Manuscript submitted to
AIMS' Journals
Volume **00**, Number **0**, Xxxx **XXXX**

doi:10.3934/xx.xxxxxxx

pp. **000–000**

AVERAGING OF HIGHLY-OSCILLATORY TRANSPORT EQUATIONS

Philippe Chartier* and Nicolas Crouseilles
Univ Rennes, Inria, Mingus Team,
Campus de Beaulieu, F-35042 Rennes, France.

Mohammed Lemou
Univ Rennes, CNRS, Mingus Team,
Campus de Beaulieu, F-35042 Rennes, France.

Florian Méhats
Univ Rennes, Mingus Team,
Campus de Beaulieu, F-35042 Rennes, France.

(Communicated by the associate editor name)

Abstract. In this paper, we develop a new strategy aimed at obtaining high-order asymptotic models for transport equations with highly-oscillatory solutions. The technique relies upon recent developments averaging theory for ordinary differential equations, in particular normal form expansions in the vanishing parameter. Noteworthy, the result we state here also allows for the complete recovery of the exact solution from the asymptotic model. This is done by solving a companion transport equation that stems naturally from the change of variables underlying high-order averaging. Eventually, we apply our technique to the Vlasov equation with external electric and magnetic fields. Both constant and non-constant magnetic fields are envisaged, and asymptotic models already documented in the literature and re-derived using our methodology. In addition, it is shown how to obtain new high-order asymptotic models.

1. Introduction.

In a large variety of situations, one is confronted to the resolution of a family of transport equations of the form

$$\partial_t f(t,y) + F^\varepsilon(y)\cdot\nabla_y f(t,y) = 0, \quad f(0,y) = f_0(y)\in\mathbb{R}, \quad t\in\mathbb{R}, \quad y\in\mathbb{R}^n, \tag{1}$$

indexed by a small positive parameter ε, whose occurrence in *real-life* models often lies at the core of numerous theoretical and numerical difficulties encountered in obtaining a(-n) (approximate-) solution. The nature of the difficulties (both theoretical and numerical) triggered by the presence of ε may vary according to the form of the vector field $y\mapsto F^\varepsilon(y)\in\mathbb{R}^n$. In this article, we shall address the highly-oscillatory situation where it can be split into two parts

$$F^\varepsilon(y) = \frac{1}{\varepsilon}\omega(y)\,G(y) + K(y) \tag{2}$$

2020 *Mathematics Subject Classification.* Primary: 34C29, 82B40, 35Q83.
Key words and phrases. Averaging, formal series, normal form, transport equation, highly-oscillatory regime, Vlasov equation, strong magnetic field.
* Corresponding author: Philippe Chartier.

where the flow $(t, y_0) \mapsto \Phi_t(y_0)$ associated with the differential equation

$$\dot{y}(t) = G(y(t)), \quad y(0) = y_0, \tag{3}$$

is assumed to be **periodic**, regardless of the specific trajectory (i.e. independently of the initial condition y_0 at time $t = 0$) and where $y \mapsto \omega(y)$ is a scalar function *bounded from below* by a positive constant. Owing to the $1/\varepsilon$-term in front of the vector field G, the solution of the transport equation evolves in a *highly-oscillatory regime* as soon as ε becomes small, which is specifically the regime under investigation here. Since our ultimate goal is the design of *high-order uniformly accurate* numerical methods (i.e. methods whose computational cost and accuracy are not influenced by the value of ε), the identification of the asymptotic models is a pre-requisite: this is the task addressed in this work.

Examples of highly-oscillatory equations of the form (1) are numerous [2, 3, 4, 5, 13, 14, 15, 16]. It is obviously out of the scope of this introductory paper to treat all of them: we will rather concentrate on the following model that will constitute hereafter our *target application*, namely the **Vlasov equation with strong magnetic field**

$$\partial_t f(t,x,v) + v \cdot \nabla_x f(t,x,v) + \left(E(x) + \frac{1}{\varepsilon} v \times B(x)\right) \cdot \nabla_v f(t,x,v) = 0, \tag{4}$$

where $x \in \mathbb{R}^3$ and $v \in \mathbb{R}^3$ denote respectively the spatial and velocity variables, $f : \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \mathbb{R}$ is the distribution function, i.e. the density of particles at time t, position x and velocity v, and where $E : \mathbb{R}^3 \mapsto \mathbb{R}^3$ and $B : \mathbb{R}^3 \mapsto \mathbb{R}^3$ are respectively the electric and magnetic fields, assumed to be *external* at this stage (i.e. not coupled with f through Maxwell equations for instance).

Our first objective is this paper is to derive *formal* asymptotic models for equation (1) with F^ε satisfying (2) and $\omega \equiv 1$. Rather then merely obtain the limit equation where ε tends to zero, we demand higher-order terms in powers of ε. The methodology we propose relies on recent results from the theory of averaging for highly-oscillatory ordinary differential equations [20, 21], and more precisely on normal forms obtained as ε-expansions. Such series have been derived with the help of B-series in [6, 7, 9] or somehow more simply in [17, 18, 19] with word-series[1]. The underlying results we shall lean onto will be presented in Section 3, but prior to that, we shall show in Section 2 how the splitting of the vector field F^ε into two *commuting* vector fields naturally leads to *two independent transport equations*[2]. The corresponding first result (for constant ω) will be stated in Section 4.

In Section 5, we will address the much more involved situation of a varying frequency (ω non-constant in (2)), which requires to work in an augmented space. In particular, the main result of this paper will be stated there. It allows to rewrite the original transport equation (1) as a set of four *non-stiff* equations for a *phase function (S)* and a *profile function (h)*. This procedure is inspired from the recent work [11], although the context here is different. The two equations for the profile function are the counterpart of the averaged equation obtained elsewhere in the literature. However, solving the equation for the phase function S allows to recover

[1] Although the effect of truncating the aforementioned formal series has been fully analysed in subsequent papers [7, 8], it is out of the scope of this first paper to present a complete error analysis.

[2] The aim of this section is to introduce the rationale underlying our methodology, i.e. the idea that decomposing the vector field F^ε in (1) into two commuting vector fields allows to separate the stiff and non-stiff parts of the transport equation.

exactly the complete solution of (1). This part is up to our knowledge completely new. Since we use series-expansions, it is possible to write down explicitly and in a systematic way the terms appearing in the four equations for S and h. In Section 6, we shall eventually envisage our target application (4) and show how to obtain the terms of these developments. Firstly, in Section 6.1, we will consider the case of a constant magnetic field $B(x) \equiv B$ in (4) with a physical space of dimension two, as it appears to be a simple application of the results of Section 4. Secondly, in Section 6.2, we will address the more involved situation of a varying magnetic field (B non-constant in (4)), which requires a preliminary treatment of the transport equation, as exposed in Section 5. At last, we shall treat equation (4) in full generality, i.e. in three dimensions and with a general magnetic field, and compare the equations we obtain with our methodology to results previously published in the literature.

2. **Decomposition of a transport equation.** Let us consider the Liouville equation

$$\partial_t f(t,y) + F(y)\cdot\nabla_y f(t,y) = 0,$$

associated to a split vector field of the form

$$F = F_1 + F_2,$$

and let us make the fundamental assumption that the Lie bracket of F_1 and F_2 vanishes, that is to say that

$$\forall y \in \mathbb{R}^n, \quad [F_1,F_2](y) := (\partial_y F_1)(y)\; F_2(y) - (\partial_y F_2)(y)\; F_1(y) = 0.$$

This *commutation* of vector fields further manifests itself as the commutation of the two flows[3] associated with F_1 and F_2, or as the commutation of the Lie operators associated with F_1 and F_2. More precisely, denoting $\mathcal{L}_{F_1}$ and $\mathcal{L}_{F_2}$ the operators defined, for any smooth function $g \in \mathcal{C}^\infty(\mathbb{R}^n;\mathbb{R}^m)$ by

$$\forall y \in \mathbb{R}^n, \quad \mathcal{L}_{F_1}(g)(y) = \partial_y g(y)\, F_1(y) \quad \text{and} \quad \forall y \in \mathbb{R}^n, \quad \mathcal{L}_{F_2}(g)(y) = \partial_y g(y)\, F_2(y),$$

we have[4]

$$\mathcal{L}_{F_1}\mathcal{L}_{F_2} = \mathcal{L}_{F_2}\mathcal{L}_{F_1}, \tag{5}$$

i.e. more explicitly

$$\forall g \in \mathcal{C}^\infty(\mathbb{R}^n;\mathbb{R}^m), \quad \mathcal{L}_{F_1}\Big(\mathcal{L}_{F_2}(g)\Big) = \mathcal{L}_{F_2}\Big(\mathcal{L}_{F_1}(g)\Big).$$

The method of characteristics immediately gives for any *smooth* solution of (1)

$$\forall t \in \mathbb{R}, \quad f(t,\cdot) = \exp\left(-t\mathcal{L}_{F_1+F_2}\right)(f_0), \tag{6}$$

which, owing to relation (5), can also be written as

$$\forall t \in \mathbb{R}, \quad f(t,\cdot) = \exp\left(-t\mathcal{L}_{F_1}\right)\exp\left(-t\mathcal{L}_{F_2}\right)(f_0) = \exp\left(-t\mathcal{L}_{F_2}\right)\exp\left(-t\mathcal{L}_{F_1}\right)(f_0). \tag{7}$$

A somehow *natural* step forward now consists in *separating the two times* in previous relation and defining the new function with additional variable τ

$$\tilde{f}(t,\tau,\cdot) = \exp\left(-\tau\mathcal{L}_{F_1}\right)\exp\left(-t\mathcal{L}_{F_2}\right)(f_0) = \exp\left(-t\mathcal{L}_{F_2}\right)\exp\left(-\tau\mathcal{L}_{F_1}\right)(f_0). \tag{8}$$

We are now in position to state the following proposition, which shows that the augmented function $\tilde{f}$ is in fact the unique solution of a system of two **independent** equations.

[3] These flows are assumed to be defined for all $t \in \mathbb{R}$ and all $y \in \mathbb{R}^n$ without further notice.

[4] Owing to the general well-known formula $\mathcal{L}_{F_1}\mathcal{L}_{F_2} - \mathcal{L}_{F_2}\mathcal{L}_{F_1} = \mathcal{L}_{[F_1,F_2]}$.

Proposition 1. *Consider the system composed of the following two transport equations*

$$\forall (t,\tau,y) \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}^n, \quad \partial_\tau \tilde{f}(t,\tau,y) + F_1(y)\cdot\nabla_y \tilde{f}(t,\tau,y) = 0 \tag{9}$$

and

$$\forall (t,\tau,y) \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}^n, \quad \partial_t \tilde{f}(t,\tau,y) + F_2(y)\cdot\nabla_y \tilde{f}(t,\tau,y) = 0, \tag{10}$$

together the with initial condition $\tilde{f}(0,0,y) = f_0(y)$. If the condition $[F_1,F_2] = 0$ is satisfied, this system has a unique solution, which furthermore satisfies

$$\forall (t,y) \in \mathbb{R}\times\mathbb{R}^n, \quad \tilde{f}(t,t,y) = f(t,y).$$

Proof. We first note that, if a solution $\tilde{f}$ exists, then equations (9) and (10) can be solved in any order. Hence, we can obtain the value of $\tilde{f}(t,\tau,y)$ by first solving (9) for $t = 0$ from the initial value $\tilde{f}(0,0,y) = f_0(y)$ -this furnishes $\tilde{f}(0,\tau,y)$- and then by solving (10) for fixed τ from this initial value. Insofar as the solution exists, it is thus unique. Now, define

$$\tilde{f}(t,\tau,\cdot) = \exp(-\tau\mathcal{L}_{F_1})\exp(-t\mathcal{L}_{F_2})(f_0) = \exp(-t\mathcal{L}_{F_2})\exp(-\tau\mathcal{L}_{F_1})(f_0).$$

It is easy to check that it satisfies both (9) and (10) by considering successively the first and the second form. The function $\tilde{f}$ defined above is thus the unique solution of system (9-10). Finally,

$$\begin{aligned}\partial_t\Big(\tilde{f}(t,t,y)\Big) + F\cdot\nabla_y\tilde{f}(t,t,y) &= \partial_t\tilde{f}(t,t,y) + \partial_\tau\tilde{f}(t,t,y) + F\cdot\nabla_y\tilde{f}(t,t,y)\\ &= \partial_t\tilde{f}(t,t,y) + F_1\cdot\nabla_y\tilde{f}(t,t,y) + \partial_\tau\tilde{f}(t,t,y) + F_2\cdot\nabla_y\tilde{f}(t,t,y)\\ &= 0.\end{aligned}$$

The initial condition $\tilde{f}(0,0,\cdot) = f_0$ and a uniqueness argument then allow to conclude. □

3. **Averaging of ordinary differential equations in a nutshell.**

Since our approach for averaging the transport equation (1) consists in averaging first the characteristics and then rewrite the corresponding Liouville equations, we hereafter recall the main results upon which we shall lean. In this paper, we content ourselves with formal expansions, thus neglecting at this stage the occurrence of error terms. This is justified by the fact that, under appropriate smoothness assumptions, these errors actually become of size ε^n for any fixed n, or even exponentially small (i.e. bounded by $Ce^{-C/\varepsilon}$ for some positive constant C). A completely rigorous treatment of these error terms for ordinary differential equations can be found for instance in [7].

3.1. **A normal form theorem.**

Consider the highly-oscillatory differential equation

$$\dot{y} = F^\varepsilon(y) := \frac{1}{\varepsilon}G(y) + K(y) \tag{11}$$

i.e. equation (2) with $\omega \equiv 1$, where both vector fields G and K are assumed to be smooth[5]. As already alluded to in the Introduction section, the fundamental assumption (**H**) required to go any further is that

[5]Either of class C^k or analytic. The precise smoothness assumption determines the type of error bounds, either polynomial or exponential in ε and is thus not essential here.

(H) G generates a **periodic flow** Φ_τ, regardless of the specific trajectory (i.e. with a period which remains independent of the initial value). By convention, we will suppose here that this period is 2π.

Since the Lie bracket of G and K has here no reason to vanish, we can not reproduce right away the analysis conducted in previous section. It is precisely the aim of averaging to rewrite F^ε as the sum of two commuting fields[6]. As already emphasized, this is in general possible only up to small error terms, so that the theorem stated below is to be understood in a formal sense.

Theorem 3.1. *Suppose that the vector field F^ε can be split according to equation (11) and that G satisfies assumption* (**H**). *Then there exist two vector fields G^ε and K^ε such that*

(i) $F^\varepsilon = \frac{1}{\varepsilon}G^\varepsilon + K^\varepsilon$;
(ii) the Lie bracket of G^ε and K^ε vanishes, i.e. $[G^\varepsilon, K^\varepsilon] = 0$;
(iii) the vector field G^ε generates a flow $\tau \mapsto \Phi^\varepsilon_\tau$ which is 2π-periodic, regardless of the specific trajectory, i.e.

$$\forall (t, y) \in \mathbb{R} \times \mathbb{R}^n, \quad \Phi^\varepsilon_{t+2\pi}(y) = \Phi^\varepsilon_t(y).$$

This result brings us back to Section 2 and indeed allows to *split* equation (1) into two equations of the form (9-10); details will be given in Section 4. We conclude this subsection with a few additional statements related to the conservation of geometric properties by stroboscopic averaging.

Theorem 3.2. *Suppose that the vector field F^ε can be split according to equation (11) and that G satisfies assumption* (**H**). *Then the two vector fields G^ε and K^ε of Theorem 3.1 have the following properties:*

(i) if both G and K are divergence-free vector fields, then so are G^ε and K^ε;
(ii) if both G and K are Hamiltonian vector fields, then so are G^ε and K^ε;
(iii) if both G and K are Poisson vector fields with the same structure matrix

$$\mathbb{R}^n \ni y \longmapsto \Omega(y) \in \mathbb{R}^{n\times n}$$

defining a Poisson bracket (i.e. skew symmetric and satisfying Jacobi identity and Leibniz' rule)

$$\{P, Q\} = (\nabla P)^T \Omega \nabla Q,$$

then G^ε and K^ε are also Poisson vector fields with structure matrix Ω.

Remark 1. The properties of Theorem 3.2 are intimately linked to the choice of stroboscopic averaging (see [9, 10]), which is the only averaging procedure preserving geometric properties of the initial vector field F^ε.

3.2. **Expansions in powers of ε of the vector fields G^ε and K^ε.** Since we wish in particular to identify the asymptotic behaviour of (1) in the limit when ε tends to zero as well as higher-order terms in ε, it is essential to consider ε-expansions of the various functions appearing in Theorem 3.1. Since this was precisely the point of view adopted in [9, 10], we shall again quote the following result[7]:

[6] At least, this is one way to envisage averaging for ordinary differential equations and this is the point of view adopted both in [9] and in the recent series of papers by Murua and Sanz-Serna [17, 18, 19].

[7] Note again that an alternative proof of this result may be found in [17] and [18].

Theorem 3.3. *Consider the Fourier series of*

$$K_\tau(y) = \left(\frac{\partial \Phi_\tau}{\partial y}(y)\right)^{-1} \quad (K \circ \Phi_\tau)(y) = \sum_{k \in \mathbb{Z}} e^{ik\tau} \hat{K}_k(y). \tag{12}$$

The averaged *vector field* K^ε *admits the following* **formal** ε*-expansion*

$$K^\varepsilon = \sum_{r=1}^{+\infty} \varepsilon^{r-1} K^{[r]} = \sum_{r=1}^{+\infty} \frac{\varepsilon^{r-1}}{r} \sum_{(i_1,\ldots,i_r) \in \mathbb{Z}^r} \bar{\beta}_{i_1\cdots i_r} \, [\ldots[\hat{K}_{i_1}, \hat{K}_{i_2}], \hat{K}_{i_3}], \ldots, \hat{K}_{i_r}] \tag{13}$$

where the coefficients $\bar{\beta}$ *are universal (problem-independent). Similarly, the vector field* G^ε *admits the following formal* ε*-expansion*

$$G^\varepsilon \quad = \quad \varepsilon(F^\varepsilon - K^\varepsilon). \tag{14}$$

Remark 2. The fact that geometric properties of G^ε and K^ε are inherited from F^ε may also be seen as a direct consequence of the form of previous expansions, which are linear combinations of embedded Lie-brackets of the $\hat{K}_k$'s. For instance, if both G and K are Poisson vector field with structure matrix $\Omega(y)$ then K_τ is of the form

$$K_\tau(y) = \Omega(y) \nabla_y H_\tau(y) \quad \text{with} \quad H_\tau(y) = \sum_{k \in \mathbb{Z}} e^{i\tau} \hat{H}_k(y)$$

and all Fourier coefficients $\hat{K}_k(y) = \Omega(y) \nabla_y \hat{H}_k(y)$ are also Poisson vector fields for the structure matrix $\Omega(y)$. Since

$$\forall (k,l) \in \mathbb{Z}^2, \quad [\hat{K}_k, \hat{K}_l] = \Omega(y) \nabla_y \{\hat{H}_k, \hat{H}_l\}$$

where $\{\cdot,\cdot\}$ denotes the Poisson bracket operation, it is then immediate to see that both G^ε and K^ε are Poisson vector fields with Hamiltonians given by formulas (13) and (14) where Lie brackets are replaced by Poisson brackets and the $\hat{K}_k$'s by the $\hat{H}_k$'s. Similarly, if $\mathit{div}(G) = \mathit{div}(K) = 0$, then $\mathit{div}(\hat{K}_k) = 0$ for all $k \in \mathbb{Z}$ and a standard computation shows that

$$\forall (k,l) \in \mathbb{Z}^2, \quad \mathit{div}\left([\hat{K}_k, \hat{K}_l]\right) = 0$$

so that again both G^ε and K^ε are divergence-free.

In order to be able to derive the expansions of G^ε and K^ε, it still remains to give the value of the coefficients $\bar{\beta}$ appearing in formula (13). This is the purpose of next proposition.

Proposition 2. *The coefficients* $\bar{\beta}$ *can be computed recursively from the following formulas, which hold for all values of* $j \in \mathbb{Z}^*$, $r, s \in \mathbb{N}^*$ *and* $(l_1, \ldots, l_s) \in \mathbb{Z}^s$*:*

$$\begin{array}{lllllll}
\bar{\beta}_0 & = & 1, & \bar{\beta}_j & = & 0, \\
\bar{\beta}_{0^{r+1}} & = & 0, & \bar{\beta}_{0^r j} & = & \frac{i}{j}\left(\bar{\beta}_{0^{r-1}j} - \bar{\beta}_{0^r}\right), \\
\bar{\beta}_{jl_1\cdots l_s} & = & \frac{i}{j}\left(\bar{\beta}_{l_1\cdots l_s} - \bar{\beta}_{(j+l_1)l_2\cdots l_s}\right), & \bar{\beta}_{0^r jl_1\cdots l_s} & = & \frac{i}{j}\left(\bar{\beta}_{0^{r-1}jl_1\cdots l_s} - \bar{\beta}_{0^r(j+l_1)l_2\cdots l_s}\right).
\end{array}$$

For the sake of illustration and later use, we now give the first terms of $K^\varepsilon = K^{[1]} + \varepsilon K^{[2]} + \varepsilon^2 K^{[3]} + \mathcal{O}(\varepsilon^3)$, as stated in [6]:

$$\begin{aligned}
K^{[1]} &= \hat{K}_0,\\
K^{[2]} &= \sum_{k>0} \frac{i}{k}\left([\hat{K}_k, \hat{K}_{-k}] + [\hat{K}_0, \hat{K}_k - \hat{K}_{-k}]\right),\\
K^{[3]} &= \sum_{k\neq 0} \frac{1}{k^2}\left([[\hat{K}_k, \hat{K}_0], \hat{K}_0] + [[\hat{K}_{-k}, \hat{K}_k], \hat{K}_k] - \frac{1}{2}[[\hat{K}_{-2k}, \hat{K}_k], \hat{K}_k] + [[\hat{K}_0, \hat{K}_k], \hat{K}_{-k}]\right)\\
&\quad - \sum_{0\neq m\neq -l\neq 0} \frac{1}{l(m+l)}[[\hat{K}_0, \hat{K}_l], \hat{K}_m] + \sum_{k<-|l|} \frac{1}{lk}[[\hat{K}_k, \hat{K}_l], \hat{K}_{-l}]\\
&\quad - \sum_{0>k<m, m+k\neq 0} \frac{1}{km}[[\hat{K}_k, \hat{K}_{-k}], \hat{K}_m]\\
&\quad - \sum_{0\neq m\neq \pm l\neq 0, m>-m-l<l} \frac{1}{m(m+l)}[[\hat{K}_{-m-l}, \hat{K}_l], \hat{K}_m].
\end{aligned} \tag{15}$$

Remark 3. The following expressions of the first three terms of the averaged equation have also been derived in various places and do not use Fourier coefficients:

$$\begin{aligned}
K^{[1]}(y) &= \frac{1}{2\pi}\int_0^{2\pi} K_\tau(y)d\tau, \quad K^{[2]}(y) = \frac{-1}{4\pi}\int_0^{2\pi}\int_0^\tau [K_s(y), K_\tau(y)]dsd\tau,\\
K^{[3]}(y) &= \frac{1}{8\pi}\int_0^{2\pi}\int_0^\tau\int_0^s [[K_r(y), K_s(y)], K_\tau(y)]drdsd\tau\\
&\qquad + \frac{1}{24\pi}\int_0^{2\pi}\int_0^\tau\int_0^\tau [K_r(y), [K_s(y), K_\tau(y)]]drdsd\tau.
\end{aligned}$$

Further terms can be formally obtained by using a non-linear Magnus expansion [1]. Each of these is a linear combination of iterated integrals of iterated brackets of K_τ.

As an illustration, we derive below the expressions of G^ε and K^ε for a simple example. We thus consider the following vector field

$$F^\varepsilon(y) = \begin{pmatrix} v \\ \frac{1}{\varepsilon}Jv + E \end{pmatrix} \tag{16}$$

where $x = (x_1, x_2)^T \in \mathbb{R}^2$, $v = (v_1, v_2)^T \in \mathbb{R}^2$, $y = (x_1, x_2, v_1, v_2)^T \in \mathbb{R}^4$, $E \in \mathbb{R}^2$ and

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

The function F^ε may be decomposed into the sum $\frac{1}{\varepsilon}G + K$ with

$$G(y) = \begin{pmatrix} 0 \\ Jv \end{pmatrix} \quad \text{and} \quad K(y) = \begin{pmatrix} v \\ E \end{pmatrix},$$

and the flow Φ_τ associated with G simply reads

$$\Phi_\tau(y) = \begin{pmatrix} x \\ e^{\tau J}v \end{pmatrix}.$$

Substituting Φ_τ into K then leads to

$$K_\tau(y) = \begin{pmatrix} e^{\tau J}v \\ e^{-\tau J}E \end{pmatrix} = e^{i\tau}\hat{K}_1(y) + e^{-i\tau}\hat{K}_{-1}(y)$$

with

$$\hat{K}_1(y) = \frac{1}{2}\begin{pmatrix} v - iJv \\ E + iJE \end{pmatrix} \quad \text{and} \quad \hat{K}_{-1}(y) = \frac{1}{2}\begin{pmatrix} v + iJv \\ E - iJE \end{pmatrix},$$

where we have used the relation $e^{\theta J} = (\cos\theta)I + (\sin\theta)J$ and have written $\cos\theta = \frac{1}{2}(e^{i\theta} + e^{-i\theta})$ and $\sin\theta = \frac{1}{2i}(e^{i\theta} - e^{-i\theta})$. Formula (13) then gives

$$\begin{aligned} K^{[1]} &= \hat{K}_0 = 0, \\ K^{[2]} &= i[\hat{K}_1, \hat{K}_{-1}] = -2\Im\left((\partial_y \hat{K}_1)\hat{K}_{-1}\right) = \begin{pmatrix} JE \\ 0 \end{pmatrix} \end{aligned}$$

and all other $K^{[r]}$ for $r \geq 3$ vanish, as can be checked by easy calculations.

4. **Averaging of transport equations with constant frequency.** Compiling the arguments of the two previous sections, it is now straightforward to obtain the following corollary, which establishes in particular the existence of a formal averaged transport equation for problems of the form (1-2).

Corollary 1. *Let $F^\varepsilon = \frac{1}{\varepsilon}G^\varepsilon + K^\varepsilon$ be the normal form splitting of a highly-oscillating vector field $F^\varepsilon = \frac{1}{\varepsilon}G + K$ satisfying* (**H**). *The solution of the transport equation*

$$\partial_t f(t,y) + F^\varepsilon(y)\cdot\nabla_y f(t,y) = 0$$

may be obtained as the diagonal value (i.e. for the value $\tau = t/\varepsilon$) of the two-scale function $\tilde{f}(t,\tau,y)$, 2π-periodic in τ, defined as the unique solution of the following system of two equations

$$\begin{cases} \forall(t,\tau,y), & \partial_\tau \tilde{f}(t,\tau,y) + G^\varepsilon(y)\cdot\nabla_y \tilde{f}(t,\tau,y) = 0, \quad (i) \\ \forall(t,\tau,y), & \partial_t \tilde{f}(t,\tau,y) + K^\varepsilon(y)\cdot\nabla_y \tilde{f}(t,\tau,y) = 0 \quad (ii) \end{cases}$$

with initial condition $\tilde{f}(0,0,\cdot) = f_0$. Moreover, the ε-expansions of G^ε and K^ε are given by formulas (13-14) of Theorem 3.3. If in addition G and K are both divergence-free, then so are G^ε and K^ε, and similarly, if G and K are both Hamiltonian, then so are G^ε and K^ε, with Hamiltonians that can be obtained again from formulas (13-14) by replacing Lie brackets by Poisson brackets.

Proof. The result follows immediately from Proposition 1 with $F_1 = \frac{1}{\varepsilon}G^\varepsilon$ and $F_2 = K^\varepsilon$ and from Theorem 3.3. $\square$

Remark 4. Equation (ii) is usually referred to as the *averaged transport equation.*

As a straightforward illustration of this corollary, we consider the simplified case of a set of particles evolving in a constant electric field (independent of time and phase-space variables) and submitted to a constant magnetic field. The corresponding equation

$$\partial_t f + v\cdot\nabla_x f + \left(\frac{1}{\varepsilon}Jv + E\right)\cdot\nabla_v f = 0, \tag{17}$$

-where f depends on time $t \in \mathbb{R}$, position $x \in \mathbb{R}^2$ and velocity $v \in \mathbb{R}^2$- is obviously of the form (1) with $y = (x_1, x_2, v_1, v_2)^T \in \mathbb{R}^4$ and F^ε given by (16). On the one hand, given the extreme simplicity of the vector field F^ε, the solution $f(t,x,v)$ of (17) can be directly written as

$$f_0\left(x - \varepsilon J(e^{\tau J} - I)v - \varepsilon^2(e^{\tau J}E - E) + \varepsilon t JE, e^{\tau J}v - \varepsilon J(e^{\tau J} - I)E\right) \tag{18}$$

for $\tau = t/\varepsilon$. On the other hand, using the computations at the end of previous section, equations (i) and (ii) of Corollary 1 for $\tilde{f}(t,\tau,x,v)$ take the following form

$$(i)\ \partial_\tau \tilde{f} + (\varepsilon v - \varepsilon^2 JE)\cdot\nabla_x \tilde{f} + (Jv + \varepsilon E)\cdot\nabla_v \tilde{f} = 0, \quad (ii)\ \partial_t \tilde{f} + \varepsilon JE\cdot\nabla_x \tilde{f} = 0.$$

By direct differentiation w.r.t. τ and then t, it can be checked that the function given in formula (18) satisfies both equations (i) and (ii).

5. **High-oscillations with varying frequency.** In this section, we again consider the transport equation

$$\partial_t f(t,y) + F^\varepsilon(y)\cdot\nabla_y f(t,y) = 0 \tag{19}$$

where the vector field F^ε is now of the form

$$F^\varepsilon(y) = \frac{1}{\varepsilon}\omega(y)G(y) + K(y) \tag{20}$$

with G still generating a 2π-periodic flow Φ_τ, independently of the initial condition. In this form, Theorem 1 does not directly apply, owing to the non-existence of a common frequency for all trajectories (if ω varies). In order to rewrite (19) in a more amenable form, we thus divide it by ω

$$\frac{1}{\omega(y)}\partial_t f(t,y) + \frac{1}{\omega(y)}F^\varepsilon(y)\cdot\nabla_y f(t,y) = 0. \tag{21}$$

Upon denoting $Y = (t,y)$, previous equation may then be rewritten as $\mathcal{L}_{\check{F}^\varepsilon}(f) = 0$, where

$$\mathcal{L}_{\check{F}^\varepsilon}(f) = (\partial_Y f)\ \check{F}^\varepsilon \tag{22}$$

is the Lie derivative of f in the direction of the **augmented** vector field

$$\begin{aligned}\check{F}^\varepsilon(Y) &= \begin{pmatrix} \frac{1}{\omega(y)} \\ \frac{1}{\omega(y)}F^\varepsilon(y)\end{pmatrix} = \frac{1}{\varepsilon}\begin{pmatrix} 0 \\ G(y)\end{pmatrix} + \begin{pmatrix} \frac{1}{\omega(y)} \\ \frac{1}{\omega(y)}K(y)\end{pmatrix} \\ &:= \frac{1}{\varepsilon}\check{G}(Y) + \check{K}(Y). \end{aligned} \tag{23}$$

In particular, note that $\check{G}$ still generates a 2π-periodic flow.

5.1. **Immersion as the stationary solution of an extended equation.** Our first idea is to interpret the function $f(t,y) = f(Y)$ as the (stationary) solution to the following augmented transport equation on $g(s,Y)$:

$$\partial_s g(s,Y) + \check{F}^\varepsilon(Y)\cdot\nabla_Y g(s,Y) = 0, \quad g(0,Y) = f(Y) = f(t,y).$$

This means that

$$g(s,\cdot) = \exp\left(-s\mathcal{L}_{\check{F}^\varepsilon}\right) f = f \quad \text{for all} \quad s \geq 0.$$

Since $\check{G}$ generates a 2π-periodic flow, the averaging Theorem 3.1 ensures that

$$\check{F}^\varepsilon = \frac{1}{\varepsilon}\check{G}^\varepsilon + \check{K}^\varepsilon,$$

where $\check{G}^\varepsilon$ still generates a 2π-periodic flow and $[\check{G}^\varepsilon, \check{K}^\varepsilon] = 0$. Proceeding as in Section 2, we then get two equations for

$$\tilde{g}(s,\tau,\cdot) = \exp\left(-\tau\mathcal{L}_{\check{G}^\varepsilon}\right)\exp\left(-s\mathcal{L}_{\check{K}^\varepsilon}\right) f$$

of the form

$$(i)\quad \partial_s \tilde g(s,\tau,Y) + \check K^\varepsilon(Y)\cdot\nabla_Y \tilde g(s,\tau,Y) = 0 \tag{24}$$

$$(ii)\quad \partial_\tau \tilde g(s,\tau,Y) + \check G^\varepsilon(Y)\cdot\nabla_Y \tilde g(s,\tau,Y) = 0 \tag{25}$$

which can be solved one after another in any order, since $[\check G^\varepsilon, \check K^\varepsilon] = 0$. Note the usual relation $\tilde g(s, s/\varepsilon, Y) = g(s, Y) = f(Y)$. However, there is here no known initial condition at $s = \tau = 0$, since $\tilde g(0,0,Y) = g(0,Y) = f(t,y)$ is precisely the unknown of the original problem.

5.2. **Eliminating the extra-variable** s**.** Our objective in this subsection is to transform the two equations (24-25) into new equations which do not involve the variable s and are provided with a proper initial condition, namely $f_0(y)$. We will then show how to recover the original solution $f(t,y)$ using only these new equations. To this aim, we will introduce a phase-function $(t,\tau,y)\mapsto S(t,\tau,y)$ in the spirit of [11], which will be defined later on as the solution of a transport equation, and a profile-function $(t,\tau,y)\mapsto h(t,\tau,y)$ defined by

$$h(t,\tau,y) = \tilde g(S(t,\tau,y),\tau,t,y), \tag{26}$$

that will also be shown to satisfy a companion transport equation. Our starting point is the following set of relations

$$\begin{aligned}
\partial_t h &= (\partial_s \tilde g(S,\tau,t,y))\,\partial_t S + \partial_t \tilde g(S,\tau,t,y),\\
\partial_\tau h &= (\partial_s \tilde g(S,\tau,t,y))\,\partial_\tau S + \partial_\tau \tilde g(S,\tau,t,y),\\
\partial_y h &= (\partial_s \tilde g(S,\tau,t,y))\,\partial_y S + \partial_y \tilde g(S,\tau,t,y),
\end{aligned}$$

where we have omitted the obvious arguments of functions h and S and which may be straightforwardly obtained. Together with equations (24) and (25), they lead immediately to

$$\begin{aligned}
&\check K_1^\varepsilon(y)\partial_t h(t,\tau,y) + \check K_2^\varepsilon(y)\cdot\nabla_y h(t,\tau,y)\\
&= (\partial_s \tilde g(S,\tau,t,y))\left(\check K_1^\varepsilon(y)\partial_t S(t,\tau,y) + \check K_2^\varepsilon(y)\partial_y S(t,\tau,y) - 1\right),
\end{aligned}$$

and

$$\begin{aligned}
&\partial_\tau h(t,\tau,y) + \check G_1^\varepsilon(y)\partial_t h(t,\tau,y) + \check G_2^\varepsilon(y)\cdot\nabla_y h(t,\tau,y)\\
&= (\partial_s \tilde g(S(t,\tau,y),\tau,t,y))\left(\check G_1^\varepsilon(y)\partial_t S(t,\tau,y) + \check G_2^\varepsilon(y)\partial_y S(t,\tau,y) + \partial_\tau S(t,\tau,y)\right),
\end{aligned}$$

where the index 1 in $\check K_1^\varepsilon$ and $\check G_1^\varepsilon$ refers to the first components of $\check K^\varepsilon$ and $\check G^\varepsilon$, while the index 2 in $\check K_2^\varepsilon$ and $\check G_2^\varepsilon$ refers to *all remaining components* of $\check K^\varepsilon$ and $\check G^\varepsilon$. Now, in order to eliminate the variable s from the previous two equations, one has to choose S such that

$$\begin{aligned}
&\check K_1^\varepsilon(y)\partial_t S(t,\tau,y) + \check K_2^\varepsilon(y)\cdot\nabla_y S(t,\tau,y) = 1,\\
&\partial_\tau S(t,\tau,y) + \check G_1^\varepsilon(y)\partial_t S(t,\tau,y) + \check G_2^\varepsilon(y)\cdot\nabla_y S(t,\tau,y) = 0,
\end{aligned} \tag{27}$$

and then

$$\begin{aligned}
&\check K_1^\varepsilon(y)\partial_t h(t,\tau,y) + \check K_2^\varepsilon(y)\cdot\nabla_y h(t,\tau,y) = 0,\\
&\partial_\tau h(t,\tau,y) + \check G_1^\varepsilon(y)\partial_t h(t,\tau,y) + \check G_2^\varepsilon(y)\cdot\nabla_y h(t,\tau,y) = 0,
\end{aligned} \tag{28}$$

with initial conditions

$$S(0,0,y) = 0, \qquad h(0,0,y) = f_0(y). \tag{29}$$

From these functions S and h, one can recover the distribution function $f(t,y)$ as follows: for any given (t,y), define $\tau(t,y)$ as a solution of

$$\tau(t,y) = \frac{S(t,\tau(t,y),y)}{\varepsilon}.$$

Then f can be obtained from the relation

$$\begin{aligned} h(t,\tau(t,y),y) &= \tilde{g}\left(S(t,\tau(t,y),y), \frac{S(t,\tau(t,y),y)}{\varepsilon}, t, y\right), \\ &= g(S(t,\tau(t,y),y),t,y) \\ &= f(t,y). \end{aligned}$$

Lemma 5.1. *Assume that $y \mapsto \check{K}_1^\varepsilon(y)$ does not vanish and consider the two vector fields*

$$\check{A}^\varepsilon := \frac{1}{\check{K}_1^\varepsilon}\,\check{K}_2^\varepsilon \quad \text{and} \quad \check{B}^\varepsilon := \check{G}_2^\varepsilon - \frac{\check{G}_1^\varepsilon}{\check{K}_1^\varepsilon}\,\check{K}_2^\varepsilon,$$

together with the two scalar functions

$$\check{\alpha}^\varepsilon := \frac{1}{\check{K}_1^\varepsilon} \quad \text{and} \quad \check{\beta}^\varepsilon := -\frac{\check{G}_1^\varepsilon}{\check{K}_1^\varepsilon}.$$

Then the following two relations hold true

$$\mathcal{L}_{\check{A}}\check{\beta} = \mathcal{L}_{\check{B}}\check{\alpha} \quad \text{and} \quad \mathcal{L}_{\check{A}}\mathcal{L}_{\check{B}} = \mathcal{L}_{\check{B}}\mathcal{L}_{\check{A}}. \tag{30}$$

Proof. Owing to Theorem 3.3, the two vector fields $\check{K}^\varepsilon$ and $\check{G}^\varepsilon$ have a vanishing Lie bracket (with respect to the $Y=(t,y)$ variable). This implies that

$$\partial_y \check{K}_1^\varepsilon(y)\;\check{G}_2^\varepsilon(y) - \partial_y \check{G}_1^\varepsilon(y)\;\check{K}_2^\varepsilon(y) = 0 \;\text{ and }\; \partial_y \check{K}_2^\varepsilon(y)\;\check{G}_2^\varepsilon(y) - \partial_y \check{G}_2^\varepsilon(y)\;\check{K}_2^\varepsilon(y) = 0. \tag{31}$$

By definition of $\check{\alpha}^\varepsilon$ and $\check{\beta}^\varepsilon$, the first relation may be rewritten as

$$\nabla_y \check{\beta}^\varepsilon \cdot \check{A}^\varepsilon - \nabla_y \check{\alpha}^\varepsilon \cdot \check{B}^\varepsilon = 0 \tag{32}$$

which proves the first statement of the lemma. Now, given a smooth vector field $L:\mathbb{R}^n \to \mathbb{R}^n$, a scalar function $a:\mathbb{R}^n \to \mathbb{R}$ and δy a vector of $\mathbb{R}^n$, the relation

$$(\partial_y(aL))\,\delta y = (\nabla_y a \cdot \delta y)\,L + a(\partial_y L)\,\delta y$$

holds true and may be used to compute the Lie bracket of $\check{A}^\varepsilon$ and $\check{B}^\varepsilon$ as follows

$$\begin{aligned} [\check{A}^\varepsilon, \check{B}^\varepsilon] =& \partial_y(\check{\alpha}^\varepsilon \check{K}_2^\varepsilon)\,\left(\check{G}_2^\varepsilon + \check{\beta}^\varepsilon \check{K}_2^\varepsilon\right) - \partial_y \check{G}_2^\varepsilon\,(\check{\alpha}^\varepsilon \check{K}_2^\varepsilon) - \partial_y(\check{\beta}^\varepsilon \check{K}_2^\varepsilon)\,(\check{\alpha}^\varepsilon \check{K}_2^\varepsilon) \\ =& \left(\nabla_y \check{\alpha}^\varepsilon \cdot \left(\check{G}_2^\varepsilon + \check{\beta}^\varepsilon \check{K}_2^\varepsilon\right)\right)\check{K}_2^\varepsilon + \check{\alpha}^\varepsilon(\partial_y \check{K}_2^\varepsilon)\,\left(\check{G}_2^\varepsilon + \check{\beta}^\varepsilon \check{K}_2^\varepsilon\right) \\ & - \check{\alpha}^\varepsilon(\partial_y \check{G}_2^\varepsilon)\,\check{K}_2^\varepsilon - \check{\alpha}^\varepsilon\left(\nabla_y \check{\beta}^\varepsilon \cdot \check{K}_2^\varepsilon\right)\check{K}_2^\varepsilon - \check{\alpha}^\varepsilon \check{\beta}^\varepsilon(\partial_y \check{K}_2^\varepsilon)\,\check{K}_2^\varepsilon. \end{aligned}$$

Using the second half of (31), the equality above simplifies to

$$\begin{aligned} [\check{A}^\varepsilon, \check{B}^\varepsilon] &= \left(\nabla_y \check{\alpha}^\varepsilon \cdot \check{G}_2^\varepsilon + \check{\beta}^\varepsilon \nabla_y \check{\alpha}^\varepsilon \cdot \check{K}_2^\varepsilon - \check{\alpha}^\varepsilon \nabla_y \check{\beta}^\varepsilon \cdot \check{K}_2^\varepsilon\right)\check{K}_2^\varepsilon \\ &= \left(\nabla_y \check{\alpha}^\varepsilon \cdot \check{B}^\varepsilon - \nabla_y \check{\beta}^\varepsilon \cdot \check{A}^\varepsilon\right)\check{K}_2^\varepsilon \end{aligned}$$

where the scalar term in factor of $\check{K}_2^\varepsilon$ now vanishes owing to (32). This implies the second statement of the lemma and completes its proof. □

Theorem 5.2. *Consider the functions $S(t,\tau,y)$ and $h(t,\tau,y)$ satisfying the following two separate systems of equations*

$$\check K_1^\varepsilon(y)\partial_t S(t,\tau,y) + \check K_2^\varepsilon(y)\cdot\nabla_y S(t,\tau,y) = 1, \tag{33}$$

$$\check K_1^\varepsilon(y)\partial_\tau S(t,\tau,y) + \Big(\check K_1^\varepsilon(y)\check G_2^\varepsilon(y) - \check G_1^\varepsilon(y)\check K_2^\varepsilon(y)\Big)\cdot\nabla_y S(t,\tau,y) = -\check G_1^\varepsilon(y), \tag{34}$$

$$S(0,0,y) = 0, \tag{35}$$

and

$$\check K_1^\varepsilon(y)\partial_t h(t,\tau,y) + \check K_2^\varepsilon(y)\cdot\nabla_y h(t,\tau,y) = 0, \tag{36}$$

$$\check K_1^\varepsilon(y)\partial_\tau h(t,\tau,y) + \Big(\check K_1^\varepsilon(y)\check G_2^\varepsilon(y) - \check G_1^\varepsilon(y)\check K_2^\varepsilon(y)\Big)\cdot\nabla_y h(t,\tau,y) = 0, \tag{37}$$

$$h(0,0,y) = f_0(y). \tag{38}$$

If the function $y\mapsto \check K_1^\varepsilon(y)$ does not vanish, then the following statements hold:

(i) system (36-37-38) has a unique solution h, periodic w.r.t. τ;
(ii) system (33-34-35) has a unique solution S, periodic w.r.t. τ;
*(iii) the **formal** expansion of the solution $f(t,y)$ of problem (19-20) satisfies*

$$f(t,y) = h(t,\tau(t,y),y),$$

where the function $(t,y)\mapsto\tau(t,y)\in\mathbb{R}$ is implicitly defined (locally) by the relation

$$\varepsilon\tau(t,y) = S(t,\tau(t,y),y).$$

Proof. A straightforward computation shows that the four equations (33), (34), (36), (37) are equivalent to the four equations in (27) and (28). Hence, if the separate systems (36-37) and (33-34) have unique solutions, they are clearly periodic w.r.t. τ. Now, proving the first statement requires to show that equations (36) and (37) can be solved in any order, i.e. that $\mathcal{L}_{\check A^\varepsilon}$ and $\mathcal{L}_{\check B^\varepsilon}$ commute, which is ensured by previous lemma. If h is the solution of (36-37-38), then $h(\cdot,0,\cdot)$ is in particular the solution of the Cauchy problem (36-38) and thus reads

$$h(t,0,\cdot) = \exp(-t\mathcal{L}_{\check A^\varepsilon})f_0.$$

Equation (37), which is a transport equation in variables (τ,y) with fixed parameter t, can then be uniquely solved. Given the initial data $h(t,0,\cdot)=\exp(-t\mathcal{L}_{\check A^\varepsilon})f_0$, this yields

$$h(t,\tau,\cdot) = \exp(-\tau\mathcal{L}_{\check B^\varepsilon})\exp(-t\mathcal{L}_{\check A^\varepsilon})f_0. \tag{39}$$

Hence, if a solution of (36-37-38) exists, it is necessarily of this form and thus unique. Conversely, one has, according to previous lemma

$$\exp(-t\mathcal{L}_{\check A^\varepsilon})\exp(-\tau\mathcal{L}_{\check B^\varepsilon})f_0 = \exp(-\tau\mathcal{L}_{\check B^\varepsilon})\exp(-t\mathcal{L}_{\check A^\varepsilon})f_0$$

and by differentiating the left-hand side w.r.t. t and the right-hand side τ, it may be checked that h given in (39) is indeed solution -thus the unique solution- of system (36-37-38). This proves (i).

Proceeding similarly for system (33-34-35), we first solve (33-35) for fixed $\tau=0$. This yields

$$S(t,0,\cdot) = \exp\big(-t\mathcal{L}_{\check A^\varepsilon}\big)\,S(0,0,\cdot) + \int_0^t \exp\big((s-t)\mathcal{L}_{\check A^\varepsilon}\big)\,ds\;\alpha = t\varphi\big(-t\mathcal{L}_{\check A^\varepsilon}\big)\,\alpha,$$

where $\varphi(z) = \frac{e^z-1}{z}$ is holomorphic on $\mathbb{C}$. The function S so-obtained then serves as initial condition for the evolution in τ through equation (34). This then leads to

$$\begin{aligned} S(t,0,\cdot) &= \exp\big(-\tau\mathcal{L}_{\breve{B}^\varepsilon}\big)t\varphi\big(-t\mathcal{L}_{\breve{A}^\varepsilon}\big)\,\alpha + \tau\varphi\big(-\tau\mathcal{L}_{\breve{B}^\varepsilon}\big)\,\beta \\ &= \tau\varphi\big(-\tau\mathcal{L}_{\breve{B}^\varepsilon}\big)\,\beta + t\varphi\big(-t\mathcal{L}_{\breve{A}^\varepsilon}\big)\alpha - t\,\tau\,\varphi\big(-t\mathcal{L}_{\breve{A}^\varepsilon}\big)\varphi\big(-\tau\mathcal{L}_{\breve{B}^\varepsilon}\big)\mathcal{L}_{\breve{B}^\varepsilon}\alpha \end{aligned}$$

where we have used the commutation of $\mathcal{L}_{\breve{A}^\varepsilon}$ and $\mathcal{L}_{\breve{B}^\varepsilon}$. Solving the equations in reverse order would have led to the symmetric variant

$$S(t,0,\cdot) = \tau\varphi\big(-\tau\mathcal{L}_{\breve{B}^\varepsilon}\big)\,\beta + t\varphi\big(-t\mathcal{L}_{\breve{A}^\varepsilon}\big)\alpha - t\,\tau\,\varphi\big(-t\mathcal{L}_{\breve{A}^\varepsilon}\big)\varphi\big(-\tau\mathcal{L}_{\breve{B}^\varepsilon}\big)\mathcal{L}_{\breve{A}^\varepsilon}\beta,$$

which, owing to Lemma 5.1 ($\mathcal{L}_{\breve{A}^\varepsilon}\beta = \mathcal{L}_{\breve{B}^\varepsilon}\alpha$), coincides with the first one. This proves *(ii)*.

It remains to prove *(iii)*. From (26) and the definition of h and $\tilde{g}$, we have

$$\forall\,(t,\tau,y),\ h(t,\tau,y) = \tilde{g}\,(S(t,\tau,y),\tau,t,y) \ \text{ and }\ \tilde{g}\left(S(t,\tau,y),\frac{S(t,\tau,y)}{\varepsilon},t,y\right) = f(t,y),$$

so that the value of $f(t,y)$ can be recovered from h and S through the formula

$$\forall\,(t,\tau,y),\quad f(t,y) = h(t,\tau(t,y),y),$$

provided that $\tau(t,y)$ satisfies

$$\varepsilon\tau(t,y) = S(t,\tau(t,y),y).$$

Given the periodicity of S w.r.t. τ, this equation always has a solution $\tau(t,y)$. $\square$

Remark 5. (truncated averaged models) If one keeps, in the expansions of the averaged fields $\breve{A}^\varepsilon$ and $\breve{B}^\varepsilon$ (defined in Lemma 5.1), only the terms of order less than (or equal to) n in ε, then the question arises whether the corresponding truncated averaged models[8] have a solution in the exact sense, and whether this solution is periodic w.r.t. τ. Generally speaking, the transport operators associated with the truncated fields $\breve{A}_n$ and $\breve{B}_n$ (i.e. $\breve{A}^\varepsilon = \breve{A}_n + O(\varepsilon^{n+1})$, $\breve{B}^\varepsilon = \breve{B}_n + O(\varepsilon^{n+1})$) do not commute exactly. More precisely, we only have $[\breve{A}_n,\breve{B}_n] = O(\varepsilon^{n+1})$. However, one can define an approximate solution by first solving

$$\partial_\tau h + \breve{B}_n(y)\cdot\nabla_y h = 0,\quad h(0,0,y) = f_0(y)$$

for fixed $t=0$ (in this way we obtain a solution $h_n(0,\tau,y)$ defined for all τ), and then solving

$$\partial_t h + \breve{A}_n(y)\cdot\nabla_y h = 0,\quad h(0,\tau,y) = h_n(0,\tau,y)$$

in order to get a solution $h_n^1(t,\tau,y)$ defined for all τ and t. At this stage, it is worth emphasizing that the function h_n^1 does not satisfy exactly the first equation for all t (only for $t=0$), since $[\breve{A}_n,\breve{B}_n]\neq 0$. Nevertheless, it does satisfy it up to terms of size ε^{n+1}. In particular, if one solves the two equations in reverse order, the function h_n^2 obtained does not coincide with h_n^1 exactly, but only up to terms of size ε^n and we have $h_n^1 - h_n^2 = O(\varepsilon^{n+1})$. In this sense, the result in previous theorem is at this stage only *formal*.

[8]i.e. the models obtained by removing all the terms of size ε^p for $p\geq n+1$.

5.3. **An illustrative elementary example.** Our aim here is to illustrate the result of Section 5 on an elementary example for which exact solutions can be easily obtained. Consider the following transport equation

$$\partial_t f + \left(\frac{1}{\varepsilon}\omega(y)Jy + y\right)\cdot\nabla_y f = 0, \qquad f(0,y) = f_0(y), \tag{40}$$

where $y \in \mathbb{R}^2$, and where

$$\omega(y) = 1 + |y|^2 = 1 + y_1^2 + y_2^2 \quad \text{and} \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

This equation can be solved as follows: let $\varphi_t^\varepsilon(y)$ be the flow of the characteristics equation

$$\dot{y} = \frac{1}{\varepsilon}\omega(y)Jy + y.$$

By taking its inner product by y, we have immediately $|\varphi_t^\varepsilon(y)| = \exp(t)|y|$, so that

$$\varphi_t^\varepsilon(y) = \exp(t)\exp\left(\frac{1}{\varepsilon}\left(t + (e^{2t}-1)\frac{|y|^2}{2}\right)J\right)y.$$

As a consequence, the explicit solution of (40) reads

$$f(t,y) = f_0\left(\exp\left(-t - \frac{1}{\varepsilon}\left(t + (1-e^{-2t})\frac{|y|^2}{2}\right)J\right)y\right). \tag{41}$$

Now, we observe that the two fields $\omega(y)Jy$ and $K(y) = y$ do not commute, and in order to transform the problem into a highly-oscillatory problem with y-independent frequency, one has to divide equation (40) by ω and immerse the equation on f into an augmented one for the unknown $g(s,t,y)$

$$\partial_s g + \frac{1}{\omega(y)}\partial_t g + \left(\frac{1}{\varepsilon}Jy + \frac{y}{\omega(y)}\right)\cdot\nabla_y g = 0, \qquad g(0,t,y) = f(t,y). \tag{42}$$

Unlike the fields $\omega(y)Jy$ and K, we now observe that the two augmented fields $\check{G}(y) = (0, Jy)^T$ and $\check{K}(y) = \left(\frac{1}{\omega(y)}, \frac{y}{\omega(y)}\right)^T$ do commute. This means that equation (42) is already written in a normal form and therefore the averaged fields in this case are simply

$$\check{G}^\varepsilon = (0, Jy)^T, \quad \check{K}^\varepsilon = (\check{K}_1^\varepsilon, \check{K}_2^\varepsilon)^T, \quad \text{with} \quad \check{K}_1^\varepsilon = \frac{1}{\omega(y)}, \quad \check{K}_2^\varepsilon = \frac{y}{\omega(y)}.$$

We now apply Theorem 5.2 in this particular case. The solution $h = h(t,0,y)$ to

$$\partial_t h + y\cdot\nabla_y h = 0, \qquad h(0,0,y) = f_0(y),$$

is $h(t,0,y) = f_0(e^{-t}y)$. As a consequence, the solution $h = h(t,\tau,y)$ to

$$\partial_\tau h + Jy\cdot\nabla_y h = 0, \qquad h(t,0,y) = f_0(e^{-t}y),$$

is

$$h(t,\tau,y) = f_0(e^{-t}e^{-\tau J}y). \tag{43}$$

The solution $S = S(t,0,y)$ to

$$\partial_t S + y\cdot\nabla_y S = \omega(y), \qquad S(0,0,y) = 0,$$

is simply $S(t,0,y) = t + (1-e^{-2t})\frac{|y|^2}{2}$, so that the solution $S = S(t,\tau,y)$ to

$$\partial_\tau S + Jy\cdot\nabla_y S = 0, \qquad S(t,0,y) = t + (1-e^{-2t})\frac{|y|^2}{2},$$

is constant w.r.t. τ, given that $|e^{\tau J}y|^2 = |y|^2$, i.e.

$$S(t,\tau,y) = t + (1-e^{-2t})\frac{|y|^2}{2}. \tag{44}$$

Theorem 5.2 asserts that $f(t,y) = h(t,\tau(t,y),y)$ where $\tau(t,y)$ is given by $\varepsilon\tau(t,y) = t + (1-e^{-2t})\frac{|y|^2}{2}$, an assertion which can be easily checked on our explicit example.

6. **Application to Vlasov equations with a strong magnetic field.** In this section, we consider the case of particles submitted to a strong magnetic field and evolving in an electric field $E(x)$ depending on the position x only. We recall hereinafter the corresponding equation (4) on the distribution function $f = f(t,x,v)$, $t \geq 0$, $x \in \mathbb{R}^3$, $v \in \mathbb{R}^3$:

$$\partial_t f + v\cdot\nabla_x f + \left(E(x) + v \times \frac{B(x)}{\varepsilon}\right)\cdot\nabla_v f = 0, \qquad f(0,x,v) = f_0(x,v), \tag{45}$$

which is closely related to the illustrative example of Section 4, though with the additional difficulty that E and B may vary. We further assume here that E derives from a potential U, i.e. that $E(x) = -\nabla_x U(x)$.

6.1. **Constant magnetic field.** Over a first phase, we assume that the magnetic field is constant. This means that, up to constant rotation, we have $B(x) = (0,0,b(x))^T$ and $b(x) \equiv b$. Upon rescaling the time $t \to t/b$ in f, i.e. considering the equation for $f(t/b,x,v)$ instead of $f(t,x,v)$ we may even assume that $b = 1$. We further assume in this first phase that the potential U depends only on the orthogonal direction (to B) of x, that is on the first two components (x_1,x_2) of x. This means that the electric field $E(x)$ is orthogonal to B and depends only on (x_1,x_2). Assume finally that the initial data f_0 only depends on (x_1,x_2) and (v_1,v_2), a property which is therefore propagated by the flow (45). All these assumptions allow us to restrict ourselves to a $2D\times 2D$ setting and to rewrite (45) in the form (1) with $n=4$, $y=(x,v)\in\mathbb{R}^2\times\mathbb{R}^2$ and

$$F^\varepsilon(y) = \begin{pmatrix} v \\ \frac{1}{\varepsilon}Jv + E(x)\end{pmatrix} = \frac{1}{\varepsilon}G + K \text{ with } G(y) = \begin{pmatrix} 0 \\ Jv\end{pmatrix} \text{ and } K(y) = \begin{pmatrix} v \\ E(x)\end{pmatrix}.$$

We now repeat the steps followed for the example of Section 4, starting first with the flow Φ_τ (associated with G)

$$\Phi_\tau(y) = \begin{pmatrix} x \\ e^{\tau J}v\end{pmatrix}.$$

The time-dependent vector field K_τ then writes

$$K_\tau(y) = \begin{pmatrix} e^{\tau J}v \\ e^{-\tau J}E(x)\end{pmatrix} = e^{i\tau}\hat{K}_1(y) + e^{-i\tau}\hat{K}_{-1}(y)$$

with

$$\hat{K}_1(y) = \frac{1}{2}\begin{pmatrix} v - iJv \\ E(x) + iJE(x)\end{pmatrix} \quad \text{and} \quad \hat{K}_{-1}(y) = \frac{1}{2}\begin{pmatrix} v + iJv \\ E(x) - iJE(x)\end{pmatrix}.$$

Formula (13) then gives

$$\begin{aligned} K^{[1]} &= \hat{K}_0 = 0, \\ K^{[2]} &= i[\hat{K}_1,\hat{K}_{-1}] = -2\Im\left((\partial_y\hat{K}_1)\hat{K}_{-1}\right) = \begin{pmatrix} JE \\ \frac{1}{2}(\Delta U)Jv\end{pmatrix} \end{aligned}$$

where we used computed successively

$$\frac{\partial \hat{K}_1}{\partial y} = \frac{1}{2}\begin{pmatrix} 0 & (I-iJ) \\ -(I+iJ)\nabla_x^2 U & 0 \end{pmatrix}$$

and[9]

$$\begin{aligned} \frac{\partial \hat{K}_1}{\partial y}\,\hat{K}_{-1} &= \frac{1}{4}\begin{pmatrix} (I-iJ)^2 E \\ -(I+iJ)\nabla_x^2 U(I+iJ)v \end{pmatrix} \\ &= \frac{1}{4}\begin{pmatrix} 2(I-iJ)E \\ -\left(\nabla_x^2 U - J\nabla_x^2 UJ + i(J\nabla_x^2 U + \nabla_x^2 UJ)\right)v \end{pmatrix} \\ &= \frac{1}{4}\begin{pmatrix} 2(I-iJ)E \\ -\left(\nabla_x^2 U + \det(\nabla_x^2 U)I + i\Delta UJ\right)v \end{pmatrix}. \end{aligned}$$

At second order in ε, equation (i) of Corollary 1 for $\tilde{f}(t,\tau,x,v)$ thus has the following form

$$\partial_\tau \tilde{f} + \varepsilon(v - \varepsilon JE)\cdot\nabla_x \tilde{f} + ((1-\varepsilon^2\Delta U)Jv + \varepsilon E)\cdot\nabla_v \tilde{f} = 0$$

while equation (ii) is simply

$$\partial_t \tilde{f} + \varepsilon JE\cdot\nabla_x \tilde{f} + \frac{\varepsilon}{2}(\Delta U)Jv\cdot\nabla_v \tilde{f} = 0. \tag{46}$$

This transport equation coincides, up to a rescaling in time, with the asymptotic model derived in [14]. We emphasize that, according to Remark 5, these two equations have to be understood in the approximate sense, which means that they cannot be satisfied exactly in general, but can only be solved approximately allowing errors of order ε^2.

Let us now stress that, although it may seem *natural* at first sight, the splitting of F^ε considered so far does not allow for the preservation of the Poisson structure of the original equation which becomes apparent by assuming that $E(x) = -\nabla_x U(x)$ and writing

$$F^\varepsilon(y) = \Omega^\varepsilon(x)\nabla_y \mathcal{H}(y) \text{ with } \Omega^\varepsilon(x) = \begin{pmatrix} 0 & I \\ -I & \frac{1}{\varepsilon}J \end{pmatrix} \text{ and } \mathcal{H}(y) = \frac{1}{2}\|v\|^2 + U(x).$$

As a matter of fact, choosing $G(y) = \begin{pmatrix} 0 \\ v \end{pmatrix}$ for the stiff term corresponds to a splitting of the structure matrix $\Omega^\varepsilon(x)$ and of the vector field as follows

$$\frac{1}{\varepsilon}G(y) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{\varepsilon}J \end{pmatrix}\nabla_y \mathcal{H}(y) \quad \text{and} \quad K(y) = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\nabla_y \mathcal{H}(y),$$

so that even though both G and K are Poisson vector fields they do not share the same structure. It follows that the systems of Corollary 1 are not Poisson. Nevertheless, another choice is possible that allows to maintain the geometric structure.

[9]Note that if S is a 2×2 symmetric matrix then

$$JS + SJ = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \alpha & \gamma \\ \gamma & \beta \end{pmatrix} + \begin{pmatrix} \alpha & \gamma \\ \gamma & \beta \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = (\alpha+\beta)J$$

so that $J\nabla^2 U + \nabla^2 UJ = (\Delta U)J$ and

$$JSJ = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \alpha & \gamma \\ \gamma & \beta \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -\det(S)I.$$

It consists in a splitting of the Hamiltonian $\mathcal{H}(y)$ in

$$\mathcal{H}(y) = A(y) + H(y) \quad \text{with} \quad A(y) = \frac{1}{2}\|v\|^2 \quad \text{and} \quad H(y) = V(x)$$

resulting into the new splitting of F^ε in

$$G(y) = \varepsilon\Omega^\varepsilon(x)A(y) = \begin{pmatrix} \varepsilon v \\ Jv \end{pmatrix} \quad \text{and} \quad K(y) = \Omega^\varepsilon(x)H(y) = \begin{pmatrix} 0 \\ E(x) \end{pmatrix}.$$

Note that whereas the notation adopted does not reflect it, the vector field G itself now depends on ε. The flow Φ_τ associated to G then reads

$$\Phi_\tau(y) = \begin{pmatrix} x + \varepsilon J^{-1}(e^{\tau J} - I)v \\ e^{\tau J} v \end{pmatrix}$$

and is clearly again 2π-periodic, so that we can again apply Corollary 1 and theorems 3.2 and 3.3 with

$$K_\tau(y) = \Omega^\varepsilon(y)\nabla_y H_\tau(y) \quad \text{where} \quad H_\tau(y) = U(x + \varepsilon J^{-1}(e^{\tau J} - I)v).$$

It then follows that

$$H^{[1]}(y) = \hat{H}_0(y) = \frac{1}{2\pi}\int_0^{2\pi} U(x + \varepsilon J^{-1}(e^{\tau J} - I)v)d\tau$$

and

$$\begin{aligned} K^{[1]}(y) &= \frac{1}{2\pi}\begin{pmatrix} 0 & I \\ -I & \frac{1}{\varepsilon}J \end{pmatrix}\begin{pmatrix} \int_0^{2\pi} \nabla_x U(x + \varepsilon J^{-1}(e^{\tau J} - I)v)d\tau \\ \varepsilon\int_0^{2\pi} J(e^{-\tau J} - I)\nabla_x U(x + \varepsilon J^{-1}(e^{\tau J} - I)v)d\tau \end{pmatrix} \\ &= \begin{pmatrix} \varepsilon\int_0^{2\pi} J(e^{-\tau J} - I)\nabla_x U(x + \varepsilon J^{-1}(e^{\tau J} - I)v)d\tau \\ -\int_0^{2\pi} e^{-\tau J}\nabla_x U(x + \varepsilon J^{-1}(e^{\tau J} - I)v)d\tau \end{pmatrix} \end{aligned}$$

so that equation (ii) of Corollary 1 for $\tilde{f}(t,\tau,x,v)$ now writes

$$\begin{aligned} \partial_t \tilde{f} + \varepsilon &\int_0^{2\pi} J(I - e^{-\tau J})E(x + \varepsilon J^{-1}(e^{\tau J} - I)v)d\tau \cdot \nabla_x \tilde{f} \\ &+ \int_0^{2\pi} e^{-\tau J}E(x + \varepsilon J^{-1}(e^{\tau J} - I)v)d\tau \cdot \nabla_v \tilde{f} = 0. \end{aligned} \tag{47}$$

In contrast with equation(46), the occurrence of ε in (47) is not the indication of a second-order approximation, but rather of a first-order approximation preserving the geometric structures of the original Vlasov equation. The corresponding equation (i) of Corollary 1 can be also deduced immediately, as well as second-order versions of (i) and (ii) from

$$\begin{aligned} H^{[2]}(y) &= \frac{-1}{4\pi}\int_0^{2\pi}\int_0^\tau \{H_s(y), H_\tau(y)\}dsd\tau \\ &= \frac{-1}{4\pi}\int_0^{2\pi}\int_0^\tau (\nabla_y H_s(y))^T\Omega^\varepsilon(x)\nabla_y H_\tau(y)dsd\tau. \end{aligned}$$

6.2. **Magnetic field with varying intensity and constant direction.** Over this second phase, we still work in a $2D \times 2D$ setting and keep the same notations as in the previous section. However, we address here the case of a magnetic field with varying intensity $b(x)$ and constant direction $B(x) = (0,0,b(x))^T$. Note that due to divergence free property of $B(x)$, the function b depends only on (x_1, x_2). In order to handle this case of varying intensity $b(x)$, one has to proceed as in Section 5. We first immerse the problem into an augmented one by adding a new parametrization variable s, then we derive averaging models at different orders for this augmented problem, and finally eliminate the extra-variable s from the averaged models and show how the original distribution function is recovered. In order to do so, we assume that $b(x)$ should not vanish for any x in $\mathbb{R}^2$ and we will make this assumption for the remaining of this section. The augmented distribution function $g(s,t,x,v)$ satisfies

$$\partial_s g + \frac{1}{b(x)}\partial_t g + \frac{1}{b(x)} v \cdot \nabla_x g + \left(\frac{1}{\varepsilon} Jv - \frac{1}{b(x)}\nabla_x U(x)\right) \cdot \nabla_v g = 0. \tag{48}$$

The original distribution function $f(t,x,v)$ is then viewed as a stationary solution of this evolution equation in s. Denoting $Y = (t, x_1, x_2, v_1, v_2) \in \mathbb{R}^5$ the now **extended** phase-space variable, we equivalently write (48) as follows

$$\partial_s g(s,Y) + \check{F}^\varepsilon(Y) \cdot \nabla_Y g(s,Y) = 0$$

where

$$\check{F}^\varepsilon(Y) = \begin{pmatrix} \frac{1}{b(x)} \\ \frac{1}{b(x)} v \\ \frac{1}{\varepsilon} Jv - \frac{1}{b(x)}\nabla_x U(x) \end{pmatrix}$$

is the **extended** vector field. We may now resume the derivation of the equations (i) and (ii) of Theorem 1, by first splitting $\check{F}^\varepsilon$ into $\check{F}^\varepsilon = \frac{1}{\varepsilon}\check{G} + \check{K}$ with

$$\check{G}(Y) = \begin{pmatrix} 0 \\ 0 \\ Jv \end{pmatrix} \quad \text{and} \quad \check{K}(Y) = \frac{1}{b(x)} \begin{pmatrix} 1 \\ v \\ -\nabla_x U(x) \end{pmatrix}.$$

It is clear that $\check{G}$ now generates a 2π-periodic flow

$$\check{\Phi}_\tau(Y) = \check{\Phi}_\tau \begin{pmatrix} t \\ x \\ v \end{pmatrix} = \begin{pmatrix} t \\ x \\ e^{\tau J} v \end{pmatrix}$$

whose period is independent of the trajectory. The function $\check{K}_\tau$ becomes

$$\check{K}_\tau(Y) = \frac{1}{b(x)} \begin{pmatrix} 1 \\ e^{\tau J} v \\ e^{-\tau J} E(x) \end{pmatrix}$$

and the corresponding Fourier modes are all vanishing except the modes 1, -1 and 0 (the additional one w.r.t. the case of a constant field):

$$\hat{K}_0(Y) = \begin{pmatrix} \frac{1}{b(x)} \\ 0 \\ 0 \end{pmatrix}, \quad \hat{K}_1(Y) = \frac{1}{2b(x)} \begin{pmatrix} 0 \\ (I - iJ)v \\ (I + iJ)E(x) \end{pmatrix},$$

and

$$\hat{K}_{-1}(Y) = \frac{1}{2b(x)} \begin{pmatrix} 0 \\ (I + iJ)v \\ (I - iJ)E(x) \end{pmatrix}.$$

According to Theorem 1, we thus have

$$K^{[1]}(Y) = \hat{K}_0(Y)$$

and

$$K^{[2]} = i\left([\hat{K}_1, \hat{K}_{-1}] + [\hat{K}_0, \hat{K}_1 - \hat{K}_{-1}]\right) = -2\Im([\hat{K}_0, \hat{K}_1]) - 2\Im\left((\partial_Y \hat{K}_1)\hat{K}_{-1}\right).$$

Omitting the argument x in E, U and b, and denoting simply ∇ for ∇_x, we have

$$\frac{\partial \hat{K}_0}{\partial Y} = \frac{-1}{b^2}\begin{pmatrix} 0 & \nabla^T b & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and

$$\frac{\partial \hat{K}_1}{\partial Y} = \frac{1}{2b^2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & -(I-iJ)v\,\nabla^T b & b(I-iJ) \\ 0 & -b(I+iJ)\nabla^2 U + (I+iJ)\nabla U\,\nabla^T b & 0 \end{pmatrix}$$

so that $(\partial_Y \hat{K}_1)\hat{K}_{-1}$ i given by

$$\begin{aligned}
&\frac{1}{4b^3}\begin{pmatrix} 0 & 0 & 0 \\ 0 & -(I-iJ)v\,\nabla^T b & b(I-iJ) \\ 0 & -b(I+iJ)\nabla^2 U + (I+iJ)\nabla U\,\nabla^T b & 0 \end{pmatrix}\begin{pmatrix} 0 \\ (I+iJ)v \\ (I-iJ)E \end{pmatrix} \\
&\qquad = \frac{1}{4b^3}\begin{pmatrix} 0 \\ -(I-iJ)v\,\nabla^T b\,(I+iJ)v + b(I-iJ)^2 E \\ -b(I+iJ)\nabla^2 U\,(I+iJ)v + (I+iJ)\nabla U\,\nabla^T b\,(I+iJ)v \end{pmatrix}
\end{aligned}$$

and finally

$$-2\Im\left((\partial_Y \hat{K}_1)\hat{K}_{-1}\right) = \frac{1}{2b^3}\begin{pmatrix} 0 \\ (\nabla b\cdot Jv)v - (\nabla b\cdot v)Jv + 2bJE \\ -\varepsilon(\nabla b\cdot v)J\nabla U - \varepsilon(\nabla b\cdot Jv)\nabla U + b(\Delta U)Jv \end{pmatrix}.$$

Besides, we have

$$\begin{aligned}
(\partial_Y \hat{K}_0)\hat{K}_1 &= \frac{-1}{2b^3}\begin{pmatrix} 0 & \nabla^T b & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ (I-iJ)v \\ (I+iJ)E \end{pmatrix} \\
&= \frac{-1}{2b^3}\begin{pmatrix} \nabla b\cdot v - i\nabla b\cdot Jv \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

and

$$\begin{aligned}
(\partial_Y \hat{K}_1)\hat{K}_0 &= \frac{1}{2b^3}\begin{pmatrix} 0 & 0 & 0 \\ 0 & -(I-iJ)v\,\nabla^T b & b(I-iJ) \\ 0 & -b(I+iJ)\nabla^2 U + (I+iJ)\nabla U\,\nabla^T b & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\
&= 0
\end{aligned}$$

so that

$$-2\Im([\hat{K}_0, \hat{K}_1]) = \frac{-1}{b^3}\begin{pmatrix} \nabla b\cdot Jv \\ 0 \\ 0 \end{pmatrix}.$$

Finally, at first order in ε, we have

$$\check{K}^\varepsilon = K^{[1]} + \varepsilon K^{[2]} = \frac{1}{b}\begin{pmatrix} 1 - \varepsilon\,\frac{\nabla b\cdot Jv}{b^2} \\ -\varepsilon\,\frac{(\nabla b\cdot v)}{2b^2}Jv + \varepsilon\,\frac{(\nabla b\cdot Jv)}{2b^2}v - \varepsilon\,\frac{1}{b}J\nabla U \\ -\frac{\varepsilon(\nabla b\cdot v)}{2b^2}J\nabla U - \frac{\varepsilon(\nabla b\cdot Jv)}{2b^2}\nabla U + \frac{\varepsilon\Delta U}{2b}Jv \end{pmatrix} = \begin{pmatrix} K_1^\varepsilon \\ K_2^\varepsilon \end{pmatrix},$$

and

$$\check{G}^\varepsilon = \varepsilon(\check{F}^\varepsilon - \check{K}^\varepsilon) = \frac{1}{b}\begin{pmatrix} 0 \\ \varepsilon v \\ bJv - \varepsilon\nabla U \end{pmatrix}.$$

Therefore the transport equations[10] on h are at first order in ε:

$$\partial_t h + \frac{\varepsilon}{2b}\left(\frac{\nabla b\cdot Jv}{b}v - \frac{\nabla b\cdot v}{b}Jv - 2J\nabla U\right)\cdot\nabla_x h$$
$$-\frac{\varepsilon}{2b}\left(\frac{\nabla b\cdot v}{b}J\nabla U + \frac{\nabla b\cdot Jv}{b}\nabla U - (\Delta U)Jv\right)\cdot\nabla_v h = 0, \tag{49}$$

and

$$\partial_\tau h + \frac{\varepsilon}{b}v\cdot\nabla_x h + \left(Jv - \frac{\varepsilon}{b}\nabla U\right)\cdot\nabla_v h = 0, \tag{50}$$

with the initial condition $h(0,0,y) = f_0(y)$. Similarly the transport equations on S are

$$\partial_t S + \frac{\varepsilon}{2b}\left(\frac{\nabla b\cdot Jv}{b}v - \frac{\nabla b\cdot v}{b}Jv - 2J\nabla U\right)\cdot\nabla_x S$$
$$-\frac{\varepsilon}{2b}\left(\frac{\nabla b\cdot v}{b}J\nabla U + \frac{\nabla b\cdot Jv}{b}\nabla U - (\Delta U)Jv\right)\cdot\nabla_v S = b(x)\left(1 + \varepsilon\,\frac{\nabla b\cdot Jv}{b^2}\right), \tag{51}$$

and

$$\partial_\tau S + \frac{\varepsilon}{b}v\cdot\nabla_x S + \left(Jv - \frac{\varepsilon}{b}\nabla U\right)\cdot\nabla_v S = 0, \tag{52}$$

with the initial condition $S(0,0,y) = 0$. Again, we wish to put the stress on the fact that these two truncated models in h and S should be understood in the sense of Remark 5.

Now we make some important comments on these transport equations. The transport equation (49) coincides with the gyro-kinetic model that has been derived in [14] in the particular case of constant b. It also contains all the terms in the models recently derived in [5] in the case of varying $b = b(x)$ when restricted to the $2D\times 2D$ geometry. However, in addition to the fact that our averaged models keep all the variables (x, v), our approach provides more information through the phase S and the dependence in τ. These informations are necessary to correctly reconstruct the full original distribution function f (and not only the averaged model) at first order in ε. This reconstruction may be performed through the relation $f(t,x,v) = h(t,\tau(t,x,v),x,v) + O(\varepsilon^2)$ where $\tau(t,x,v)$ is a solution to $\varepsilon\tau = S(t,\tau,x,v)$. Up to our knowledge, no such construction can be found in the literature.

6.3. **Magnetic field in 3D with varying intensity and varying direction.** We now consider the transport kinetic equation in its general form (4) and in a $3D\times 3D$ setting. This means in particular that now we allow variations of B in both amplitude and direction. Our aim in this part is to extend our previous approach to this more general case.

We first immerse the model (4) into an augmented problem in the unknown $g(s,t,x,v)$, as follows

$$\partial_s g + \frac{1}{|B(x)|}\partial_t g + \frac{v}{|B(x)|}\cdot\nabla_x g + \left(\frac{E(x)}{|B(x)|} + \frac{1}{\varepsilon}v\times\frac{B(x)}{|B(x)|}\right)\cdot\nabla_v g = 0 \tag{53}$$

with the initial condition $g(0,t,x,v) = f(t,x,v)$. The main interest of this form is that the oscillatory part in the variable s is now driven by the vector field $v\times\frac{B(x)}{|B(x)|}$, which, as we shall see, generates a periodic flow with a constant period 2π. More precisely, the trajectories

[10] An easy calculation using polar coordinates shows that the magnetic moment is preserved.

$$\dot{x}(s) = 0, \qquad \dot{v}(s) = v(s) \times \frac{B(x(s))}{|B(x(s))|}, \qquad (x(0), v(0)) = (x_0, v_0) \in \mathbb{R}^3 \times \mathbb{R}^3,$$

are all periodic with the same period 2π independently of (x_0, v_0).

In particular the period does not depend on the trajectory although the unit vector $\frac{B(x)}{|B(x)|}$ depends on this trajectory. Indeed let e_0 be a unit vector and let (e_1, e_2, e_0) be an orthonormal basis such that $e_0 \times e_1 = e_2$ and $e_1 \times e_2 = e_0$. The matrix representing the skew-symmetric linear map $\mathcal{J}_{e_0} : v \mapsto v \times e_0$ in the basis (e_1, e_2, e_0), is simply $\mathcal{J} = \begin{pmatrix} J & 0 \\ 0 & 0 \end{pmatrix}$. Since $\exp(t\mathcal{J})$ is 2π-periodic, the flow $\exp(t\mathcal{J}_{e_0})$ is 2π-periodic. We now apply our methodology to model (53). Here the vector field $\check{F}^\varepsilon = \frac{1}{\varepsilon}\check{G} + \check{K}$ is given by

$$\check{K}(t, x, v) = \frac{1}{|B(x)|} \begin{pmatrix} 1 \\ v \\ E(x) \end{pmatrix}, \qquad \check{G}(t, x, v) = \begin{pmatrix} 0 \\ 0 \\ v \times \frac{B(x)}{|B(x)|} \end{pmatrix}.$$

We introduce the following notations

$$e(x) = \frac{B(x)}{|B(x)|}, \quad \mathcal{J}_e v = v \times e, \quad \mathcal{P}_e v = (e \cdot v)e, \quad \forall e \in \mathbb{S}^2, v \in \mathbb{R}^3, x \in \mathbb{R}^3. \tag{54}$$

Using Theorem 3.3, the vector field K_τ can be easily computed to get

$$\Phi_\tau(t, x, v) = \begin{pmatrix} t \\ x \\ \exp\left(\tau \mathcal{J}_{e(x)}\right) v \end{pmatrix}.$$

The following elementary identities

$$\mathcal{J}_e^2 = -I + \mathcal{P}_e, \qquad \mathcal{J}_e \mathcal{P}_e = \mathcal{P}_e \mathcal{J}_e = 0$$

imply that

$$\begin{aligned} \Phi_\tau(t, x, v) &= \begin{pmatrix} t \\ x \\ (\cos\tau)v + (1 - \cos\tau)\mathcal{P}_{e(x)} v + (\sin\tau)\mathcal{J}_{e(x)} v \end{pmatrix} \\ &= \begin{pmatrix} t \\ x \\ (\cos\tau)v + (1 - \cos\tau)(e(x) \cdot v)e(x) + (\sin\tau) v \times e(x) \end{pmatrix}. \end{aligned} \tag{55}$$

We then deduce the expression of the Jacobian matrix $\partial_{t,x,v}\Phi_\tau = (\partial_t \Phi_\tau, \partial_x \Phi_\tau, \partial_v \Phi_\tau)$:

$$\partial_{t,x,v}\Phi_\tau = \begin{pmatrix} 1 & 0 & 0 \\ 0 & I & 0 \\ 0 & R_\tau & Q_\tau \end{pmatrix},$$

where

$$\begin{aligned} R_\tau &= (1 - \cos\tau)\partial_x\left(\mathcal{P}_{e(x)} v\right) + (\sin\tau)\partial_x\left(\mathcal{J}_{e(x)} v\right) \\ &= \alpha_0 + \alpha e^{i\tau} + \overline{\alpha} e^{-i\tau}, \\ Q_\tau &= (\cos\tau)I + (1 - \cos\tau)\mathcal{P}_{e(x)} + (\sin\tau)\mathcal{J}_{e(x)} \\ &= a_0 + a e^{i\tau} + \overline{a} e^{-i\tau}, \end{aligned}$$

and

$$\begin{aligned} a_0 &= \mathcal{P}_{e(x)}, \qquad \alpha_0 = \partial_x\left(\mathcal{P}_{e(x)} v\right), \\ 2a &= I - \mathcal{P}_{e(x)} - i\mathcal{J}_{e(x)}, \\ 2\alpha &= -\partial_x\left(\mathcal{P}_{e(x)} v + i\mathcal{J}_{e(x)} v\right). \end{aligned}$$

Note that the matrix R_τ takes care with the so-called curvature terms which are the terms coming from the space variation of the direction $e(x)$ of the magnetic field. In order to compute the inverse of the matrix $\partial_{t,x,v}\Phi_\tau$, we observe that

$$(\partial_{t,x,v}\Phi_\tau)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & I & 0 \\ 0 & -Q_\tau^{-1}R_\tau & Q_\tau^{-1} \end{pmatrix},$$

which means that we only need to compute Q_τ^{-1}. Using again the identity $\mathcal{J}_e^2 = -I + \mathcal{P}_e$, one may check

$$\begin{aligned} Q_\tau^{-1} &= (\cos\tau)I + (1-\cos\tau)\mathcal{P}_{e(x)} - (\sin\tau)\mathcal{J}_{e(x)}, \\ &= a_0 + \overline{a}e^{i\tau} + ae^{-i\tau} = Q_{-\tau}. \end{aligned}$$

Now we also have

$$\check{K} \circ \Phi_\tau(t,x,v) = \frac{1}{|B(x)|}\begin{pmatrix} 1 \\ Q_\tau v \\ E(x) \end{pmatrix},$$

and therefore

$$\check{K}_\tau(t,x,v) = \frac{1}{|B(x)|}\begin{pmatrix} 1 \\ Q_\tau v \\ -Q_{-\tau}R_\tau Q_\tau v + Q_{-\tau}E(x) \end{pmatrix}.$$

One can easily see that the Fourier expansion of $\check{K}_\tau$ (in the periodic variable τ) only contains modes $k \in \mathbb{Z}$ with $|k| \le 3$. Note that we can recover the previous case (in which $B(x)$ had a constant direction and $(x,v) \in \mathbb{R}^2 \times \mathbb{R}^2$) by taking $\mathcal{P}_{e(x)}v = 0$, $\mathcal{J}_{e(x)} \equiv \mathcal{J} = \begin{pmatrix} J & 0 \\ 0 & 0 \end{pmatrix}$ and $\alpha = 0$, which means that $R_\tau = 0$ and $Q_\tau = e^{\tau\mathcal{J}}$.

Although all the Fourier coefficients of $\check{K}_\tau$ can be derived from this expression, we just give for simplicity the 0^{th} mode:

$$\hat{K}_0(x,v) = \frac{1}{|B(x)|}\begin{pmatrix} 1 \\ \mathcal{P}_{e(x)}v = (e(x)\cdot v)e(x) \\ (\hat{K}_0)_3 \end{pmatrix} = K^{[1]},$$

with

$$\begin{aligned} (\hat{K}_0)_3(x,v) &= a_0E(x) - (a_0\alpha_0a_0 + a_0\alpha\overline{a} + a_0\overline{\alpha}a + \overline{a}\alpha_0\overline{a} + \overline{a}\overline{\alpha}a_0 + a\alpha_0a + a\alpha a_0)v \\ &= \mathcal{P}_{e(x)}E(x) - \Big[4\mathcal{P}_{e(x)}\partial_x\left(\mathcal{P}_{e(x)}v\right)\mathcal{P}_{e(x)} + \frac{1}{2}\mathcal{P}_{e(x)}\partial_x\left(\mathcal{J}_{e(x)}v\right)\mathcal{J}_{e(x)} \\ &-\frac{1}{2}\mathcal{J}_{e(x)}\partial_x\left(\mathcal{P}_{e(x)}v\right)\mathcal{J}_{e(x)} - \frac{1}{2}\mathcal{J}_{e(x)}\partial_x\left(\mathcal{J}_{e(x)}v\right)\mathcal{P}_{e(x)} \\ &-\mathcal{P}_{e(x)}\partial_x\left(\mathcal{P}_{e(x)}v\right) - \partial_x\left(\mathcal{P}_{e(x)}v\right)\mathcal{P}_{e(x)} + \frac{1}{2}\partial_x\left(\mathcal{P}_{e(x)}v\right)\Big]v. \end{aligned}$$

We then deduce the vector field G^ε at the 0^{th} order in ε:

$$\begin{aligned} G^{[1]} &= \varepsilon(\check{F}^\varepsilon - K^{[1]}) + \mathcal{O}(\varepsilon) \\ &= \tfrac{1}{|B(x)|}\begin{pmatrix} 0 \\ \varepsilon\left(v - \mathcal{P}_{e(x)}v\right) \\ |B(x)|\mathcal{L}_{e(x)}v + \varepsilon\left(E(x) - (\hat{K}_0)_3(x,v)\right) \end{pmatrix} + \mathcal{O}(\varepsilon) \\ &= \begin{pmatrix} 0 \\ 0 \\ \mathcal{L}_{e(x)}v \end{pmatrix} + \mathcal{O}(\varepsilon). \end{aligned}$$

The averaged model at the 0^{th} order in ε can now be written in terms of $h(t,\tau,x,v)$ and $S(t,\tau,x,v)$. We have

$$\begin{aligned} &\partial_t h + \left(\tfrac{B(x)}{|B(x)|}\cdot v\right)\tfrac{B(x)}{|B(x)|}\cdot\nabla_x h + (\hat{K}_0)_3(x,v)\cdot\nabla_v h = 0, \\ &\partial_\tau h + \left(v \times \tfrac{B(x)}{|B(x)|}\right)\cdot\nabla_v h = 0, \end{aligned} \tag{56}$$

and

$$\begin{aligned} &\partial_t S + \left(\tfrac{B(x)}{|B(x)|}\cdot v\right)\tfrac{B(x)}{|B(x)|}\cdot\nabla_x S + (\hat{K}_0)_3(x,v)\cdot\nabla_v S = |B(x)|, \\ &\partial_\tau S + \left(v \times \tfrac{B(x)}{|B(x)|}\right)\cdot\nabla_v S = 0, \end{aligned}$$

with the initial conditions: $h(0,0,x,v) = f_0(x,v)$ and $S(0,0,x,v) = 0$. Note that in the particular case where $B(x)$ has a constant direction $B(x) = b(x)e_0 = (0,0,b(x))^T$, we get

$$\begin{aligned} &\partial_t h + \ v_\parallel\partial_{x_\parallel}h + E_\parallel\partial_{v_\parallel}h = 0, \\ &\partial_\tau h + Jv_\perp\cdot\partial_{v_\perp}h = 0. \end{aligned}$$

and

$$\begin{aligned} &\partial_t S + \ v_\parallel\partial_{x_\parallel}S + E_\parallel\partial_{v_\parallel}S = b(x), \\ &\partial_\tau S + Jv_\perp\cdot\partial_{v_\perp}S = 0, \end{aligned}$$

where we used the standard notations $v_\parallel = v\cdot e_0,\ E_\parallel = E\cdot e_0$ and

$$v = (v_1, v_2, v_\parallel) = (v_\perp, v_\parallel), \quad E = (E_1, E_2, E_\parallel) = (E_\perp, E_\parallel), \quad \partial_{v_\perp}h = (\partial_{v_1}h, \partial_{v_2}h),$$

and the same notations for the space variable x. Observe that the exact solution of the two equations for S (for the 0^{th} order in ε) is simply $S(t,\tau,x,v) = b(x)t$.

The averaged equations at the first order in ε can also be derived in the case of a magnetic field $B(x)$ with constant direction $B(x) = (0,0,b(x))^T$, with $b(x) > 0$. In this case we have $R_\tau = 0$, $e(x)$ is the constant unit vector e_0, $|B(x)| = b(x)$, and therefore

$$\check{K}_\tau(t,x,v) = \frac{1}{b(x)}\begin{pmatrix} 1 \\ Q_\tau v \\ Q_{-\tau}E(x) \end{pmatrix}.$$

The non-zero Fourier modes in τ of this quantity $\check{K}_\tau$ are

$$\hat{K}_0 = \frac{1}{b}\begin{pmatrix} 1 \\ a_0 v \\ a_0 E \end{pmatrix}, \quad \hat{K}_1 = \frac{1}{b}\begin{pmatrix} 0 \\ av \\ \overline{a}E \end{pmatrix}, \quad \hat{K}_{-1} = \frac{1}{b}\begin{pmatrix} 0 \\ \overline{a}v \\ aE \end{pmatrix}.$$

The computation of K^ε at first order in ε can then be derived from Theorem 3.3 as follows. We know from Theorem 3.3 that $\check{K}^\varepsilon = K^{[1]} + \varepsilon K^{[2]}$ with

$$K^{[1]} = \hat{K}_0, \quad K^{[2]} = -2\Im\left((\partial_Y\hat{K}_1)\hat{K}_{-1}\right) - 2\Im\left([\hat{K}_0, \hat{K}_1]\right).$$

Since

$$\partial_Y \hat{K}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -a\left(v \otimes \frac{\nabla b}{b^2}\right) & \frac{a}{b} \\ 0 & \overline{a}\partial_x\left(\frac{E}{b}\right) & 0 \end{pmatrix}, \quad \partial_Y \hat{K}_0 = \begin{pmatrix} 0 & -\frac{(\nabla b)^T}{b^2} & 0 \\ 0 & -a_0\left(v \otimes \frac{\nabla b}{b^2}\right) & \frac{a_0}{b} \\ 0 & a_0\partial_x\left(\frac{E}{b}\right) & 0 \end{pmatrix}$$

we get for $2\Im\left((\partial_Y \hat{K}_1)\hat{K}_{-1}\right)$

$$\frac{1}{2b}\begin{pmatrix} 0 \\ -(I-\mathcal{P})\left(v \otimes \frac{\nabla b}{b^2}\right)\mathcal{J}v + \mathcal{J}\left(v \otimes \frac{\nabla b}{b^2}\right)(I-\mathcal{P})v - \frac{2}{b}\mathcal{J}E \\ (I-\mathcal{P})\partial_x\left(\frac{E}{b}\right)\mathcal{J}v + \mathcal{J}\partial_x\left(\frac{E}{b}\right)(I-\mathcal{P})v \end{pmatrix}$$

where we have denoted $\mathcal{P} = \mathcal{P}_{e_0}$ and $\mathcal{J} = \mathcal{J}_{e_0}$. We also have

$$2\Im\left([\hat{K}_0, \hat{K}_1]\right) = \frac{1}{b}\begin{pmatrix} \mathcal{J}v \cdot \frac{\nabla b}{b^2} \\ \mathcal{P}\left(v \otimes \frac{\nabla b}{b^2}\right)\mathcal{J}v - \mathcal{J}\left(v \otimes \frac{\nabla b}{b^2}\right)\mathcal{P}v \\ -\mathcal{P}\partial_x\left(\frac{E}{b}\right)\mathcal{J}v - \mathcal{J}\partial_x\left(\frac{E}{b}\right)\mathcal{P}v \end{pmatrix},$$

therefore

$$K^{[2]} = \frac{1}{b}\begin{pmatrix} -\mathcal{J}v \cdot \frac{\nabla b}{b^2} \\ \frac{1}{2}\left(\mathcal{J}v \cdot \frac{\nabla b}{b^2}\right)(I-3\mathcal{P})v - \frac{1}{2}\left((I-3\mathcal{P})v \cdot \frac{\nabla b}{b^2}\right)\mathcal{J}v + \frac{1}{b}\mathcal{J}E \\ -\frac{1}{2}(I-3\mathcal{P})\partial_x\left(\frac{E}{b}\right)\mathcal{J}v - \frac{1}{2}\mathcal{J}\partial_x\left(\frac{E}{b}\right)(I-3\mathcal{P})v \end{pmatrix},$$

and

$$\begin{aligned} \check{K}^\varepsilon =& K^{[1]} + \varepsilon K^{[2]} + O(\varepsilon^2) \\ =& \frac{1}{b}\begin{pmatrix} 1 - \varepsilon\mathcal{J}v \cdot \frac{\nabla b}{b^2} \\ v_\parallel e_0 + \frac{\varepsilon}{2}\left[\left(\mathcal{J}v \cdot \frac{\nabla b}{b^2}\right)(I-3\mathcal{P})v - \left((I-3\mathcal{P})v \cdot \frac{\nabla b}{b^2}\right)\mathcal{J}v\right] + \frac{\varepsilon}{b}\mathcal{J}E \\ E_\parallel e_0 - \frac{\varepsilon}{2}\left[(I-3\mathcal{P})\partial_x\left(\frac{E}{b}\right)\mathcal{J}v + \mathcal{J}\partial_x\left(\frac{E}{b}\right)(I-3\mathcal{P})v\right] \end{pmatrix} \\ &+ O(\varepsilon^2). \end{aligned}$$

We finally deduce the field G^ε at first order in ε

$$\check{G}^\varepsilon = \varepsilon(\check{F}^\varepsilon - K^{[1]} - \varepsilon K^{[2]}) + O(\varepsilon^2) = \begin{pmatrix} 0 \\ 0 \\ \mathcal{J}v \end{pmatrix} + \frac{\varepsilon}{b}\begin{pmatrix} 0 \\ v_\perp \\ E_\perp \end{pmatrix} + O(\varepsilon^2).$$

Therefore, the evolution in time t of h at the first order in ε is driven by the following equation (with the above described notations)

$$\begin{aligned} &\left[1 - \varepsilon J v_\perp \cdot \frac{\partial_{x_\perp} b}{b^2}\right]\partial_t h + \ v_\parallel\left[1 - \varepsilon J v_\perp \cdot \frac{\partial_{x_\perp} b}{b^2}\right]\partial_{x_\parallel} h + \left[E_\parallel + \varepsilon\partial_{x_\perp}\left(\frac{E_\parallel}{b}\right) \cdot v_\perp\right]\partial_{v_\parallel} h \\ &\qquad -\frac{\varepsilon}{2b}\left[|v_\perp|^2 J\frac{\partial_{x_\perp} b}{b} - 2JE_\perp\right] \cdot \partial_{x_\perp} h \\ &\qquad +\frac{\varepsilon}{2}\left[\left(\frac{\partial_{x_\perp} b}{b^2} \cdot JE_\perp\right)v_\perp + 2v_\parallel\partial_{x_\parallel}\left(\frac{E_\perp}{b}\right) - \partial_{x_\parallel}\left(\frac{E_\perp}{b}\right)Jv_\perp\right] \cdot \partial_{v_\perp} h = 0, \end{aligned}$$

which simplifies into

$$\begin{aligned} &\partial_t h + \ v_\parallel\partial_{x_\parallel} h + \left[E_\parallel + \varepsilon E_\parallel J v_\perp \cdot \frac{\partial_{x_\perp} b}{b^2} + \varepsilon\partial_{x_\perp}\left(\frac{E_\parallel}{b}\right) \cdot v_\perp\right]\partial_{v_\parallel} h \\ &\qquad -\frac{\varepsilon}{2b}\left[|v_\perp|^2 J\frac{\partial_{x_\perp} b}{b} - 2JE_\perp\right] \cdot \partial_{x_\perp} h \\ &\qquad +\frac{\varepsilon}{2}\left[\left(\frac{\partial_{x_\perp} b}{b^2} \cdot JE_\perp\right)v_\perp + 2v_\parallel\partial_{x_\parallel}\left(\frac{E_\perp}{b}\right) - \partial_{x_\parallel}\left(\frac{E_\perp}{b}\right)Jv_\perp\right] \cdot \partial_{v_\perp} h = 0. \end{aligned} \tag{57}$$

Note that we have used the identity $\nabla_x \cdot B = 0$ which implies that $b(x) = b(x_\perp)$. This provides an asymptotic model which is identical to the one recently derived in [5] or, up to a rescaling in time, to the one derived in [12]. However our approach provides more informations since this equation still contains all the original variables (x, v) of the distribution function and has to be coupled with an equation describing its dependence on a periodic variable τ which has to fit with a suitable phase function S. As we shall see, this equation in τ will provide a suitable initial data for equation (57). The second equation on h writes

$$\partial_\tau h + Jv_\perp \cdot \partial_{v_\perp} h + \frac{\varepsilon}{b} v_\perp \cdot \partial_{x_\perp} h + \varepsilon \frac{E_\perp}{b} \cdot \partial_{v_\perp} h = 0. \tag{58}$$

The system of the two equations (57-58) is subjected to the initial data $h(0, 0, x, v) = f_0(x, v)$.

Once again, we recall that system (57-58) with initial condition $h(0, 0, x, v) = f_0(x, v)$ is only valid up to ε^2 terms, and solutions to this system have to be understood in the sense of Remark 5.

Similarly the equations on S are

$$\begin{aligned} \partial_t S + \ v_\| \partial_{x_\|} S &+ \left[E_\| + \varepsilon E_\| J v_\perp \cdot \frac{\partial_{x_\perp} b}{b^2} + \varepsilon \partial_{x_\perp} \left(\frac{E_\|}{b} \right) \cdot v_\perp \right] \partial_{v_\|} S \\ & - \frac{\varepsilon}{2b} \left[|v_\perp|^2 \frac{\partial_{x_\perp} b}{b} - 2JE_\perp \right] \cdot \partial_{x_\perp} S \\ & + \frac{\varepsilon}{2} \left[\left(\frac{\partial_{x_\perp} b}{b^2} \cdot JE_\perp \right) v_\perp + 2 v_\| \partial_{x_\|} \left(\frac{E_\perp}{b} \right) - \partial_{x_\|} \left(\frac{E_\perp}{b} \right) J v_\perp \right] \cdot \partial_{v_\perp} S \\ & \quad = b + \varepsilon J v_\perp \cdot \frac{\partial_{x_\perp} b}{b}, \end{aligned} \tag{59}$$

and

$$\partial_\tau S + Jv_\perp \cdot \partial_{v_\perp} S + \frac{\varepsilon}{b} v_\perp \cdot \partial_{x_\perp} S + \varepsilon \frac{E_\perp}{b} \cdot \partial_{v_\perp} S = 0, \tag{60}$$

with the initial data $S(0, 0, x, v) = 0$.

We now observe that $S(t, \tau, x, v) = b(x)t + O(\varepsilon)$, and therefore it is more convenient to write these equations in terms of

$$\tilde{S}(t, \tau, x, v) = \frac{S(t, \tau, x, v) - b(x)t}{\varepsilon}$$

and get

$$\begin{aligned} \partial_t \tilde{S} + \ v_\| \partial_{x_\|} \tilde{S} &+ \left[E_\| + \varepsilon E_\| J v_\perp \cdot \frac{\partial_{x_\perp} b}{b^2} + \varepsilon \partial_{x_\perp} \left(\frac{E_\|}{b} \right) \cdot v_\perp \right] \partial_{v_\|} \tilde{S} \\ & - \frac{\varepsilon}{2b} \left[|v_\perp|^2 \frac{\partial_{x_\perp} b}{b} - 2JE_\perp \right] \cdot \partial_{x_\perp} \tilde{S} \\ & + \frac{\varepsilon}{2} \left[\left(\frac{\partial_{x_\perp} b}{b^2} \cdot JE_\perp \right) v_\perp + 2 v_\| \partial_{x_\|} \left(\frac{E_\perp}{b} \right) - \partial_{x_\|} \left(\frac{E_\perp}{b} \right) J v_\perp \right] \cdot \partial_{v_\perp} \tilde{S} \\ & = J v_\perp \cdot \frac{\partial_{x_\perp} b}{b}, \end{aligned}$$

and

$$\partial_\tau \tilde{S} + Jv_\perp \cdot \partial_{v_\perp} \tilde{S} + \varepsilon \frac{t}{b} v_\perp \cdot \partial_{x_\perp} b + \frac{\varepsilon}{b} v_\perp \cdot \partial_{x_\perp} \tilde{S} + \varepsilon \frac{E_\perp}{b} \cdot \partial_{v_\perp} \tilde{S} = 0,$$

with the initial data $\tilde{S}(0, 0, x, v) = 0$. We then recover the solution f by the relation

$$f(t, x, v) = h(t, \tau(t, x, v), x, v),$$

where $(t,x,v) \mapsto \tau(t,x,v) \in \mathbb{R}$ is implicitly defined (locally) from $\tilde{S}$ by the equation

$$\tau(t,x,v) = \frac{b(x)t}{\varepsilon} + \tilde{S}(t,\tau(t,x,v),x,v).$$

Acknowledgments. This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom Research and Training Programme 2019-2020. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

REFERENCES

[1] S. Blanes, F. Casas, J.A. Oteo, J. Ros, *A pedagogical approach to the Magnus expansion*, European Journal of Physics **31**, 907-918, 2010.

[2] M. Bostan, *The Vlasov-Maxwell system with strong initial magnetic field. Guiding-center approximation*, Multiscale Model. Simul. **6**, 1026-1058, 2007.

[3] M. Bostan, *Transport equations with disparate advection fields. Application to the gyrokinetic models in plasma physics*, J. Differential Equations, **249**, 2010.

[4] M. Bostan, *Multi-scale analysis for linear first order PDEs. The finite Larmor radius regime*, SIAM J. Math. Anal., **48**, No. 3, pp.2133-2188, 2016.

[5] M. Bostan, *Gyro-kinetic Vlasov equation in three dimensional setting. Second order approximation*, SIAM J. Multiscale Model. Simul., **8**, No. 5, pp. 1923-1957, 2010.

[6] P. Chartier, A. Murua, J.-M. Sanz-Serna, *Erratum to: Higher-order averaging, formal series and numerical integration II: the quasi-periodic case*, Found Comput Math **17**, 2017.

[7] P. Chartier, A. Murua, J.-M. Sanz-Serna, Higher-order averaging, formal series and numerical integration III: error bounds, Found Comput Math **15**, 2015.

[8] P. Chartier, A. Murua, J.-M. Sanz-Serna, A formal series approach to averaging: exponentially small error estimates, Discrete and Continuous Dynamical Systems (DCDS-A), **32**, no. 9, 2012.

[9] P. Chartier, A. Murua, J.-M. Sanz-Serna, *Higher-order averaging, formal series and numerical integration II: the quasi-periodic case*, Found Comput Math **4**, 2012.

[10] P. Chartier, A. Murua, J.M. Sanz-Serna, *Higher-order averaging, formal series and numerical integration I: B-series*, Found Comput Math **10**, 2010.

[11] N. Crouseilles, S. Jin, M. Lemou, *Nonlinear Geometric Optics method based multi-scale numerical schemes for highly-oscillatory transport equations*, Math. Mod. Meth. App. Sc. **27**, pp. 2031-2070, 2017.

[12] P. Degond, F. Filbet, *On the asymptotic limit of the three dimensional Vlasov-Poisson system for large magnetic field : formal derivation*, J. Stat. Physics **165**, pp. 765-784, 2016.

[13] E. Frénod, P.-A. Raviart, E. Sonnendrücker, *Two scale expansion of a singularly perturbed convection equation*, J. Maths. Pures Appl., **80**, No 8, pp 815–843, 2001.

[14] E. Frénod, E. Sonnendrücker, *Long time behavior of the Vlasov equation with strong external magnetic field*, Math. Models Methods Appl. Sci. **10**, 539-553, 2000.

[15] E. Frénod, E. Sonnendrücker, *The finite Larmor radius approximation*, SIAM J. Math. Anal. **32**, 1227-1247, 2001.

[16] F. Golse, L. Saint-Raymond, *The Vlasov-Poisson system with strong magnetic field*, J. Math. Pures Appl. **78**, no. 8, 791-817, 1999.

[17] A. Murua, J.-M. Sanz-Serna, *Averaging and computing normal forms with word series algorithms*, Discrete Mechanics, Geometric Integration and Lie-Butcher Series: DMGILBS, Madrid, May 2015, Springer, 5 nov. 2018.

[18] A. Murua, J.-M. Sanz-Serna, *Computing normal forms and formal invariants of dynamical systems by means of word series*, Nonlinear Analysis **138**, 2016.

[19] A. Murua, J.-M. Sanz-Serna, *Word series for dynamical systems and their numerical integrators*, Found Comput Math **17**, 2017.

[20] L. M. Perko, *Higher order averaging and related methods for perturbed periodic and quasi-periodic systems*, SIAM J. Applied. Math. **17**, 698-724, 1969.

[21] J. A. Sanders, F. Verhulst, *Averaging methods in nonlinear dynamical systems*, Applied Mathematical Sciences, Vol. 59. Springer-Verlag, 1985.

E-mail address: Philippe.Chartier@inria.fr
E-mail address: Nicolas.Crouseilles@inria.fr
E-mail address: Mohammed.Lemou@univ-rennes1.fr
E-mail address: Florian.Mehats@univ-rennes1.fr