,\xi) := \big(s, \mathcal{F}_s(y,\eta)\big) \quad ; \quad \mathcal{F}_s(y,\eta) := \big(X(\varepsilon; s,y,\eta), \Xi(\varepsilon; s,y,\eta)\big) \tag{4.24}$$

which is area preserving, and which should not be confused with the *approximated* flow $\mathrm{F}(\cdot)$ defined by (3.65,3.66). Solving (2.21) on $[0,T)$ by integration along the complete characteristics gives rise to the following $\mathcal{C}^1$-solution

$$\begin{aligned} f(t,x,\xi) = {} & f^{in}\big(X(-t,x,\xi), \Xi(-t,x,\xi)\big) \\ & + \int_0^t \Big(M'(|\Xi|) \frac{\Xi \cdot \boldsymbol{E}}{|\Xi|} \Big) \big(s, X(s-t,x,\xi), \Xi(s-t,x,\xi)\big) ds \end{aligned} \tag{4.25}$$

which leads directly to (4.18). □

4.2.2. *L^∞-bounds on the fields.* Next, we consider the fields $\boldsymbol{E}$ and $\boldsymbol{B}$. To this end, as indicated in Subsection 2.2, we can interpret the MRVM system as a VW system in order to use (2.45,2.46) with $\boldsymbol{A}^0 \equiv 0$, or (3.60,3.61), or other identities established in Section 3.

Lemma 4.8 (control of $\boldsymbol{E}$ and $\boldsymbol{B}$ in sup norm)**.** *Any classical $\mathcal{C}^1$-solution $(f, \boldsymbol{E}, \boldsymbol{B})$ satisfying the MRVM system on $[0, T]$ as well as the momentum support condition*

$$\exists R^\infty \in [1, +\infty[\,; \quad \forall t \in [0, T], \quad \sup f(t, \cdot) \subset \big\{(x, \xi)\,;\, |\xi| \leq R^\infty\big\} \tag{4.26}$$

is, for all $t \in [0, T]$, satisfying

$$\begin{aligned} \mathcal{N}_t(\boldsymbol{E}, \boldsymbol{B}) \leq Cc(R^\infty)\Big[t\mathcal{N}_0(f) + \big(c(R^\infty)(t + (R^\infty)^3 + t(R^\infty)^4\big) \int_0^t \mathcal{N}_s ds \\ + \varepsilon tc(R^\infty)(R^\infty)^3 \int_0^t \mathcal{N}_s^2 ds\Big] \end{aligned} \tag{4.27}$$

Proof. We start by estimating the electric field $\boldsymbol{E}$. To this end, we can exploit (3.84), where the distributions D^i are given by (3.80,3.81,3.82). We first study $\boldsymbol{E}^1$. After one integration by parts with respect to the time variable s, we find

$$\begin{aligned} \boldsymbol{E}^1 = & - \int\int p_\varepsilon\big(t - 0, x - \mathrm{X}(0, y, \eta), \Xi(0, y, \eta)\big) Y\big(t - 0, x - \mathrm{X}(0, y, \eta)\big) g(0, y, \eta) dy d\eta \\ & - \int_0^t \int\int p_\varepsilon\big(t - s, x - \mathrm{X}(s, y, \eta), \Xi(s, y, \eta)\big) Y\big(t - s, x - \mathrm{X}(s, y, \eta)\big) \partial_s g(s, y, \eta) ds dy d\eta \end{aligned}$$

In the above line, replace $\partial_s g$ as indicated in (3.71). Then, make the change of variables based on the approximated flow $\mathrm{F}_s(\cdot)$ defined in (3.67). As already noted, this is (for all s) area preserving. To avoid confusions, introduce the notations

$$(y', \eta') := \big(\mathrm{X}(s, y, \eta), \Xi(s, y, \eta)\big)$$

By this way, we find that $\boldsymbol{E}^1 = \boldsymbol{E}^{1,1} + \boldsymbol{E}^{1,2} + \boldsymbol{E}^{1,3}$ with

$$\boldsymbol{E}^{1,1} := - \int\int p_\varepsilon(t, x - y, \eta) Y(t, x - y) f^{in}(y, \eta) dy d\eta \tag{4.28}$$

$$\boldsymbol{E}^{1,2} := - \int_0^t \int\int p_\varepsilon(t - s, x - y', \eta') Y(t - s, x - y') M'(|\eta'|) \frac{\eta'}{|\eta'|} \cdot \boldsymbol{E}(s, y') ds dy' d\eta' \tag{4.29}$$

$$\begin{aligned} \boldsymbol{E}^{1,3} := - \int_0^t \int\int p_\varepsilon\big(t - s, x - y', \eta'\big) Y(t - s, x - y') \\ \times \nabla_{\eta'} \cdot \big[\big(\boldsymbol{E}(s, y') + \nu(\varepsilon\eta') \times \boldsymbol{B}(s, y')\big) f\big] ds dy' d\eta' \end{aligned} \tag{4.30}$$

In (4.28), there is no time integration. Thus, we cannot directly apply Lemma 3.1. But we can proceed as in the proof of Lemma 3.1 to obtain

$$\boldsymbol{E}^{1,1}(t, x) = - \int \int_{\mathbb{S}^2} p_\varepsilon(1, \omega, \eta) \frac{\mathbb{1}_{t>0}}{4\pi} f^{in}(x - t\omega, \eta) t d\sigma d\eta \tag{4.31}$$

With R^{in} as in (2.29), we have $R^{in} \leq R^\infty < +\infty$. By Assumption 2.4, we know that

$$|\eta| \geq R^{in} \quad \Longrightarrow \quad f^{in}(\cdot, \eta) \equiv 0 \tag{4.32}$$

In view of (4.31) and (4.32), we have

$$|\boldsymbol{E}^{1,1}(t, x)| \leq \frac{|\mathbb{S}^2|}{4\pi} \int_{|\eta| \leq R^{in}} ||p_\varepsilon(1, \cdot, \eta)||_{L^\infty(\mathbb{S}^2)} ||f^{in}(\cdot, \eta)||_{L_x^\infty} t d\eta \tag{4.33}$$

On the other hand, for all $\varepsilon \in]0, 1]$ and $|\eta| \leq R^\infty$, we find that

$$\frac{1}{|\nu(\varepsilon\eta) \cdot \omega - 1|} \leq c(R^\infty) := \frac{\sqrt{1 + (R^\infty)^2}}{\sqrt{1 + (R^\infty)^2} - R^\infty} < +\infty \tag{4.34}$$

It follows that

$$||p_\varepsilon(1,\cdot,\eta)||_{L^\infty(\mathbb{S}^2)} \le 2c(R^\infty) \tag{4.35}$$

and therefore

$$|\boldsymbol{E}^{1,1}(t,x)| \le \frac{8\pi}{9}(R^{in})^3 c(R^\infty) t \mathcal{N}_0(f) \tag{4.36}$$

Look at the part $\boldsymbol{E}^{1,2}$ given by (4.29). Knowing that $p_\varepsilon(\cdot,\eta') \in \mathcal{M}_0$, we can apply Lemma 3.1 with $m=0$ to find

$$|\boldsymbol{E}^{1,2}(t,x)| \le \frac{t}{3}\int_0^t \int ||p_\varepsilon(1,\cdot,\eta')||_{L^\infty(\mathbb{S}^2)} |M'(|\eta'|)| \mathcal{N}_s(\boldsymbol{E},\boldsymbol{B}) ds d\eta' \tag{4.37}$$

where we used the convention (2.13). With (4.35), there remains

$$|\boldsymbol{E}^{1,2}(t,x)| \le \frac{2t}{3} c(R^\infty) ||M'||_1 \int_0^t \mathcal{N}_s(\boldsymbol{E},\boldsymbol{B}) ds \tag{4.38}$$

Now, consider the quantity $\boldsymbol{E}^{1,3}$ given by (4.30). Exploit Lemma 3.6 to extract

$$|\boldsymbol{E}^{1,3}(t,x)| \le \frac{\varepsilon t}{3}\int_0^t \int ||\nabla_\xi p(1,\cdot,\varepsilon\eta')||_{L^\infty(\mathbb{S}^2)} ||\big[\boldsymbol{E}(s,\cdot)+\nu(\varepsilon\eta')\times \boldsymbol{B}(s,\cdot)\big] f(s,\cdot,\eta')||_{L^\infty_x} ds d\eta'$$

From the definition (3.47) of p, we can compute $\nabla_\xi p(\cdot)$ and use (4.34) to obtain

$$|\eta'| \le R^\infty \implies ||\nabla_\xi p(1,\cdot,\varepsilon\eta')||_{L^\infty(\mathbb{S}^2)} \le 8c(R^\infty)^2 \tag{4.39}$$

Then, using the key information (4.26), there remains a quadratic form in $(\boldsymbol{E},\boldsymbol{B})$ and f, namely

$$|\boldsymbol{E}^{1,3}(t,x)| \le \frac{8\varepsilon t}{3} c(R^\infty)^2 \int_0^t \int_{|\eta'|\le R^\infty} \mathcal{N}_s(\boldsymbol{E},\boldsymbol{B}) \mathcal{N}_s(f) ds d\eta' \tag{4.40}$$

$$\le \frac{32\varepsilon t}{9}\pi c(R^\infty)^2 (R^\infty)^3 \int_0^t \mathcal{N}_s(\boldsymbol{E},\boldsymbol{B}) \mathcal{N}_s(f) ds \tag{4.41}$$

Combining (4.36), (4.38) and (4.41), we find

$$\begin{aligned} |\boldsymbol{E}^{1}(t,x)| \le C c(R^\infty) t \Big[\mathcal{N}_0(f) + c(R^\infty) \int_0^t \mathcal{N}_s(\boldsymbol{E},\boldsymbol{B}) ds \\ + \varepsilon c(R^\infty)(R^\infty)^3 \int_0^t \mathcal{N}_s(\boldsymbol{E},\boldsymbol{B}) \mathcal{N}_s(f) ds \Big] \end{aligned} \tag{4.42}$$

The quantity $\boldsymbol{E}^2$ is given by (3.84) with D^2 as in (3.81). The preceding strategy gives rise to

$$\begin{aligned} \boldsymbol{E}^2 &= \int_0^t \int\int q_\varepsilon\big(t-s, x-\mathrm{X}(s,y,\eta), \Xi(s,y,\eta)\big) Y\big(t-s, x-\mathrm{X}(s,y,\eta)\big) g(s,y,\eta) \\ &= \int_0^t \int\int q_\varepsilon(t-s, x-y', \eta') Y\big(t-s, x-y'\big) f(s,y',\eta') ds dy' d\eta' \end{aligned}$$

Since $q_\varepsilon \in \mathcal{M}_{-1}$, Lemma 3.1 with $m=-1$ furnishes

$$|\boldsymbol{E}^{2}(t,x)| \le \frac{1}{3}\int_0^t \int_{|\eta'|\le R^\infty} ||q_\varepsilon(1,\cdot,\eta')||_{L^\infty(\mathbb{S}^2)} ||f(s,\cdot,\eta')||_{L^\infty_x} ds d\eta' \tag{4.43}$$

From the definition (4.34) of q, we can deduce

$$|\eta'| \le R^\infty \implies ||q_\varepsilon(1,\cdot,\eta')||_{L^\infty(\mathbb{S}^2)} \le 2c(R^\infty)^2$$

It follows that

$$|\boldsymbol{E}^{2}(t,x)| \le \frac{8\pi}{9} c(R^\infty)^2 (R^\infty)^3 \int_0^t \mathcal{N}_s(f) ds \tag{4.44}$$

The part $\boldsymbol{E}^3$ is defined by (3.84) with $\mathcal{D}^3$ as in (3.82). For the same reasons as above, we find

$$\boldsymbol{E}^3 = \int_0^t \int\int (K_\varepsilon^1 \cdot K_\varepsilon^2)(t-s, x-y', \eta') Y\big(t-s, x-y'\big) f(s, y', \eta') ds dy' d\eta'$$

The two auxiliary functions $K_\varepsilon^1(\cdot)$ and $K_\varepsilon^2(\cdot)$ are given by (3.77) and (3.78). For $|\xi| \leq R^\infty$, we have $|K_\varepsilon^1| \leq CR^\infty$ uniformly in ε, t and x. On the other hand, the function $K_\varepsilon^2(\cdot,\xi) \equiv \nabla_\xi p(\cdot, \varepsilon\xi)$ is (near the cone $\mathscr{L}\mathscr{C}$) in $\mathcal{M}_0$. It can be estimated exactly as in (4.39). This time, Lemma 3.1 applied with $m = 0$ leads to

$$|\boldsymbol{E}^3(t,x)| \leq Ctc(R^\infty)^2 (R^\infty)^4 \int_0^t \mathcal{N}_s(f) ds \tag{4.45}$$

Finally, combining (4.42), (4.44) and (4.45), we find

$$\begin{aligned} |\boldsymbol{E}(t,x)| \leq & \, Cc(R^\infty)\Big[t\mathcal{N}_0(f) + tc(R^\infty)\int_0^t \mathcal{N}_s(\boldsymbol{E},\boldsymbol{B})ds + c(R^\infty)(R^\infty)^3 \int_0^t \mathcal{N}_s(f)ds \\ & + tc(R^\infty)(R^\infty)^4 \int_0^t \mathcal{N}_s(f)ds + \varepsilon tc(R^\infty)(R^\infty)^3 \int_0^t \mathcal{N}_s(\boldsymbol{E},\boldsymbol{B})\mathcal{N}_s(f)ds\Big] \end{aligned} \tag{4.46}$$

It remains to consider the expression $\boldsymbol{B}$ which is determined by (3.61). The discussion is exactly as above. It suffices to replace p_ε and q_ε by the functions a_ε^0 and a_ε^1 of (3.57), where a^0 and a^1 are as in (3.34,3.35), whereas a_0 and a_1 are given by (3.25). The expressions a_ε^0 and a_ε^1 satisfy the same features as p_ε and q_ε. Like p_ε and q_ε, they belong respectively to $\mathcal{M}_0$ and $\mathcal{M}_{-1}$. As a consequence, the bound on $|\boldsymbol{B}(t,x)|$ is the same as in the right hand side of (4.46), and therefore, with $\mathcal{N}_t$ as in (4.17), we can retain (4.27). $\square$

4.2.3. *Proof of Proposition 4.6.* Consider a family $(B_\varepsilon)_\varepsilon$ of bounded subsets in $\mathcal{X}_\varepsilon^1$. We can find some $\delta_0^\infty \in \mathbb{R}_+^*$ such that

$$\sup_{\varepsilon\in]0,1]} \sup_{\boldsymbol{U}\in B_\varepsilon} \mathcal{N}(\boldsymbol{U}) < \delta_0^\infty < +\infty \tag{4.47}$$

Fix some $T \in \mathbb{R}_+^*$. The matter is to show the subparagraphs (i) and (ii) of Definition 4.3. As already explained, under Assumption 2.4, there exists a unique classical solution to the MRVM system. This solution is defined on a time interval $[0, T_\varepsilon)$ with $0 < T_\varepsilon$. For the moment, select some $R^\infty > R^{in}$. Define $\mathcal{T}_\varepsilon$ as the maximal time inside $\big[0, \min(T, T_\varepsilon)\big]$ such that (4.13) is verified with $R(\cdot) \equiv R^\infty$ on $[0, \mathcal{T}_\varepsilon)$. By the continuity of the flow, we have $0 < \mathcal{T}_\varepsilon \leq T$ and

$$\forall t \in [0, \mathcal{T}_\varepsilon), \quad \sup f(t,\cdot) \subset \big\{(x,\xi)\,;\, |x| \leq \mathrm{R}^{in} + t\ ,\ |\xi| \leq R^\infty\big\} \tag{4.48}$$

In view of (4.48), for $t \in [0, \mathcal{T}_\varepsilon)$, we can apply Lemmas 4.7 and 4.8. Adding (4.18) and (4.27), we can easily see that, for all $t \in [0, \mathcal{T}_\varepsilon)$, we have

$$\mathcal{N}_t \leq \alpha + C \int_0^t g(\mathcal{N}_s) ds, \quad g(z) := \alpha + \beta z + \varepsilon\gamma z^2 \tag{4.49}$$

where $g \equiv g_{\varepsilon,\alpha,\beta,\gamma}$ depends on parameters α, β and γ given by

$$0 \leq \alpha \equiv \alpha(R^\infty, f^{in}) := c(R^\infty)\mathcal{N}(f^{in}) < \alpha^\infty := c(R^\infty)\delta_0^\infty \tag{4.50}$$

$$1 \leq \beta \equiv \beta(R^\infty, T) := 1 + c(R^\infty)^2\big[T + (R^\infty)^3 + T(R^\infty)^4\big] \tag{4.51}$$

$$1 \leq \gamma \equiv \gamma(R^\infty, T) := 1 + Tc(R^\infty)^2(R^\infty)^3 \tag{4.52}$$

The function $g(\cdot)$ is positive and nondecreasing on $[0, +\infty[$. It is therefore compatible with non linear extensions of Grönwall's inequalities [23]. Define

$$G(\lambda) \equiv G_{\varepsilon,\alpha,\beta,\gamma}(\lambda) := \int_{\sqrt{\alpha}}^{\lambda} \frac{dz}{g(z)} \quad ; \quad G(+\infty) := \int_{\sqrt{\alpha}}^{+\infty} \frac{dz}{g(z)} < +\infty \tag{4.53}$$

The function $G(\cdot)$ is positive on the interval $[0,\sqrt{\alpha}[$. It is nondecreasing on $]0,+\infty[$ onto the interval $]G(0),G(+\infty)[$. By construction, we have $G(\sqrt{\alpha})=0$. As a result, the Bihari-Lasalle inequality can be applied as long as the time t is bounded according to

$$0 \le Ct < G(+\infty) - G(\alpha) = \int_\alpha^{+\infty} \frac{dz}{g(z)} \tag{4.54}$$

In view of (4.50), it suffices to adjust t in such a way that

$$0 \le Ct \le \mathcal{T}(\alpha^\infty, R^\infty, T) := \int_\alpha^{+\infty} \frac{dz}{\alpha^\infty + \beta z + \gamma z^2} \tag{4.55}$$

The continuous function $\mathcal{T}(\alpha^\infty, R^\infty, \cdot)$ is decreasing on $[0,+\infty[$ with

$$0 < \mathcal{T}(\alpha^\infty, R^\infty, 0) \quad ; \quad \lim_{T\to+\infty} \mathcal{T}(\alpha^\infty, R^\infty, T) = 0$$

The choice of T can now be optimized by adjusting T in such a way that $T \equiv \mathrm{T} = \mathrm{T}(\alpha^\infty, R^\infty)$ with $C\mathrm{T} = \mathcal{T}(\alpha^\infty, R^\infty, \mathrm{T})$. Up to the time $\mathcal{T}_\varepsilon \le \mathrm{T}$, the Bihari-Lasalle inequality can be applied with $g_{1,\alpha,\beta,\gamma}(\cdot) \ge g_{\varepsilon,\alpha,\beta,\gamma}(\cdot)$ in place of $g_{\varepsilon,\alpha,\beta,\gamma}(\cdot)$ to obtain

$$\forall t \in [0,\mathcal{T}_\varepsilon), \quad \mathcal{N}_t \le \omega(\alpha) \equiv \omega_{\beta,\gamma}(\alpha) := G^{-1}_{1,\alpha,\beta,\gamma}\big(G_{1,\alpha,\beta,\gamma}(\alpha) + C\mathrm{T}\big) \tag{4.56}$$

where C is the constant coming from (4.49). The above function $\omega(\cdot)$ is clearly continuous, positive and nondecreasing on $[0,+\infty[$. On the other hand, remark that

$$\begin{aligned} G_{1,\alpha,\beta,\gamma}(\alpha) &= -\int_1^{1/\sqrt{\alpha}} \frac{dz}{1+\beta z+\gamma\alpha z^2} \le -\int_1^{1/\sqrt{\alpha}} \frac{dz}{1+\gamma+\beta z} \\ &\le \frac{1}{2\beta}\ln\alpha + \frac{1}{\beta}\ln(1+\gamma+\beta) \end{aligned}$$

It follows that, for $\alpha_0 \in \mathbb{R}_+^*$ small enough, we have

$$0 \le \alpha \le \alpha_0 \implies G_{1,\alpha,\beta,\gamma}(\alpha) + C \le 0 \implies 0 \le \omega(\alpha) \le G^{-1}(0) = \sqrt{\alpha} \implies \lim_{\alpha\to 0}\omega(\alpha) = 0$$

Thus, the function $\omega(\cdot)$ is indeed a modulus of continuity. Looking at (4.56), we can retain that

$$\forall t \in [0,\mathcal{T}_\varepsilon), \quad \mathcal{N}_t \le \omega\big(\mathcal{N}_0(f)\big) \le \omega_{\beta,\gamma}(\delta_0^\infty) \tag{4.57}$$

Coming back to (4.23), it follows that

$$\forall t \in [0,\mathcal{T}_\varepsilon), \quad |\Xi(t)| \le |\Xi(0)| + 2\omega_{\beta,\gamma}(\delta_0^\infty)t \tag{4.58}$$

which implies that

$$\forall t \in [0,\mathcal{T}_\varepsilon), \quad \sup f(t,\cdot) \subset \big\{(x,\xi)\,;\, |\xi| \le R(t) := R^{in} + 2\omega_{\beta,\gamma}(\delta_0^\infty)t\big\} \tag{4.59}$$

Given δ_0^∞, we can now adjust $T \equiv T(\delta_0^\infty)$ in such a way that

$$0 < T < \sup_{R^\infty > R^{in}} \min\Big(\mathrm{T}(\alpha^\infty, R^\infty)\,;\, \frac{R^\infty - R^{in}}{2\omega_{\beta(R^\infty,\mathrm{T}),\gamma(R^\infty,\mathrm{T})}(\alpha^\infty)}\Big) \tag{4.60}$$

This guarantees the existence of some finite $R_m^\infty > R^{in}$ satisfying $T < \mathrm{T}(\alpha^\infty, R_m^\infty)$ as well as

$$0 \le t \le T \implies R^{in} + 2\omega_{\beta(R^\infty,\mathrm{T}),\gamma(R^\infty,\mathrm{T})}(\alpha^\infty)t < R_m^\infty \tag{4.61}$$

In view of (4.59) and (4.61), as long as $\mathcal{T}_\varepsilon < \min(T,T_\varepsilon)$, we can exploit (4.48) with $R^\infty = R_m^\infty$. Furthermore, since the inequality inside (4.61) is strict, the control (4.48) remains true on some extended interval $[0,\mathcal{T}_\varepsilon + \bar{t})$ with $0<\bar{t}$, which contradicts the preceding definition of $\mathcal{T}_\varepsilon$. This means that $\mathcal{T}_\varepsilon = \min(T,T_\varepsilon)$. Then, the line (4.56) gives rise to (4.12) with $\mathcal{N}^\star \equiv \mathcal{N}$ and $\omega(\cdot)$ as in (4.56). On the other hand, the line (4.59) becomes the same as (4.13). □

Proposition 4.6 furnishes no information on T_ε. A main difficulty is to prove that T_ε can be bounded from below by a positive threshold $T \in \mathbb{R}_+^*$ which do not depend on $\varepsilon \in]0,1]$. To this end, Lipschitz estimates are necessary. They are also very useful to obtain stability or convergence results. All these aspects are investigated in the next subsection.

4.3. **Estimates in Lipschitz norm.** In order to clarify expectations, in Paragraph 4.3.1, we first examine what happens concerning a simple case.

4.3.1. *A toy model.* Assume for the moment that the external magnetic field $\mathbf{B}_e(\cdot)$ points towards a fixed direction, say the vertical direction ${}^t(0,0,1)$, so that

$$\mathbf{B}_e(x) = \mathrm{b}_e(x)\,{}^t(0,0,1) \quad ; \quad \big(\xi \times \mathbf{B}_e(x)\big)\cdot\nabla_\xi = \mathrm{b}_e(x)(\xi_2\partial_{\xi_1} - \xi_1\partial_{\xi_2}) \tag{4.62}$$

Then, from the Vlasov equation (2.21), just retain the singular part, that is

$$\partial_t \mathfrak{f} - \frac{\mathrm{b}_e(x)}{\varepsilon\langle\varepsilon\xi\rangle}(\xi_2\partial_{\xi_1} - \xi_1\partial_{\xi_2})\mathfrak{f} = 0, \quad \mathfrak{f}_{|t=0} = \mathfrak{f}^{in} \tag{4.63}$$

The momentum flow on $\mathbb{R}^3$ generated by (4.63) is made of fast rotations around the vertical axis

$$\Xi^m_\varepsilon(t,x,\xi) := \begin{pmatrix} \cos\big(t\mathrm{b}_e(x)/\varepsilon\langle\varepsilon\xi\rangle\big)\xi_1 - \sin\big(t\mathrm{b}_e(x)/\varepsilon\langle\varepsilon\xi\rangle\big)\xi_2 \\ \sin\big(t\mathrm{b}_e(x)/\varepsilon\langle\varepsilon\xi\rangle\big)\xi_1 + \cos\big(t\mathrm{b}_e(x)/\varepsilon\langle\varepsilon\xi\rangle\big)\xi_2 \\ \xi_3 \end{pmatrix} \tag{4.64}$$

and the solution issued from (4.63) is just

$$\mathfrak{f}(t,x,\xi) = \mathfrak{f}^{in}\big(x, \Xi^m_\varepsilon(-t,x,\xi)\big) \tag{4.65}$$

Introduce polar coordinates in the plane $\{(\xi_1,\xi_2)\in\mathbb{R}^2\}$, so that

$$(\xi_1,\xi_2) = r(\cos\theta, \sin\theta), \qquad \partial_\theta = \xi_2\partial_{\xi_1} - \xi_1\partial_{\xi_2}, \qquad (r,\theta)\in\mathbb{R}_+\times\mathbb{R} \tag{4.66}$$

With this convention, we find

$$\partial_t\mathfrak{f}(t,x,\xi) = \frac{\mathrm{b}_e(x)}{\varepsilon\langle\varepsilon\xi\rangle}\partial_\theta\mathfrak{f}^{in}\big(x,\Xi^m_\varepsilon(t,x,\xi)\big) = \mathcal{O}\Big(\frac{1}{\varepsilon}\Big) \tag{4.67}$$

$$\partial_i\mathfrak{f}(t,x,\xi) = \frac{t\partial_i\mathrm{b}_e(x)}{\varepsilon\langle\varepsilon\xi\rangle}\partial_\theta\mathfrak{f}^{in}\big(x,\Xi^m_\varepsilon(t,x,\xi)\big) = \mathcal{O}\Big(\frac{1}{\varepsilon}\Big) \tag{4.68}$$

$$|\partial_{\xi_i}\mathfrak{f}(t,x,\xi)| \leq \parallel \nabla_\xi\mathfrak{f}^{in}(x,\cdot)\parallel_{L^\infty_{x,\xi}} = \mathcal{O}(1) \tag{4.69}$$

Fix an initial condition $\mathfrak{f}^{in}(\cdot)$ satisfying Assumption 2.4. From (4.67,4.68,4.69), it can be easily seen that

$$\exists C\in\mathbb{R}_+^*; \quad \forall(\varepsilon,t)\in]0,1]\times\mathbb{R}, \quad \mathcal{N}^1_\varepsilon\big(\mathfrak{f}(t,\cdot)\big)\leq C \tag{4.70}$$

This basic example is very instructive. First, it clearly indicates the relevance of the norm $\mathcal{N}^1_\varepsilon$. Secondly, we can see that the family $(\mathfrak{f}_\varepsilon)_\varepsilon$ is bounded in $\mathcal{X}^1_1$ if and only if the initial data $\mathfrak{f}^{in}(\cdot)$ is prepared in the sense of Definition 4.2, that is if we have (4.10). As this will be seen in the next paragraphs, consideration of the MRVM system can mix things and make them more difficult, but this does not change the above conclusions.

4.3.2. *Preparation of the Vlasov equation in view of commutator estimates.* To get Lipschitz bounds on $f(\cdot)$, a difficulty is to commute the Vlasov part with derivatives ∂ such as ∂_t, ∂_{x_i} or ∂_{ξ_i}. Most complications are due to the variable coefficients (in both x and ξ) which appear at the level of (2.21) in front of the singular factor ε^{-1}. As a matter of fact, we find

$$\Big[\,\partial\,;\ \frac{1}{\varepsilon^2}\nu(\varepsilon\xi)\times\mathbf{B}_e(x)]\cdot\nabla_\xi\,\Big] = \mathcal{O}\Big(\frac{1}{\varepsilon}\Big)\nabla_\xi \tag{4.71}$$

This information is not sufficient in view of uniform estimates. To remedy this, we will use two types of arguments. The first **(a)** is adapted to spatial derivatives; the second **(b)** is aimed to deal with momentum derivatives.

- (a) Straightening of the field lines. The purpose here is to recover (4.62). To this end, select a smooth frame field $\mathrm{O}(\cdot)$ such that

$$\begin{array}{rcl} \mathrm{O}:\mathbb{S}^2 & \longrightarrow & SO(3) \\ \xi & \longmapsto & \big(e_1(\xi), e_2(\xi), e_3(\xi) := \xi\big) \end{array}$$

In other words, the vector fields $e_i(\cdot)$ are $\mathcal{C}^\infty$ on the sphere $\mathbb{S}^2$, and we have

$$\forall (i,j) \in \{1,2,3\}^2, \quad e_i(\xi) \cdot e_j(\xi) = \delta_{ij}$$

Then, define

$$O(x) := \mathrm{O}\big(\mathrm{b}_e(x)^{-1}\mathbf{B}_e(x)\big) = {}^tO(x)^{-1}, \quad \mathfrak{f}(t,x,\xi) := f\big(t,x,O(x)\xi\big) \tag{4.72}$$

The role of the orthonormal matrix $O(x)$ is to fix the direction of $\mathbf{B}_e(x)$ through the relation

$${}^tO(x)\mathbf{B}_e(x) = \mathrm{b}_e(x)^t(0,0,1) \tag{4.73}$$

Expressed in terms of $\mathfrak{f}$, the transport equation (2.21) becomes

$$\begin{aligned} \partial_t \mathfrak{f} + & \frac{\varepsilon}{\langle \varepsilon\xi\rangle} O(x)\xi \cdot \nabla_x \mathfrak{f} + \frac{\varepsilon}{\langle \varepsilon\xi\rangle} Q(x,\xi)\cdot \nabla_\xi \mathfrak{f} - \frac{\mathrm{b}_e(x)}{\varepsilon\langle \varepsilon\xi\rangle}(\xi_2\partial_{\xi_1} - \xi_1\partial_{\xi_2})\mathfrak{f} \\ & = M'(|\xi|)\,\frac{O(x)\xi\cdot \boldsymbol{E}}{|\xi|} + [{}^tO(x)\boldsymbol{E} + \nu(\varepsilon\xi)\times {}^tO(x)\boldsymbol{B})]\cdot\nabla_\xi\mathfrak{f} \end{aligned} \tag{4.74}$$

where $Q(x,\xi) := \big(O(x)\xi\cdot\nabla_x\big)O(x)\,O(x)\xi$ is some vector valued quadratic form in ξ. In Paragraph 4.3.4, to show Proposition 4.9, we will directly commute (4.74) with $\varepsilon\nabla_{t,x}$. In Paragraph 4.3.8, to show Proposition 4.12, we will first divide (4.74) by $\mathrm{b}_e(x)$ and then commute (4.74) with $\nabla_{t,x}$. These two sorts of arguments are inspired from works in geometrical optics [25, 26].

- **(b)** Filtering of the equation. Inspired by (4.65), we can further replace $\mathfrak{f}(\cdot)$ by $\mathsf{f}(\cdot)$ with

$$\mathsf{f}(t,x,\xi) := \mathfrak{f}\big(t,x,\Xi^m_\varepsilon(t,x,\xi)\big) \quad ; \quad \mathfrak{f}(t,x,\xi) := \mathsf{f}\big(t,x,\Xi^m_\varepsilon(-t,x,\xi)\big) \tag{4.75}$$

The effect of this filtering is to suppress the penalized term, while large amplitude oscillations appear in the coefficients. More precisely, there remains

$$\begin{aligned} \partial_t \mathsf{f} + & \frac{\varepsilon}{\langle \varepsilon\xi\rangle} O(x)\xi\cdot\nabla_x\mathsf{f} + \frac{\varepsilon}{\langle \varepsilon\xi\rangle}(D_\xi\Xi^m_\varepsilon)Q(x,\xi)\cdot\nabla_\xi\mathsf{f} \\ - & \frac{1}{\langle \varepsilon\xi\rangle}(D_\xi\Xi^m_\varepsilon)(\varepsilon D_x\Xi^m_\varepsilon)O(x)\xi\cdot\nabla_\xi\mathsf{f} \\ & = M'(|\xi|)\,\frac{O(x)\xi\cdot\boldsymbol{E}}{|\xi|} + (D_\xi\Xi^m_\varepsilon)[{}^tO(x)\boldsymbol{E} + \nu(\varepsilon\xi)\times{}^tO(x)\boldsymbol{B}]\cdot\nabla_\xi\mathsf{f} \end{aligned} \tag{4.76}$$

where $D_x\Xi^m_\varepsilon$ and $D_\xi\Xi^m_\varepsilon$ must be evaluated at the position t, x and $\Xi^m_\varepsilon(-t,x,\xi)$. Note that the quantities $\varepsilon^k\partial^k_{x_i}\partial^l_{\xi_j}\Xi^m_\varepsilon$ are, for all $(k,l)\in\mathbb{N}^2$, of size one. This means that the coefficients can (and do) oscillate in (t,x) but not in ξ. It follows that we can commute (4.76) with ∂_{ξ_j}. By this way, we can control $\nabla_\xi\mathsf{f}$ (resp. $\nabla_\xi\mathfrak{f}$) in terms of $\varepsilon\nabla_x\mathsf{f}$ (resp. $\varepsilon\nabla_x\mathfrak{f}$).

Another way to proceed is to express (4.74) in cylindrical coordinates, with (ξ_1,ξ_2) as in (4.66) and the direction of ξ_3 as a vertical axis. Since $\partial_\theta = \xi_2\partial_{\xi_1} - \xi_1\partial_{\xi_2}$, the effect is to remove the dependence on ξ in the coefficient (there remains some harmless $\varepsilon\xi$). Then, we can commute the equation with ∂_θ and ∂_r. But this procedure introduces a singularity near the origin ($r=0$), when computing $\nabla_\xi\mathfrak{f}$ in terms of $(\partial_\theta\mathfrak{f},\partial_r\mathfrak{f})$. Therefore, it works only away from $(\xi_1,\xi_2)=(0,0)$.

4.3.3. *Commutator estimates to control the electromagnetic field.* In order to recover estimates of Lipschitz type on the field $(\boldsymbol{E},\boldsymbol{B})(\cdot)$, a good strategy is to commute the VW system with derivatives, and then to apply the procedure of Subsection 4.2.2. From (2.41), we obtain easily

$$\Box_{t,x}\partial_i u = -\partial_i f \tag{4.77}$$

The initial data associated with $\partial_i u$ are like in (2.42), that is

$$\partial_i u(t,x)|_{t=0} = 0, \quad \partial_t(\partial_i u)(t,x)|_{t=0} = 0 \tag{4.78}$$

From (2.45) and (2.46), we can deduce

$$\forall i \in \{0,1,2,3\}, \quad \partial_i \boldsymbol{E} = -\int [\nu(\varepsilon\xi)\partial_t + \nabla_x]\partial_i u d\xi \tag{4.79}$$

$$\forall i \in \{0,1,2,3\}, \quad \partial_i \boldsymbol{B} = \int \nabla_x \times [\partial_i u \nu(\varepsilon\xi)] d\xi \tag{4.80}$$

It is clear that the VW system is very suitable for commutations. Problems can only appear due to the term $\partial_i f$ inside (4.77), which requires to deal with (2.21). For $i \in \{0,1,2,3\}$, we find

$$\begin{aligned} &\partial_t(\partial_i f) + [\nu(\varepsilon\xi)\cdot\nabla_x](\partial_i f) - \frac{1}{\varepsilon^2}[\nu(\varepsilon\xi)\times \mathbf{B}_e(x)]\cdot\nabla_\xi(\partial_i f) \\ &\quad = |\xi|^{-1}M'(|\xi|)\xi\cdot\partial_i\boldsymbol{E} + [\boldsymbol{E}+\nu(\varepsilon\xi)\times\boldsymbol{B})]\cdot\nabla_\xi(\partial_i f) \\ &\qquad + \frac{1}{\varepsilon^2}\nabla_\xi\cdot\big[\nu(\varepsilon\xi)\times\partial_i\mathbf{B}_e(x) f\big] + \nabla_\xi\cdot\big[\big(\partial_i\boldsymbol{E}+\nu(\varepsilon\xi)\times\partial_i\boldsymbol{B}\big)f\big] \end{aligned} \tag{4.81}$$

For $i=0$, the first (singular) term in the last line of (4.81) simply disappears. But, we have to deal with the expression $\partial_t f|_{t=0}$ as determined by (1.6). The multiplication of $\partial_t f$ by ε is crucial to obtain something that is uniformly bounded. With this in mind, introduce

$$\tilde{\mathcal{N}}^1_\varepsilon\big(\boldsymbol{U}(t,\cdot)\big) := \mathcal{N}\big(\varepsilon\partial_t\boldsymbol{U}(t,\cdot)\big) + \mathcal{N}^1_\varepsilon\big(\boldsymbol{U}(t,\cdot)\big) \quad ; \quad \tilde{\mathcal{N}}^1_{\varepsilon,t} := \sup_{0\le s\le t}\tilde{\mathcal{N}}^1_\varepsilon\big(\boldsymbol{U}(s,\cdot)\big) \tag{4.82}$$

For $i \neq 0$, we find that $\partial_i f|_{t=0} = \partial_i f^{in}$ is bounded.

The situation concerning $(\partial_i u, \partial_i f)$ is very similar to that of (u,f) in Subsection 4.2. To control the amplitudes of $\partial_i\boldsymbol{E}$ and $\partial_i\boldsymbol{B}$, we can proceed as in Subsection 4.2.2 with (concerning $\partial_i\boldsymbol{E}$) decompositions like

$$\partial_i\boldsymbol{E} = (\partial_i\boldsymbol{E})^1 + (\partial_i\boldsymbol{E})^2 + (\partial_i\boldsymbol{E})^3 \quad ; \quad (\partial_i\boldsymbol{E})^1 = (\partial_i\boldsymbol{E})^{1,1} + (\partial_i\boldsymbol{E})^{1,2} + (\partial_i\boldsymbol{E})^{1,3} + (\partial_i\boldsymbol{E})^{1,4}$$

where the new part $(\partial_i\boldsymbol{E})^{1,4}$ comes from the additional source terms which are collected in the last line of (4.81). By this way, coming back to (4.46) where $\partial_i f$, $\partial_i\boldsymbol{E}$ and $\partial_i\boldsymbol{B}$ must (except in the nonlinear part) come to replace respectively f, $\boldsymbol{E}$ and $\boldsymbol{B}$, we find

$$\begin{aligned} |\partial_i\boldsymbol{E}(t,x)| \le & Cc(R^\infty)\Big[t\mathcal{N}_0(\partial_i f) + tc(R^\infty)\int_0^t \mathcal{N}_s(\partial_i\boldsymbol{E},\partial_i\boldsymbol{B})ds \\ & + c(R^\infty)(R^\infty)^3\int_0^t \mathcal{N}_s(\partial_i f)ds + tc(R^\infty)(R^\infty)^4\int_0^t\mathcal{N}_s(\partial_i f)ds \\ & + \varepsilon tc(R^\infty)(R^\infty)^3\int_0^t \mathcal{N}_s(\boldsymbol{E},\boldsymbol{B})\mathcal{N}_s(\partial_i f)ds\Big] + |(\partial_i\boldsymbol{E})^{1,4}(t,x)| \end{aligned} \tag{4.83}$$

Now, consider the new contribution $(\partial_i\boldsymbol{E})^{1,4}$. Since all the terms in the last line of (4.81) are in divergence form (in ξ), we can apply Lemma 3.6 to this situation. The same two remarkable properties occur: first, a gain of one power of ε (due to the cold framework); secondly, a gain of one derivative (due to an integration by parts in ξ). Taking into account (4.39), there remains

$$|(\partial_i\boldsymbol{E})^{1,4}(t,x)| \le Ctc(R^\infty)^2(R^\infty)^4\Big[\int_0^t\mathcal{N}_s(f)ds + \varepsilon\int_0^t\mathcal{N}_s(f)\mathcal{N}_s(\partial_i\boldsymbol{E},\partial_i\boldsymbol{B})ds\Big] \tag{4.84}$$

We work with $t \le \min(T_\varepsilon, T)$. Exploiting Proposition 4.6 to control by some uniform constant in the above nonlinear parts the multiplication by $\mathcal{N}_s(\boldsymbol{E},\boldsymbol{B})$ and $\mathcal{N}_s(f)$, we can simply retain

$$\begin{aligned} |\nabla_{t,x}\boldsymbol{E}(t,x)| \le C(t,R^\infty)\Big[&\mathcal{N}_0(\nabla_{t,x}f) + \int_0^t\mathcal{N}_s(f)ds \\ & + \int_0^t\mathcal{N}_s(\nabla_{t,x}f)ds + \int_0^t\mathcal{N}_s(\partial_i\boldsymbol{E},\partial_i\boldsymbol{B})ds\Big] \end{aligned} \tag{4.85}$$

By multiplying (4.85) by ε, we get

$$\tilde{\mathcal{N}}^1_\varepsilon\big((\boldsymbol{E},\boldsymbol{B})(t,\cdot)\big) \leq C(t,R^\infty)\Big[\tilde{\mathcal{N}}^1_{\varepsilon,0} + \int_0^t \tilde{\mathcal{N}}^1_{\varepsilon,s} ds\Big] \tag{4.86}$$

The term $\mathcal{N}_0(\partial_t f)$ inside $\tilde{\mathcal{N}}^1_{\varepsilon,0}$ is not necessarily bounded, but $\mathcal{N}_0(\varepsilon\partial_t f)$ is bounded.

4.3.4. *Weighted Lipschitz estimates for general data.* The Lipschitz regularity is asymptotically preserved in the following sense.

Proposition 4.9. *The Cauchy problem for the MRVM system is locally well-posed with uniform bounds in $\mathcal{X}^1_\varepsilon$ (in the sense of Definition 4.3).*

Proof. There is nothing to do about (4.13), which has been already obtained. Consider (4.12). In view of (4.86), the missing piece is about $\varepsilon\partial_i f$ and $\nabla_\xi f$. The passage from $f(\cdot)$ to $\mathfrak{f}(\cdot)$ up to $\mathsf{f}(\cdot)$ (and vice versa) does not change (modulo a uniform constant) the norms $\mathcal{N}$ and $\mathcal{N}^1_\varepsilon$. Thus, we can work with $\mathsf{f}(\cdot)$, and then come back. The formulation (4.76) is suitable for commutations with the derivatives $\varepsilon\nabla_{t,x}$ and ∇_ξ. Of course, extra terms are produced but always implying $\varepsilon\nabla_x\mathsf{f}$ or $\nabla_\xi\mathsf{f}$. There is no term with ε^{-1} in factor and no term involving $\nabla_x\mathsf{f}$ (without ε in factor). By this way, we get

$$\begin{aligned}\| \varepsilon\nabla_{t,x} f(t,\cdot)\|_{L^\infty_{x,\xi}} + \|\nabla_\xi f(t,\cdot)\|_{L^\infty_{x,\xi}} \leq{}& \|\varepsilon\nabla_{t,x} f(0,\cdot)\|_{L^\infty_{x,\xi}} + \|\nabla_\xi f(0,\cdot)\|_{L^\infty_{x,\xi}} \\ &+ C\int_0^t \tilde{\mathcal{N}}^1_\varepsilon\big(\boldsymbol{U}(s,\cdot)\big)ds + C\int_0^t \tilde{\mathcal{N}}^1_\varepsilon\big(\boldsymbol{U}(s,\cdot)\big)^2 ds\end{aligned} \tag{4.87}$$

The electromagnetic part $(\boldsymbol{E},\boldsymbol{B})$ can be estimated as in (4.86). Add (4.86) and (4.87) to find

$$\tilde{\mathcal{N}}^1_{\varepsilon,t} \leq C\Big[\tilde{\mathcal{N}}^1_{\varepsilon,0} + \int_0^t \tilde{\mathcal{N}}^1_{\varepsilon,s} ds + \int_0^t (\tilde{\mathcal{N}}^1_{\varepsilon,s})^2 ds\Big] \tag{4.88}$$

Restricting T if necessary, by Grönwall's inequalities, we can deduce the important bound (4.12) for the norm $\mathcal{N}^\star \equiv \mathcal{N}^1_\varepsilon \leq \tilde{\mathcal{N}}^1_\varepsilon$. □

4.3.5. *Proof of Theorem 1.1.* From Proposition 4.9, we know that

$$\forall(\varepsilon,\boldsymbol{U}^{in},t) \in]0,1]\times B_\varepsilon \times \big[0,\min(T_\varepsilon,T)\big), \quad \mathcal{N}^1_\varepsilon\big(\mathrm{S}^t_\varepsilon(\boldsymbol{U}^{in})\big) \leq \omega\big(\mathcal{N}^1_\varepsilon(\boldsymbol{U}^{in})\big) \leq C \tag{4.89}$$

If $T \leq T_\varepsilon$, there is nothing to do. Suppose that $T_\varepsilon < T$. Recalling [16, 28], we can extend the smooth solution up to a time $T_\varepsilon + \delta_\varepsilon$ for some $\delta_\varepsilon \in \mathbb{R}^*_+$. The threshold δ_ε may of course depend on ε, α^∞, T and R^{in}. It may even diminish when ε goes to zero or when R^{in} grows. But, these parameters being fixed, it does not change. This is because the Lipschitz norm of $\boldsymbol{U}(\cdot)$ near the *a priori* life span T_ε, which is well controlled by (4.89), allows to determine a minimum threshold for δ_ε. We can even repeat the continuation argument to attein $T_\varepsilon + 2\delta_\varepsilon$, and so on up to T. In other words, the time T gives a lower bound for the life span T_ε of the classical solution. This is clearly in contradiction with the assumption $T_\varepsilon < T$.

4.3.6. *Uniform long time well-posedness for small data.* The determination of T is not only built upon δ^∞_0 and R^{in}, but also on $\delta^\infty_1(\varepsilon)$ with $\delta^\infty_1(\varepsilon)$ as in (4.11). For general bounded families $(B_\varepsilon)_\varepsilon$, there are strong restrictions on the size of T. These restrictions come from the nonlinear term inside (4.88), and also from the condition on T inside (4.60). However, for $\delta^\infty_1(\varepsilon)$ small enough, they can be lifted.

Proposition 4.10. *The Cauchy problem for the MRVM system is for small data uniformly well-posed for a long time in the sense of Definition 4.5.*

Proof. The relation (1.6) implies that

$$\exists C \in \mathbb{R}^*_+; \quad \forall \varepsilon \in]0,1], \quad \tilde{\mathcal{N}}^1_{\varepsilon,0} \leq C\delta^\infty_1(\varepsilon) \tag{4.90}$$

Take $R^\infty = 2R^{in}$, and select any time $T \in \mathbb{R}_+^*$. Then, the coefficients β and γ of (4.51) and (4.52), as well as the constant C inside (4.88), become fixed. Whether this is at the level of (4.49) or (4.88), we have to deal with inequalities like

$$\mathcal{Z}(t) \leq C\delta_1^\infty(\varepsilon) + C\int_0^t \big(\mathcal{Z}(s) + \mathcal{Z}(s)^2\big)ds \tag{4.91}$$

The Grönwall's inequality can be made explicit. It furnishes

$$\forall t \leq \frac{1}{C}\ln\Big(1 + \frac{1}{C\delta_1^\infty(\varepsilon)}\Big), \quad \mathcal{Z}(t) \leq \frac{C\delta_1^\infty(\varepsilon)e^{Ct}}{C\delta_1^\infty(\varepsilon) + 1 - C\delta_1^\infty(\varepsilon)e^{Ct}} \tag{4.92}$$

Adjust δ_1^∞ in such a way that

$$\delta_1^\infty \leq \frac{1}{2C}\frac{1}{e^{CT}-1} \iff T \leq \frac{1}{C}\ln\Big(1 + \frac{1}{2C\delta_1^\infty}\Big) \iff \frac{1}{2} \leq C\delta_1^\infty(\varepsilon) + 1 - C\delta_1^\infty(\varepsilon)e^{Ct}$$

Then, by construction, for $\delta_1^\infty(\varepsilon) \leq \delta_1^\infty$, we have

$$\forall t \leq T, \quad \mathcal{Z}(t) \leq 2C\delta_1^\infty(\varepsilon)e^{Ct}$$

Replacing $\mathcal{Z}$ by $\mathcal{Z} \equiv \mathcal{N}$, we can infer that

$$\forall t \leq T, \quad |\varXi(t)| \leq |\varXi(0)| + 2\delta_1^\infty(\varepsilon)e^{Ct}$$

We can further restrict δ_1^∞ to have $2\delta_1^\infty e^{CT} < R^{in}$, which furnishes (4.48) with $R^\infty = 2R^{in}$. On the other hand, with $\mathcal{Z} \equiv \tilde{\mathcal{N}}_{\varepsilon,t}^1$, we obtain (4.15) together with a control on $[0,T]$ of the weighted Lipschitz norm. This means that $T \leq T_\varepsilon$. Briefly, all the properties inside the paragraphs (i), (ii) and (iii) of Definition 4.5 have been obtained. □

4.3.7. *Access to uniform continuity.* We come back here to the sup norm. Fix some $t \in]0,T]$. One question that arises is whether the family of solution maps $\{\mathrm{S}_t^\varepsilon\}_\varepsilon$ is uniformly continuous with values in the space L^∞.

Proposition 4.11 (uniform continuity)**.** *Let B a bounded set according to $\mathcal{X}_\varepsilon^1$. The family of mappings $\{\mathrm{S}_\varepsilon\}_\varepsilon$ is uniformly continuous on B. More precisely, for T small enough and for all $\delta > 0$, we can find $\eta > 0$ such that, for all couple $(\tilde{\boldsymbol{U}}^{in}, \boldsymbol{U}^{in}) \in B \times B$ of initial data, we have*

$$\forall(\varepsilon,t) \in]0,1] \times [0,T], \quad \| \tilde{\boldsymbol{U}}^{in} - \boldsymbol{U}^{in} \|_{L^\infty_{x,\xi}} \leq \eta \implies \| \mathrm{S}_\varepsilon^t(\tilde{\boldsymbol{U}}^{in}) - \mathrm{S}_\varepsilon^t(\boldsymbol{U}^{in}) \|_{L^\infty_{x,\xi}} \leq \delta \tag{4.93}$$

Proof. Select two $\mathcal{C}^1$-solutions $\boldsymbol{U} = (f, \boldsymbol{E}, \boldsymbol{B})$ and $\tilde{\boldsymbol{U}} = (\tilde{f}, \tilde{\boldsymbol{E}}, \tilde{\boldsymbol{B}})$ of the MRVM system, with corresponding initial data $\boldsymbol{U}^{in}$ and $\tilde{\boldsymbol{U}}^{in}$. Consider the system of equations which is satisfied by the difference $\tilde{\boldsymbol{U}} - \boldsymbol{U}$. This is still a VW system with, at the level of the Vlasov equation, a supplementary source term coming from the non linearities and given by

$$\nabla_\xi \cdot \big[(\boldsymbol{E} + \nu(\varepsilon\xi) \times \boldsymbol{B})(\tilde{f} - f) + \big((\tilde{\boldsymbol{E}} - \boldsymbol{E}) + \nu(\varepsilon\xi) \times (\tilde{\boldsymbol{B}} - \boldsymbol{B})\big)f\big] \tag{4.94}$$

When looking at the Vlasov equation on $(\tilde{f} - f)(\cdot)$, the left hand side of (4.94) can be incorporated in the transport part. The right hand side gives rise to

$$\forall t \in [0,T], \quad \mathcal{N}_t(\tilde{f} - f) \leq \mathcal{N}_0(\tilde{f} - f) + \mathcal{N}_t(\nabla_\xi f)\int_0^t \mathcal{N}_s(\tilde{\boldsymbol{E}} - \boldsymbol{E}, \tilde{\boldsymbol{B}} - \boldsymbol{B})ds \tag{4.95}$$

Exploit Proposition 4.9 to control $\mathcal{N}_t(\nabla_\xi f)$ by a fixed constant. Then, the situation concerning the difference $\tilde{f} - f$ is completely similar to what we had at the level of (4.18). There remains to control $\tilde{\boldsymbol{E}} - \boldsymbol{E}$ and $\tilde{\boldsymbol{B}} - \boldsymbol{B}$. To this end, we can repeat what we did in Subsection 4.2.2. For instance, concerning $\tilde{\boldsymbol{E}} - \boldsymbol{E}$, we still have decompositions like

$$\tilde{\boldsymbol{E}} - \boldsymbol{E} = (\tilde{\boldsymbol{E}} - \boldsymbol{E})^1 + (\tilde{\boldsymbol{E}} - \boldsymbol{E})^2 + (\tilde{\boldsymbol{E}} - \boldsymbol{E})^3$$

$$(\tilde{\boldsymbol{E}} - \boldsymbol{E})^1 = (\tilde{\boldsymbol{E}} - \boldsymbol{E})^{1,1} + (\tilde{\boldsymbol{E}} - \boldsymbol{E})^{1,2} + (\tilde{\boldsymbol{E}} - \boldsymbol{E})^{1,3} + (\tilde{\boldsymbol{E}} - \boldsymbol{E})^{1,4}$$

where the new part $(\tilde{\boldsymbol{E}}-\boldsymbol{E})^{1,4}$ comes from (4.94). Since all the terms inside (4.94) are in divergence form (in ξ), we can again apply Lemma 3.6, which furnishes

$$|(\tilde{\boldsymbol{E}}-\boldsymbol{E})^{1,4}(t,x)| \leq C\varepsilon t(R^\infty)^3 \int_0^t \mathcal{N}_s(\boldsymbol{U})\mathcal{N}_s(\tilde{\boldsymbol{U}}-\boldsymbol{U})ds \tag{4.96}$$

The proof of (4.12) can be readily repeated with $\tilde{\boldsymbol{U}}-\boldsymbol{U}$ in place of $\boldsymbol{U}$. Consideration of (4.96) only induces a change in the definition of $g(\cdot)$ and consequently of $\omega(\cdot)$. We can still find a modulus of continuity $\tilde{\omega}(\cdot)$ such that

$$\forall(\varepsilon,t)\in]0,\varepsilon_0]\times[0,T], \quad \| (\tilde{\boldsymbol{U}}-\boldsymbol{U})(t,\cdot)\|_{\mathcal{X}} \leq \tilde{\omega}\big(\| \tilde{\boldsymbol{U}}^{in}-\boldsymbol{U}^{in}\|_{\mathcal{X}}\big) \tag{4.97}$$

This is enough to deduce (4.93). □

4.3.8. *Lipschitz estimates for prepared data.* As already explained, in general, uniform Lipschitz estimates are not available. The condition (4.9) is necessary. It turns out to be sufficient.

Proposition 4.12. *[uniform bound in Lipschitz norm for prepared data] Let $(B_\varepsilon)_\varepsilon$ be a family that is prepared in the sense of Definition 4.2. Then, there exists a time $T\in\mathbb{R}_+^*$ such that, for all $\varepsilon\in]0,1]$ and for all initial data $\boldsymbol{U}^{in}\in B_\varepsilon$, the MRVM system has a unique solution $\boldsymbol{U}_\varepsilon(\cdot)$ which is defined on $[0,T]$ and which, for some finite constant $C\in\mathbb{R}_+^*$, satisfies*

$$\forall(\varepsilon,\boldsymbol{U}^{in},t)\in]0,1]\times B_\varepsilon\times[0,T], \quad \mathcal{N}_1^1\big(\mathrm{S}_\varepsilon^t(\boldsymbol{U}^{in})\big)\leq \omega\big(\mathcal{N}_1^1(\boldsymbol{U}^{in})\big)\leq C \tag{4.98}$$

Proof. The idea is to reiterate the main lines of what we did before, but there are also subtle and important variations. First and foremost, we have to estimate $\partial_t\boldsymbol{U}(\cdot)$. Consider $\partial_t f(\cdot)$. Coming back to (4.81), we have

$$\| \partial_t f(t,\cdot)\|_{L^\infty_{x,\xi}} \leq \| \partial_t f(0,\cdot)\|_{L^\infty_{x,\xi}} + \mathcal{N}_t(\nabla_\xi f)\int_0^t \mathcal{N}_s(\partial_t\boldsymbol{E},\partial_t\boldsymbol{B})ds \tag{4.99}$$

Knowing (4.9), the initial data $\partial_t f|_{t=0}$ is bounded as expected. In view of Proposition 4.9, the quantity $\mathcal{N}_t(\nabla_\xi f)$ is already bounded. The problems can arise when estimating the spatial derivatives $\nabla_x f(\cdot)$. The equations (2.21) and (4.76) can, in no way, be commuted with $\partial_i\equiv\partial_{x_i}$ when $i\neq 0$ because this would introduce the factor ε^{-1}. Things must be done differently.

The trick is well known in nonlinear geometric optics [25, 26]. It is to work at the level of (4.74). In view of (2.1), the function $\mathrm{b}_e(\cdot)$ does not vanish. Thus, we can divide (4.74) by $\mathrm{b}_e(\cdot)$, apply ∂_i, and then come back. This operation yields

$$\begin{aligned}
\partial_t(\partial_i\mathfrak{f}) &+ \frac{\varepsilon}{\langle\varepsilon\xi\rangle}O(x)\xi\cdot\nabla_x(\partial_i\mathfrak{f}) + \frac{\varepsilon}{\langle\varepsilon\xi\rangle}Q(x,\xi)\cdot\nabla_\xi(\partial_i\mathfrak{f}) - \frac{\mathrm{b}_e(x)}{\varepsilon\langle\varepsilon\xi\rangle}(\xi_2\partial_{\xi_1}-\xi_1\partial_{\xi_2})(\partial_i\mathfrak{f})\\
&= M'(|\xi|)\,\frac{O(x)\xi\cdot\partial_i\boldsymbol{E}}{|\xi|} + [{}^tO(x)\boldsymbol{E}+\nu(\varepsilon\xi)\times{}^tO(x)\boldsymbol{B}]\cdot\nabla_\xi(\partial_i\mathfrak{f})\\
&+ [{}^tO(x)\partial_i\boldsymbol{E}+\nu(\varepsilon\xi)\times{}^tO(x)\partial_i\boldsymbol{B})]\cdot\nabla_\xi\mathfrak{f}\\
&+ \partial_i(\ln\mathrm{b}_e)\Big\{\partial_t\mathfrak{f}+\frac{\varepsilon}{\langle\varepsilon\xi\rangle}O(x)\xi\cdot\nabla_x\mathfrak{f}+\frac{\varepsilon}{\langle\varepsilon\xi\rangle}Q(x,\xi)\cdot\nabla_\xi\mathfrak{f}\\
&\qquad\qquad - M'(|\xi|)\,\frac{O(x)\xi\cdot\boldsymbol{E}}{|\xi|} - [{}^tO(x)\boldsymbol{E}+\nu(\varepsilon\xi)\times{}^tO(x)\boldsymbol{B}]\cdot\nabla_\xi\mathfrak{f}\Big\}\\
&- \frac{\varepsilon}{\langle\varepsilon\xi\rangle}(\partial_iO)(x)\xi\cdot\nabla_x\mathfrak{f} - \frac{\varepsilon}{\langle\varepsilon\xi\rangle}(\partial_iQ)(x,\xi)\cdot\nabla_\xi\mathfrak{f}\\
&+ M'(|\xi|)\,\frac{(\partial_iO)(x)\xi\cdot\boldsymbol{E}}{|\xi|} + [{}^t(\partial_iO)(x)(\boldsymbol{E}+\nu(\varepsilon\xi)\times{}^t(\partial_iO)(x)\boldsymbol{B})]\cdot\nabla_\xi\mathfrak{f}
\end{aligned} \tag{4.100}$$

The remarkable point is that (4.100) involves only $\varepsilon\nabla_x\mathfrak{f}$, $\nabla_\xi\mathfrak{f}$ and $\partial_t\mathfrak{f}$, but not $\nabla_x\mathfrak{f}$. Now, we can follow the bicharacteristics to find

$$\begin{aligned}\| \partial_i f(t,\cdot) \|_{L^\infty_{x,\xi}} \leq \| \partial_i f(0,\cdot) \|_{L^\infty_{x,\xi}} &+ C \int_0^t \mathcal{N}_s(\partial_t\mathfrak{f}) ds \\ &+ C \int_0^t \mathcal{N}_s(\varepsilon\nabla_x\mathfrak{f}) ds + C \int_0^t \mathcal{N}_s(\nabla_\xi\mathfrak{f}) ds \\ &+ C\big(1+\mathcal{N}_t(\nabla_\xi f)\big) \int_0^t \big[\mathcal{N}_s(\boldsymbol{E},\boldsymbol{B}) + \mathcal{N}_s(\partial_i\boldsymbol{E},\partial_i\boldsymbol{B})\big] ds\end{aligned} \tag{4.101}$$

On the other hand, we can exploit the argument **(b)** of Paragraph 4.3.2 to estimate $\nabla_\xi\mathfrak{f}$. Recall the definition (4.82) of $\tilde{\mathcal{N}}_1^1$. Combine (4.85) with (4.99), (4.101) and this remark to find that

$$\tilde{\mathcal{N}}_1^1\big(\boldsymbol{U}(t,\cdot)\big) \leq \tilde{\mathcal{N}}_1^1\big(\boldsymbol{U}(0,\cdot)\big) + C + C\int_0^t \tilde{\mathcal{N}}_1^1\big(\boldsymbol{U}(s,\cdot)\big) ds \tag{4.102}$$

The condition (4.9) is designed to ensure that $\tilde{\mathcal{N}}_1^1\big(\boldsymbol{U}(0,\cdot)\big)$ remains uniformly bounded for the prepared data under consideration. This also holds true concerning $\tilde{\mathcal{N}}_1^1\big(\boldsymbol{U}(t,\cdot)\big)$ by Grönwall's inequality, and therefore concerning $\mathcal{N}_1^1\big(\boldsymbol{U}(t,\cdot)\big)$ □

References

[1] K. Asano and S. Ukai. On the Vlasov-Poisson limit of the Vlasov-Maxwell equation. In *Patterns and waves*, volume 18 of *Stud. Math. Appl.*, pages 369–383. North-Holland, Amsterdam, 1986.

[2] A. Babin, A. Mahalov, and B. Nicolaenko. Regularity and integrability of 3D Euler and Navier-Stokes equations for rotating fluids. *Asymptot. Anal.*, 15(2):103–150, 1997.

[3] J. Benameur, S. Ibrahim, and M. Majdoub. Asymptotic study of a magneto-hydrodynamic system. *Differential Integral Equations*, 18(3):299–324, 2005.

[4] F. Bouchut, F. Golse, and C. Pallard. Classical solutions and the Glassey-Strauss theorem for the 3D Vlasov-Maxwell system. *Arch. Ration. Mech. Anal.*, 170(1):1–15, 2003.

[5] 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.

[6] C. Cheverry. Can One Hear Whistler Waves? *Comm. Math. Phys.*, 338(2):641–703, 2015.

[7] C. Cheverry. Anomalous transport. *J. Differential Equations*, 262(3):2987–3033, 2017.

[8] C. Cheverry. Mathematical perspectives in plasma turbulence. *Research and Reports on Mathematics*, 2:2, 2018.

[9] C. Cheverry and A. Fontaine. Dispersion relations in cold magnetized plasmas. *Kinet. Relat. Models*, 10(2):373–421, 2017.

[10] C. Cheverry and A. Fontaine. Dispersion relations in hot magnetized plasmas. *Journal of Mathematical Analysis and Applications*, 2018.

[11] P. Degond and F. Filbet. On the asymptotic limit of the three dimensional vlasov-poisson system for large magnetic field: formal derivation. *Journal of Statistical Physics*, 165:765–784, 2016.

[12] I. Gallagher and L. Saint-Raymond. On pressureless gases driven by a strong inhomogeneous magnetic field. *SIAM J. Math. Anal.*, 36(4):1159–1176 (electronic), 2005.

[13] P. Gérard. Nonlinear Schrödinger equations in inhomogeneous media: wellposedness and illposedness of the Cauchy problem. In *International Congress of Mathematicians. Vol. III*, pages 157–182. Eur. Math. Soc., Zürich, 2006.

[14] R. T. Glassey. *The Cauchy problem in kinetic theory.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996.

[15] 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.

[16] R. T. Glassey and W. A. Strauss. Singularity formation in a collisionless plasma could occur only at high velocities. *Arch. Rational Mech. Anal.*, 92(1):59–90, 1986.

[17] R. T. Glassey and W. A. Strauss. Absence of shocks in an initially dilute collisionless plasma. *Comm. Math. Phys.*, 113(113):191–208, 1987.

[18] F. Golse. *Distributions, analyse de Fourier, équations aux dérivées partielles.* Cours de l'École Polytechnique, 2012.

[19] S. Ibrahim and T. Yoneda. Long-time solvability of the Navier-Stokes-Boussinesq equations with almost periodic initial large data. *J. Math. Sci. Univ. Tokyo*, 20(1):1–25, 2013.

[20] P.-Y. Jeanne. Optique géométrique pour des systèmes semi-linéaires avec invariance de jauge. (Geometric optics for gauge invariant semilinear systems). *Mém. Soc. Math. Fr., Nouv. Sér.*, 90:vi + 160, 2002.

[21] S. Klainerman. Uniform decay estimates and the Lorentz invariance of the classical wave equation. *Commun. Pure Appl. Math.*, 38:321–332, 1985.

[22] D. Lannes. Space time resonances [after Germain, Masmoudi, Shatah]. In *Séminaire Bourbaki. Volume 2011/2012. Exposés 1043–1058*, pages 355–388, ex. Paris: Société Mathématique de France (SMF), 2013.

[23] J. Lasalle. Uniqueness theorems and successive approximations. *Ann. of Math.*, 50:722–730, 1949.

[24] J. Luk and R. M. Strain. A new continuation criterion for the relativistic Vlasov-Maxwell system. *Commun. Math. Phys.*, 331(3):1005–1027, 2014.

[25] G. Métivier. The mathematics of nonlinear optics. *Handbook of Differential Equations: Evolutionary Equations*, 5:169–313, 2009.

[26] G. Métivier and S. Schochet. The incompressible limit of the non-isentropic Euler equations. *Arch. Ration. Mech. Anal.*, 158(1):61–90, 2001.

[27] C. Pallard. Smooth solutions to the relativistic Vlasov-Maxwell system. *Commun. Math. Sci.*, (suppl. 1):87–94, 2007.

[28] G. Rein. Generic global solutions of the relativistic Vlasov-Maxwell system of plasma physics. *Comm. Math. Phys.*, 135(1):41–78, 1990.

[29] G. Rein. Global weak solutions to the relativistic Vlasov-Maxwell system revisited. *Commun. Math. Sci.*, 2(2):145–158, 2004.

[30] J. Schaeffer. The classical limit of the relativistic Vlasov-Maxwell system. *Comm. Math. Phys.*, 104(3):403–421, 1986.

[31] R. Sospedra-Alfonso and R. Illner. Classical solvability of the relativistic vlasov-maxwell system with bounded spatial density. *Math. Meth. Appl. Sci.*, 33:751–757, 2010.

(Christophe Cheverry) Institut Mathématique de Rennes, Campus de Beaulieu, 263 avenue du Général Leclerc CS 74205 35042 Rennes Cedex, FRANCE

(Slim Ibrahim) Department of Mathematics and Statistics, University of Victoria 3800 Finnerty Road, Victoria, B.C., V8P 5C2, CANADA