ESAIM: COCV 27 (2021) S19
https://doi.org/10.1051/cocv/2020069

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

OPTIMAL CONTROL OF THE TWO-DIMENSIONAL VLASOV-MAXWELL SYSTEM

Jörg Weber*

Abstract. The time evolution of a collisionless plasma is modeled by the Vlasov-Maxwell system which couples the Vlasov equation (the transport equation) with the Maxwell equations of electrodynamics. We only consider a two-dimensional version of the problem since existence of global, classical solutions of the full three-dimensional problem is not known. We add external currents to the system, in applications generated by coils, to control the plasma properly. After considering global existence of solutions to this system, differentiability of the control-to-state operator is proved. In applications, on the one hand, we want the shape of the plasma to be close to some desired shape. On the other hand, a cost term penalizing the external currents shall be as small as possible. These two aims lead to minimizing some objective function. We restrict ourselves to only such control currents that are realizable in applications. After that, we prove existence of a minimizer and deduce first order optimality conditions and the adjoint equation.

Mathematics Subject Classification. 49J20, 35Q61, 35Q83, 82D10.

Received October 9, 2019. Accepted October 15, 2020.

1. Introduction

1.1. The system

The time evolution of a collisionless plasma is modeled by the Vlasov-Maxwell system. Collisions among the plasma particles can be neglected if the plasma is sufficiently rarefied or hot. The particles only interact through electromagnetic fields created collectively. We only consider plasmas consisting of just one particle species, for example, electrons. This work can immediately be adapted to the case of several particle species. For the sake of simplicity, we choose units such that physical constants like the speed of light, the charge and rest mass of an individual particle are normalized to unity. Also, for simplicity, we do not consider material parameters, for example for modeling superconductors in a fusion reactor, that is to say permittivity and permeability, which would appear in the Maxwell equations. Allowing the particles to move at relativistic speeds, the three-dimensional Vlasov-Maxwell system (on whole space) is given by

$$\partial_t f + \widehat{p} \cdot \partial_x f + (E + \widehat{p} \times B) \cdot \partial_p f = 0, \tag{1.1a}$$

$$\partial_t E - \operatorname{curl}_x B = -j_f, \tag{1.1b}$$

Keywords and phrases: Relativistic Vlasov-Maxwell system, optimal control with PDE constraints, nonlinear partial differential equations, calculus of variations.

Department of Mathematics, University of Bayreuth, 95440 Bayreuth, Germany.

* Corresponding author: Joerg.Weber@uni-bayreuth.de

$$\partial_t B + \operatorname{curl}_x E = 0, \tag{1.1c}$$
$$\operatorname{div}_x E = \rho, \tag{1.1d}$$
$$\operatorname{div}_x B = 0, \tag{1.1e}$$
$$\rho_f = 4\pi \int f \,\mathrm{d}p, \tag{1.1f}$$
$$j_f = 4\pi \int \widehat{p} f \,\mathrm{d}p. \tag{1.1g}$$

Here, the Vlasov equation is (1.1a) and the Maxwell equations of electrodynamics are (1.1b) to (1.1e). Vlasov and Maxwell equations are coupled *via* (1.1f) and (1.1g) rendering the whole system nonlinear due to the product term $(E + \widehat{p} \times B) \cdot \partial_p f$. In particular, $f = f(t, x, p)$ denotes the density of the particles on phase space, and $E = E(t, x)$, $B = B(t, x)$ are the electromagnetic fields, whereby $t \in \mathbb{R}$, x, and $p \in \mathbb{R}^3$ stand for time, position in space, and momentum. The abbreviation $\widehat{p} = \frac{p}{\sqrt{1+|p|^2}}$ denotes the velocity of a particle with momentum p. Furthermore, some moments of f appear as source terms in the Maxwell equations, that is to say j_f and ρ_f which equal the current and charge density up to the constant 4π.

However, we have not readily explained the source term ρ in (1.1d). If we would demand $\operatorname{div}_x E = \rho_f$ this would lead to a seeming contradiction: Formally integrating this equation with respect to x (and assuming $E \to 0$ rapidly enough at ∞) leads to $\int \rho_f \,\mathrm{d}x = 0$ and hence $f = 0$ by $\mathring{f} \geq 0$. This problem is caused by our simplifying restriction to one species of particles and is resolved by adding some terms to ρ_f, for example a neutralizing background density, so that we have a total charge density ρ with vanishing space integral. However, usually this background density is neglected, see [6] for example.

Considering the Cauchy problem for the above system, we moreover demand

$$f(0, x, p) = \mathring{f}(x, p), E(0, x) = \mathring{E}(x), B(0, x) = \mathring{B}(x),$$

where $\mathring{f} \geq 0$, $\mathring{E}$, and $\mathring{B}$ are some given initial data.

Unfortunately, existence of global (*i.e.*, global in time), classical (*i.e.*, continuously differentiable) solutions for general (smooth) data is an open problem in the three-dimensional setting. It is only known that global weak solutions can be obtained. This was proved by R.J. Di Perna and P.L. Lions [1]. For a detailed insight concerning this matter we recommend the review article [14] by G. Rein. As for global existence of classical solutions, the strategy was to first consider lower dimensional settings. R. Glassey and J. Schaeffer proved global existence of classical solutions in the one and one-half [4], the two [6, 7], and the two and one-half dimensional setting [5].

Since it is convenient to have global existence of classical solutions on hand, we consider a two-dimensional version of the problem in this work. Notice that *mutatis mutandis* all results and techniques can be applied to the full three-dimensional setting once global existence of classical solutions has been proved. The restriction to 'two-dimensionality' is to be understood in the following sense: All functions shall be independent of the third variables x_3 and p_3. This new model describes a plasma where the particles only move in the (x_1, x_2)-plane, but the plasma extends in the x_3-direction infinitely. To ensure that these properties are preserved in time, we have to demand that the electric field lies in the plane and that the magnetic field is perpendicular to the plane so that $E = (E_1(t, x), E_2(t, x), 0)$ and $B = (0, 0, B(t, x))$. Here and in the following, let $x = (x_1, x_2)$ and $p = (p_1, p_2)$ be two-dimensional variables. Note that hence the magnetic field is always divergence free with respect to x, so that (1.1e) is always satisfied and will no longer be mentioned. The two-dimensional Vlasov-Maxwell system reads

$$\partial_t f + \widehat{p} \cdot \partial_x f + (E + (\widehat{p}_2, -\widehat{p}_1)B) \cdot \partial_p f = 0,$$
$$\partial_t E_1 - \partial_{x_2} B = -j_{f,1},$$
$$\partial_t E_2 + \partial_{x_1} B = -j_{f,2},$$

$$\begin{aligned}\partial_t B + \partial_{x_1} E_2 - \partial_{x_2} E_1 &= 0,\\ \operatorname{div}_x E &= \rho,\\ (f, E, B)|_{t=0} &= \left(\mathring{f}, \mathring{E}, \mathring{B}\right).\end{aligned}$$

The goal is to control the plasma properly. Thereto we add external currents U to the system, in applications generated by electric coils, that induce external electromagnetic fields affecting the plasma particles. These currents, like the electric field and the current density of the plasma particles, have to lie in the plane and have to be independent of the third space coordinate. Of course, there will be an external charge density ρ_{ext} corresponding to the external current. It is natural to assume local conservation of the external charge, *i.e.*,

$$\partial_t \rho_{\text{ext}} + \operatorname{div}_x U = 0.$$

Hence, we can eliminate ρ_{ext} *via*

$$\rho_{\text{ext}} = \mathring{\rho}_{\text{ext}} - \int_0^t \operatorname{div}_x U \,\mathrm{d}\tau.$$

The initial value $\mathring{\rho}_{\text{ext}}$ will be added to the background density, which is then neglected, as was already mentioned above.

In the following, we consider the controlled relativistic Vlasov-Maxwell system

$$\left.\begin{aligned}\partial_t f + \widehat{p}\cdot\partial_x f + \left(E - \widehat{p}^{\perp} B\right)\cdot\partial_p f &= 0,\\ \partial_t E + \nabla_x^{\perp} B &= -j_f - U,\\ \partial_t B + \operatorname{curl}_x E &= 0,\\ \operatorname{div}_x E &= \rho_f - \int_0^t \operatorname{div}_x U \,\mathrm{d}\tau,\\ (f, E, B)|_{t=0} &= \left(\mathring{f}, \mathring{E}, \mathring{B}\right)\end{aligned}\right\} \tag{CVM}$$

on a finite time interval $[0, T]$ with given $T > 0$; here we introduced the abbreviations $a^{\perp} = (-a_2, a_1)$ for $a \in \mathbb{R}^2$, $\nabla_x^{\perp} B = (-\partial_{x_2} B, \partial_{x_1} B)$, and the scalar curl operator $\operatorname{curl}_x E = \partial_{x_1} E_2 - \partial_{x_2} E_1$ in 2D.

It is well-known that L^q-norms (with respect to (x, p), $1 \le q \le \infty$) of f are preserved in time by f solving the Vlasov equation since the vector field $\left(\widehat{p}, E - \widehat{p}^{\perp} B\right)$ is divergence free in (x, p). Therefore, especially, the L^1-norm (with respect to x) of the charge density ρ_f is constant in time.

The outline of our work is the following: In the first part, we have to prove unique solvability of (CVM). Of course, some regularity assumptions on the external current and the initial data have to be made in order to prove existence of classical solutions. In the second part, we consider an optimal control problem. On the one hand, we want the shape of the plasma to be close to some desired shape. On the other hand, the external currents (the costs) shall be as small as possible. These two aims lead to minimizing some objective function. To analyze the optimal control problem, it is convenient to show differentiability of the control-to-state operator first. After that, we prove existence of a minimizer and deduce first order optimality conditions and the adjoint equation.

The steps mentioned above were carried out by P. Knopf [10] and, only considering realizable control fields, by Knopf and the author [11] for the three-dimensional Vlasov-Poisson system with an external magnetic field. The consideration of the latter setting has the advantage of being able to work in three dimensions, but has the disadvantage of only imposing Poisson's equation, that is, Maxwell's equations with an internal magnetic field sufficiently small to be neglected, for the electromagnetic fields, which make things easier due to the elliptic nature of Poisson's equation in contrast to the hyperbolic nature of the (time evolutionary) Maxwell equations.

Also other approaches for controlling a Vlasov-Maxwell plasma have been considered in the literature, but they are different in nature compared to our approach. We refer to [3, 13] and the references therein.

1.2. Some notation and simple computations

We denote by $B_r(x)$ the open ball with radius $r > 0$ and center $x \in X$ where X is a normed space. Furthermore, we abbreviate $B_r := B_r(0)$. For a function $g\colon [0,T] \times \mathbb{R}^j \to \mathbb{R}^k$ we abbreviate $g(t) := g(t,\cdot)\colon \mathbb{R}^j \to \mathbb{R}^k$ for $0 \le t \le T$. Also, we write $\operatorname{supp} g$ for the support of g, and $\operatorname{supp}_x g$ (and likewise $\operatorname{supp}_p g$) for the support of a function $g = g(t,x,p)$ with respect to x, that is, the closure of the set of all x such that there are t and p with $g(t,x,p) \neq 0$. Sometimes, denoting certain function spaces, we omit the set where these functions are defined. Which set is meant should be obvious, in fact the largest possible set like $[0,T] \times \mathbb{R}^j$ (including time) or $\mathbb{R}^j$ (not including time). Moreover, C_b^k denotes the space of k-times continuously differentiable functions (on a given set) such that all derivatives up to order k are bounded. The index c, as in C_c^k, indicates that such functions are compactly supported. Furthermore, $X \hookrightarrow Y$ means that X is continuously embedded in Y. Finally, we denote by $C > 0$ some generic constant that may change from line to line (also inside a line) and may depend on some quantities; we will clarify at the beginning of each section on what quantities C may depend in this section or write the dependence explicitly as $C(r)$, for example.

1.3. Maxwell equations

We will have to consider first order and second order Maxwell equations. It is well-known that they are equivalent and that the divergence equations propagate in time if local conservation of charge holds, *i.e.*,

$$\partial_t \rho + \operatorname{div}_x j = 0. \tag{LC}$$

In our two-dimensional setting with fields $(E_1, E_2, 0)$ and $(0, 0, B)$ we conclude:

Lemma 1.1. *Let $\mathring{E}$ and $\mathring{B}$ be of class C^2 and E, $B \in \mathrm{C}^2$, and ρ, $j \in \mathrm{C}^1$. If the conditions*

$$\operatorname{div}_x \mathring{E} = \rho(0) \tag{CC}$$

and

$$\partial_t \rho + \operatorname{div}_x j = 0 \tag{LC}$$

are satisfied, then the systems of first order Maxwell equations

$$\left.\begin{aligned} \partial_t E + \nabla_x^\perp B &= -j, \\ \partial_t B + \operatorname{curl}_x E &= 0, \\ (E,B)(0) &= \left(\mathring{E}, \mathring{B}\right), \end{aligned}\right\} \tag{1stME}$$

and second order Maxwell equations

$$\left.\begin{aligned} \partial_t^2 E - \Delta_x E &= -\partial_t j - \partial_x \rho, \\ E(0) &= \mathring{E}, \\ \partial_t E(0) &= -\nabla_x^\perp \mathring{B} - j(0), \\ \partial_t^2 B - \Delta_x B &= \operatorname{curl}_x j, \\ B(0) &= \mathring{B}, \\ \partial_t B(0) &= -\operatorname{curl}_x \mathring{E}, \end{aligned}\right\} \tag{2ndME}$$

are equivalent. Moreover, then also $\operatorname{div}_x E = \rho$ *globally in time.*

We give a quite general condition that guarantees (LC).

Lemma 1.2. *Let* $g \in \mathrm{C}^0$, *and* f, d, *and* K *of class* C^1 *with* $\operatorname{div}_p K = 0$ *and* $f(t, x, \cdot)$ *compactly supported for each* $t \in [0, T]$ *and* $x \in \mathbb{R}^2$. *Assume* $\partial_t f + \widehat{p} \cdot \partial_x f + K \cdot \partial_p f = g$ *and that* $\int g \,\mathrm{d}p = 0$ *holds. Then* $\rho = \rho_f - \int_0^t \operatorname{div}_x d \,\mathrm{d}\tau$ *and* $j = j_f + d$ *satisfy* (LC).

Proof. First, $\partial_t\left(-\int_0^t \operatorname{div}_x d\,\mathrm{d}\tau\right) + \operatorname{div}_x d = 0$ is obvious. Furthermore, integrating the Vlasov equation with respect to p instantly yields $\partial_t \rho_f + \operatorname{div}_x j_f = 0$. □

Since (2ndME) consists of Cauchy problems for wave equations, we will need a solution formula for the 2D wave equation. In two dimensions, the (in C^2 unique) solution of the Cauchy problem

$$\begin{aligned} \partial_t^2 u - \Delta_x u &= f, \\ u(0) &= g, \\ \partial_t u(0) &= h, \end{aligned}$$

is given by the well known formula

$$u(t,x) = \frac{1}{2\pi}\int_0^t \int_{|x-y|<t-\tau} \frac{f(\tau, y)}{\sqrt{(t-\tau)^2 - |x-y|^2}}\,\mathrm{d}y\mathrm{d}\tau + \frac{1}{2\pi}\int_{B_1} \frac{g(x+ty) + t\nabla g(x+ty)\cdot y + th(x+ty)}{\sqrt{1-|y|^2}}\,\mathrm{d}y$$

if the data are smooth.

1.4. Control space for classical solutions

In the following let $L > 0$,

$$U \in \mathcal{V}_L := \left\{ d \in \mathrm{W}^{2,1}\left(0, T; \mathrm{C}_b^4\left(\mathbb{R}^2; \mathbb{R}^2\right)\right) \mid d(t,x) = 0 \text{ for } |x| \geq L \right\},$$

and let $\mathcal{V}_L$ be equipped with the $\mathrm{W}^{2,1}\left(0, T; \mathrm{C}_b^4\left(\mathbb{R}^2; \mathbb{R}^2\right)\right)$-norm.

2. Existence results

2.1. Estimates on the fields

2.1.1. A generalized system

The most important tool to get certain bounds is to have representations of the fields. One can use the solution formula for the wave equation and after some transformation of the integral expressions Gronwall-like estimates on the density and the fields can be derived. These bounds, for instance, will imply that the sequences constructed in Section 2.3 converge in a certain sense. Having that in mind it is useful not to work with the system (CVM) but with a somewhat generalized one with second order Maxwell equations:

$$\left.\begin{aligned}
&\partial_t f + \widehat{p}\cdot\partial_x f + \alpha(p)K\cdot\partial_p f = g,\\
&\partial_t^2 E - \Delta_x E = -\partial_t j_f - \partial_t d - \partial_x\rho_f + \partial_x\int_0^t \operatorname{div}_x d\,\mathrm{d}\tau,\\
&\partial_t^2 B - \Delta_x B = \operatorname{curl}_x j_f + \operatorname{curl}_x d,\\
&(f,E,B)(0) = \left(\mathring{f},\mathring{E},\mathring{B}\right),\\
&\partial_t E(0) = -\nabla_x^\perp \mathring{B} - j_{\mathring{f}} - d(0),\\
&\partial_t B(0) = -\operatorname{curl}_x \mathring{E},
\end{aligned}\right\} \tag{GVM}$$

with initial data $\mathring{f}$ of class C^1_c and $\mathring{E}$, $\mathring{B}$ of class C^2_b. We assume that we already have functions f, K of class C^1, E, B of class C^2, g of class C^0_b, d of class $\mathrm{C}^1\left(0,T;\mathrm{C}^2_b\right)$ and α of class C^1_b satisfying (GVM). Furthermore, we assume that $\operatorname{div}_p K = 0$ and that there is a $r>0$ such that $f(t,x,p) = g(t,x,p) = 0$ if $|p| > r$.

2.1.2. Estimates on the density

Lemma 2.1. *The density f and its (x,p)-derivatives are estimated by*

i)

$$\|f(t)\|_\infty \le \left\|\mathring{f}\right\|_\infty + \int_0^t \|g(\tau)\|_\infty \mathrm{d}\tau$$

if $g\in\mathrm{C}^0$ and

ii)

$$\|\partial_{x,p} f(t)\|_\infty \le \left(\left\|\partial_{x,p}\mathring{f}\right\|_\infty + \int_0^t \|\partial_{x,p} g(\tau)\|_\infty \mathrm{d}\tau\right)\exp\left(\int_0^t \|\partial_{x,p}(\alpha K)(\tau)\|_\infty \mathrm{d}\tau\right)$$

if $g\in\mathrm{C}^1$.

Proof. This is easily proved by considering the characteristics $X = X(s,t,x,p)$, $P = P(s,t,x,p)$ of the Vlasov equation in (GVM), which are defined *via*

$$\dot{X} = \widehat{P},\ \dot{P} = \alpha(P)K(s,X,P)$$

with initial condition $(X,P)(t,t,x,p) = (x,p)$. Then

$$f(t,x,p) = \mathring{f}((X,P)(0,t,x,p)) + \int_0^t g(s,(X,P)(s,t,x,p))\mathrm{d}s$$

and, if $g\in\mathrm{C}^1$,

$$\begin{aligned}
\partial_{x,p} f(t,x,p) = \left(\partial_{x,p}\mathring{f}\right)((X,P)(0,t,x,p)) + \int_0^t (\partial_{x,p} g)(s,(X,P)(s,t,z))\mathrm{d}s\\
- \int_0^t (\partial_{x,p} f)(s,(X,P)(s,t,z))(\partial_{x,p}(\alpha K))(s,(X,P)(s,t,z))\mathrm{d}s;
\end{aligned}$$

see [12], Section 5. The asserted estimates are hence straightforwardly derived. □

The p-support condition on f is satisfied if $\operatorname{supp}\alpha \subset B_R$ for some $R > 0$: Obviously for $|p| > \max\{R, r, r_0\}$ (where $\operatorname{supp}_p \mathring{f} \subset B_{r_0}$) we have $\dot{P}(s,t,x,p) = 0$, hence $P(s,t,x,p) = p$ and therefore $\mathring{f}((X,P)(0,t,x,p)) = g(s,(X,P)(s,t,x,p)) = 0$.

2.1.3. *Representation of the fields*

In the following the constant C may depend on T, r, and α (*i.e.*, its C^1_b-norm), and we use the abbreviations

$$\xi = \frac{y-x}{t-\tau},\ es = \frac{-2(\xi + \widehat{p})}{1+\widehat{p}\cdot\xi},\ bs = \frac{-2\xi\cdot\widehat{p}^\perp}{1+\widehat{p}\cdot\xi},\ et = \frac{-2\left(1-|\widehat{p}|^2\right)(\xi+\widehat{p})}{(1+\widehat{p}\cdot\xi)^2},\ bt = \frac{-2\left(1-|\widehat{p}|^2\right)\xi\cdot\widehat{p}^\perp}{(1+\widehat{p}\cdot\xi)^2},$$

where $t,\tau \in [0,T]$, $x,y,p \in \mathbb{R}^2$.

We state some fundamental properties which will be used several times:

Remark 2.2. i) For $|p| \le r$ and $|\xi| \le 1$ we can estimate

$$|\partial_p(bs)|, |\partial_p(es)|, |\partial_p\partial_\xi(bs)|, |\partial_p\partial_\xi(es)|, |bt|, |et|, \left|\partial_{(\xi,p)}(bt)\right|, \left|\partial_{(\xi,p)}(et)\right|$$

by a constant $C(r) > 0$ only depending on r, since

$$|1+\widehat{p}\cdot\xi| \ge 1 - |\widehat{p}||\xi| \ge 1 - \frac{r}{\sqrt{1+r^2}} > 0.$$

ii) We compute

$$\int_{|x-y|<t-\tau} \frac{\mathrm{d}y}{\sqrt{(t-\tau)^2 - |x-y|^2}} = 2\pi \int_0^{t-\tau} s\left((t-\tau)^2 - s^2\right)^{-\frac{1}{2}} \mathrm{d}s = 2\pi(t-\tau)$$

and

$$\int_0^t \int_{|x-y|<t-\tau} \frac{\mathrm{d}y\mathrm{d}\tau}{(t-\tau)^{l+1}\sqrt{1-|\xi|^2}} = \int_0^t \int_{|x-y|<t-\tau} \frac{\mathrm{d}y\mathrm{d}\tau}{(t-\tau)^l\sqrt{(t-\tau)^2-|x-y|^2}} = 2\pi\int_0^t (t-\tau)^{-l+1}\mathrm{d}\tau$$
$$\le \frac{2\pi}{2-l}T^{2-l} = C(T,l) < \infty$$

for $l < 2$.

Now we can derive integral expressions for the fields E and B proceeding similarly to [6].

Lemma 2.3. *We have $E = E^0 + ES + ET + ED$ and $B = B^0 + BS + BT + BD$ where E^0, B^0 are functionals of the initial data and $d(0)$, and where*

$$ES_j = \int_0^t \int_{|x-y|<t-\tau} \int \frac{(\alpha\partial_p(es_j) + es_j\nabla\alpha)\cdot Kf + (es_j)g}{\sqrt{(t-\tau)^2-|x-y|^2}} \mathrm{d}p\mathrm{d}y\mathrm{d}\tau,$$

$$BS = \int_0^t \int_{|x-y|<t-\tau} \int \frac{(\alpha\partial_p(bs) + bs\nabla\alpha)\cdot Kf + (bs)g}{\sqrt{(t-\tau)^2-|x-y|^2}} \mathrm{d}p\mathrm{d}y\mathrm{d}\tau,$$

$$ET_j = \int_0^t \int_{|x-y|<t-\tau} \int \frac{et_j}{(t-\tau)\sqrt{(t-\tau)^2 - |x-y|^2}} f \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau,$$

$$BT = \int_0^t \int_{|x-y|<t-\tau} \int \frac{bt}{(t-\tau)\sqrt{(t-\tau)^2 - |x-y|^2}} f \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau,$$

$$ED_j = -\frac{1}{2\pi} \int_0^t \int_{|x-y|<t-\tau} \frac{\partial_t d_j - \int_0^\tau \partial_{x_j} \operatorname{div}_x d \,\mathrm{d}s}{\sqrt{(t-\tau)^2 - |x-y|^2}} \mathrm{d}y\mathrm{d}\tau,$$

$$BD = \frac{1}{2\pi} \int_0^t \int_{|x-y|<t-\tau} \frac{\operatorname{curl}_x d}{\sqrt{(t-\tau)^2 - |x-y|^2}} \mathrm{d}y\mathrm{d}\tau.$$

Furthermore, the estimate

$$\|E(t)\|_\infty + \|B(t)\|_\infty \leq C\left(\left\|\mathring{f}\right\|_\infty + \left\|\mathring{E}\right\|_{\mathrm{C}_b^1} + \left\|\mathring{B}\right\|_{\mathrm{C}_b^1} + \|d\|_{\mathrm{W}^{1,1}\left(0,T;\mathrm{C}_b^1\right)}\right) \tag{2.1}$$

$$+ C\int_0^t ((1 + \|K(\tau)\|_\infty)\|f(\tau)\|_\infty + \|g(\tau)\|_\infty)\mathrm{d}\tau \tag{2.2}$$

holds.
If additionally $\mathring{E}$, $\mathring{B} \in \mathrm{C}_c^0$, and d is compactly supported in x uniformly in t, so are also the fields.

Proof. The representation formula are derived in much the same way as in [6], Theorem 1, the only difference is that here the source terms g and d appear. The support assertion is an immediate consequence of the representation formula. Physically, this is a result of the fact that electromagnetic fields can not propagate faster than the speed of light. Furthermore, the remaining estimate is a consequence of Remark 2.2. □

Remark 2.4. If $f(t,x,\cdot)$ is compactly supported for every t, x, but not necessarily uniformly in t, x, nevertheless the fields are given by the formula above. For this, one does not need the uniformity. However, (2.1) can not be obtained in this situation.

2.1.4. First derivatives of the fields

The next step is to differentiate these representation formulas and deriving certain estimates. The method is similar to the previous one. The constant C may now only depend on T, r, the initial data (*i.e.*, their C_b^2-norms), and $\|\alpha\|_{\mathrm{C}_b^1}$.

Lemma 2.5. *If $g \in \mathrm{C}^1$ and $d \in \mathrm{W}^{2,1}\left(0,T;\mathrm{C}_b^3\right)$, then the derivatives of the S-, T-, and D-terms are given by*

$$\partial_{x_i} BS = \int_0^t \int_{|x-y|<t-\tau} \int \frac{(\alpha\partial_p(bs) + bs\nabla\alpha)\cdot(f\partial_{x_i}K + K\partial_{x_i}f) + bs\partial_{x_i}g}{\sqrt{(t-\tau)^2 - |x-y|^2}} \mathrm{d}p\mathrm{d}y\mathrm{d}\tau,$$

$$\partial_{x_i} BT = \int_0^t \int_{|x-y|<t-\tau} \int \frac{bt}{(t-\tau)\sqrt{(t-\tau)^2 - |x-y|^2}} \partial_{x_i} f \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau,$$

$$\partial_{x_i} BD = \frac{1}{2\pi} \int_0^t \int_{|x-y|<t-\tau} \frac{\partial_{x_i} \operatorname{curl}_x d}{\sqrt{(t-\tau)^2 - |x-y|^2}} \mathrm{d}y\mathrm{d}\tau,$$

$$\partial_{x_i} ES = \int_0^t \int_{|x-y|<t-\tau} \int \frac{(\alpha \partial_p(es) + es\nabla\alpha)\cdot(f\partial_{x_i}K + K\partial_{x_i}f) + es\partial_{x_i}g}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau,$$

$$\partial_{x_i} ET = \int_0^t \int_{|x-y|<t-\tau} \int \frac{et}{(t-\tau)\sqrt{(t-\tau)^2 - |x-y|^2}} \partial_{x_i} f \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau,$$

$$\partial_{x_i} ED = \frac{1}{2\pi} \int_0^t \int_{|x-y|<t-\tau} \frac{\partial_t\partial_{x_i} d - \int_0^\tau \partial_{x_i}\partial_x \operatorname{div}_x d \,\mathrm{d}s}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}y\mathrm{d}\tau,$$

$$\partial_t BS = \int_0^t \int_{|x-y|<t-\tau} \int \frac{(\alpha \partial_p(bs) + bs\nabla\alpha)\cdot(f\partial_t K + K\partial_t f) + bs\partial_t g}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau$$
$$+ \int_{|x-y|<t} \int \frac{(\alpha\partial_p(bs) + bs\nabla\alpha)|_{\tau=0}\cdot K(0)\mathring{f} + bs|_{\tau=0} g(0)}{\sqrt{t^2 - |x-y|^2}} \,\mathrm{d}p\mathrm{d}y,$$

$$\partial_t BT = \int_0^t \int_{|x-y|<t-\tau} \int \frac{bt}{(t-\tau)\sqrt{(t-\tau)^2 - |x-y|^2}} \partial_t f \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau + \int_{|x-y|<t} \int \frac{bt|_{\tau=0}}{t\sqrt{t^2 - |x-y|^2}} \mathring{f} \,\mathrm{d}p\mathrm{d}y,$$

$$\partial_t BD = \frac{1}{2\pi} \int_0^t \int_{|x-y|<t-\tau} \frac{\partial_t \operatorname{curl}_x d}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}y\mathrm{d}\tau + \frac{1}{2\pi} \int_{|x-y|<t} \frac{\operatorname{curl}_x d(0)}{\sqrt{t^2 - |x-y|^2}} \,\mathrm{d}y,$$

$$\partial_t ES = \int_0^t \int_{|x-y|<t-\tau} \int \frac{(\alpha \partial_p(es) + es\nabla\alpha)\cdot(f\partial_t K + K\partial_t f) + es\partial_t g}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau$$
$$+ \int_{|x-y|<t} \int \frac{(\alpha\partial_p(es) + es\nabla\alpha)|_{\tau=0}\cdot K(0)\mathring{f} + es|_{\tau=0} g(0)}{\sqrt{t^2 - |x-y|^2}} \,\mathrm{d}p\mathrm{d}y,$$

$$\partial_t ET = \int_0^t \int_{|x-y|<t-\tau} \int \frac{et}{(t-\tau)\sqrt{(t-\tau)^2 - |x-y|^2}} \partial_t f \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau + \int_{|x-y|<t} \int \frac{et|_{\tau=0}}{t\sqrt{t^2 - |x-y|^2}} \mathring{f} \,\mathrm{d}p\mathrm{d}y,$$

$$\partial_t ED = -\frac{1}{2\pi} \int_0^t \int_{|x-y|<t-\tau} \frac{\partial_t^2 d - \partial_x \operatorname{div}_x d}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}y\mathrm{d}\tau - \frac{1}{2\pi} \int_{|x-y|<t} \frac{\partial_t d_j(0)}{\sqrt{t^2 - |x-y|^2}} \,\mathrm{d}y.$$

Furthermore, the derivatives are estimated by

$$\begin{aligned}
&\|\partial_{t,x}E(t)\|_\infty + \|\partial_{t,x}B(t)\|_\infty \\
&\leq C(1 + \|K\|_\infty + \|f\|_\infty + \|g\|_\infty)(1 + \|K\|_\infty)^2 \left(1 + \ln_+\left(|||\partial_{x,p} f|||_{[0,t]}\right) + \int_0^t \|\partial_{t,x,p}K(\tau)\|_\infty \mathrm{d}\tau\right) \\
&\quad + C\int_0^t \|\partial_{t,x} g(\tau)\|_\infty \mathrm{d}\tau + C\|d\|_{\mathrm{W}^{2,1}(0,T;\mathrm{C}_b^3)}
\end{aligned}$$

if $\|K\|_\infty < \infty$. *Here* $|||a|||_{[0,t]} \coloneqq \sup_{0\leq\tau\leq t} \|a(\tau)\|_\infty$.

Proof. Similarly as before, this is proved by following [6], now considering Theorem 3 therein. □

2.2. *A priori* bounds on the support with respect to p

The most important property that is exploited later while showing global existence of a solution of (CVM), is to have *a priori* bounds on the p-support of f. This means: If we have a solution (f, E, B) of (CVM) on $[0, T[$ with $f \in \mathrm{C}^1$ and E, B of class C^2, we have to show that

$$P(t) \coloneqq \inf\{a > 0 \mid f(\tau, x, p) = 0 \text{ for all } |p| \geq a,\, 0 \leq \tau \leq t\} + 3$$

is bounded, *i.e.*, $P(t) \leq Q$ for $0 \leq t < T$ where $Q > 0$ is some constant only dependent on T, the initial data (*i.e.*, their C^1_b-norms and $P(0)$), L, and $\|U\|_{\mathcal{V}_L}$. In the following, the constant C may also only depend on these numbers. Note that, per definition, P is monotonically increasing and that $|f| \leq \left\|\mathring{f}\right\|_\infty$. Moreover, $P(t) < \infty$ for each $0 \leq t < T$ because we have an *a priori* estimate on the x-support of f *via* $\left|\dot{X}\right| \leq 1$, so that $\mathrm{supp}_x f \subset B_s$, and on the compact set $[0, t] \times \overline{B_s}$ the electromagnetic fields are bounded; hence the force field $E - \widehat{p}^\perp B$ is bounded there. Furthermore, (LC) holds by Lemma 1.2. Therefore, and with Remark 2.4 we have the representations of the fields as given in Lemma 2.3. Moreover, we can also demand that (f, E, B) solves

$$\left.\begin{aligned}
&\partial_t f + \widehat{p} \cdot \partial_x f + \left(E - \widehat{p}^\perp B\right) \cdot \partial_p f = 0,\\
&\partial_t^2 E - \Delta_x E = -\partial_t j_f - \partial_t U - \partial_x \rho_f + \partial_x \int_0^t \mathrm{div}_x\, U \,\mathrm{d}\tau,\\
&\partial_t^2 B - \Delta_x B = \mathrm{curl}_x\, j_f + \mathrm{curl}_x\, U,\\
&(f, E, B)(0) = \left(\mathring{f}, \mathring{E}, \mathring{B}\right),\\
&\partial_t E(0) = -\nabla_x^\perp \mathring{B} - j_{\mathring{f}} - U(0),\\
&\partial_t B(0) = -\,\mathrm{curl}_x\, \mathring{E}
\end{aligned}\right\} \qquad \text{(CVM2nd)}$$

instead of (CVM) since both systems are equivalent by Lemma 1.1.

We use the notation

$$\omega \coloneqq \frac{y - x}{|y - x|},\; a \wedge b \coloneqq a_1 b_2 - a_2 b_1,\; K \coloneqq E - \widehat{p}^\perp B$$

and follow [7].

2.2.1. *Energy estimates*

The key in [7] is a sample of estimates that follow from the local energy conservation law

$$\partial_t\left(\frac{1}{2}|E|^2 + \frac{1}{2}B^2 + 4\pi \int f\sqrt{1 + |p|^2}\mathrm{d}p\right) + \mathrm{div}_x\left(-BE^\perp + 4\pi \int fp\,\mathrm{d}p\right) = 0.$$

However, this equation is false in our situation due to the external currents U. But still we are able to prove an analogue of [7], Lemma 1:

Lemma 2.6. *Let $0 \leq R \leq T$. The estimates*

i)

$$\sup_{x\in\mathbb{R}^2} \int_{|y-x|<R} \left(\frac{1}{2}|E|^2 + \frac{1}{2}B^2 + 4\pi \int f\sqrt{1 + |p|^2}\mathrm{d}p\right)\mathrm{d}y \leq C,$$

ii)

$$\sup_{x\in\mathbb{R}^2}\int_0^t\int_{|y-x|=t-\tau+R}\left(\frac{1}{2}(E\cdot\omega)^2+\frac{1}{2}(B+\omega\wedge E)^2+4\pi\int f\sqrt{1+|p|^2}(1+\widehat{p}\cdot\omega)\mathrm{d}p\right)\mathrm{d}S_y\mathrm{d}\tau\le C,$$

iii)

$$\sup_{x\in\mathbb{R}^2}\int_{|y-x|<R}\rho_f^{\frac{3}{2}}\,\mathrm{d}y\le C,$$

iv)

$$\sup_{x\in\mathbb{R}^2}\int_{|y-x|<R}\left(\int\frac{f}{\sqrt{1+|p|^2}}\,\mathrm{d}p\right)^3\mathrm{d}y\le C$$

hold for all $t\in[0,T[$.

Proof. We split the electromagnetic fields into internal and external fields; precisely, they are defined by

$$\begin{aligned}\partial_t E_{\mathrm{int}}+\nabla_x^\perp B_{\mathrm{int}}&=-j_f,\\ \partial_t B_{\mathrm{int}}+\mathrm{curl}_x\,E_{\mathrm{int}}&=0,\\ (E_{\mathrm{int}},B_{\mathrm{int}})(0)&=\left(\mathring{E},\mathring{B}\right)\end{aligned}$$

and

$$\begin{aligned}\partial_t E_{\mathrm{ext}}+\nabla_x^\perp B_{\mathrm{ext}}&=-U,\\ \partial_t B_{\mathrm{ext}}+\mathrm{curl}_x\,E_{\mathrm{ext}}&=0,\\ (E_{\mathrm{ext}},B_{\mathrm{ext}})(0)&=0.\end{aligned}$$

Indeed, the existence of $(E_{\mathrm{ext}},B_{\mathrm{ext}})$ is guaranteed since the (time evolutionary) Maxwell equations form a linear, symmetric, hyperbolic system, see [9], Theorem I. Because of $U\in\mathcal{V}_L$ we have $E_{\mathrm{ext}},\ B_{\mathrm{ext}}\in\mathrm{C}^0\left(0,T;\mathrm{H}^3\right)\cap\mathrm{C}^1\left(0,T;\mathrm{H}^2\right)\subset\mathrm{C}^1$; furthermore

$$\|(E_{\mathrm{ext}},B_{\mathrm{ext}})(t)\|_\infty\le C\|(E_{\mathrm{ext}},B_{\mathrm{ext}})(t)\|_{\mathrm{H}^2}\le C\int_0^T\|U(\tau)\|_{\mathrm{H}^2}\mathrm{d}\tau\le C\|U\|_{\mathcal{V}_L}=C$$

by Sobolev's embedding theorem and the support condition on U. Because of the linearity of the Maxwell equations it holds that $E_{\mathrm{int}}:=E-E_{\mathrm{ext}}$ and $B_{\mathrm{int}}:=B-B_{\mathrm{ext}}$ solve their equations mentioned earlier and are of class C^1. Now let

$$e_{\mathrm{int}}:=\frac{1}{2}|E_{\mathrm{int}}|^2+\frac{1}{2}B_{\mathrm{int}}^2+4\pi\int f\sqrt{1+|p|^2}\mathrm{d}p$$

which is physically the energy density of the internal system and

$$e:=\frac{1}{2}|E|^2+\frac{1}{2}B^2+4\pi\int f\sqrt{1+|p|^2}\mathrm{d}p.$$

We have

$$\begin{aligned}
&\partial_t e_{\text{int}} + \operatorname{div}_x\left(-B_{\text{int}}E_{\text{int}}^{\perp} + 4\pi\int fp\,\mathrm{d}p\right)\\
&= E_{\text{int}}\cdot\partial_t E_{\text{int}} + B_{\text{int}}\partial_t B_{\text{int}} + 4\pi\int \partial_t f\sqrt{1+|p|^2}\mathrm{d}p + E_{\text{int},2}\partial_{x_1}B_{\text{int}} + B_{\text{int}}\partial_{x_1}E_{\text{int},2}\\
&\quad - E_{\text{int},1}\partial_{x_2}B_{\text{int}} - B_{\text{int}}\partial_{x_2}E_{\text{int},1} + 4\pi\int\partial_x f\cdot p\,\mathrm{d}p\\
&= -E_{\text{int}}\cdot j_f - 4\pi\int K\cdot\partial_p f\sqrt{1+|p|^2}\mathrm{d}p\\
&= -E_{\text{int}}\cdot j_f + 4\pi E\cdot\int f\partial_p\sqrt{1+|p|^2}\mathrm{d}p + 4\pi B\int f\operatorname{div}_p p^{\perp}\mathrm{d}p\\
&= E_{\text{ext}}\cdot j_f
\end{aligned}$$

where we made use of the respective Vlasov-Maxwell equations, $\partial_p\sqrt{1+|p|^2} = \widehat{p}$, and $\operatorname{div}_p p^{\perp} = 0$. We integrate this identity over a suitable set and arrive at

$$\begin{aligned}
&\int_0^t\int_{|y-x|<t-\tau+R} E_{\text{ext}}\cdot j_f\,\mathrm{d}y\mathrm{d}\tau = \int_0^t\int_{|y-x|<t-\tau+R}\left(\partial_\tau e_{\text{int}} + \operatorname{div}_y\left(-B_{\text{int}}E_{\text{int}}^{\perp} + 4\pi\int fpdp\right)\right)\mathrm{d}y\mathrm{d}\tau\\
&= -\int_{|y-x|<t+R} e_{\text{int}}(0,y)\mathrm{d}y + \int_{|y-x|<R} e_{\text{int}}(t,y)\mathrm{d}y\\
&\quad + \frac{1}{\sqrt{2}}\int_0^t\int_{|y-x|=t-\tau+R}\left(e_{\text{int}} + \omega\cdot\left(-B_{\text{int}}E_{\text{int}}^{\perp} + 4\pi\int fpdp\right)\right)dS_y\mathrm{d}\tau
\end{aligned}\tag{2.3}$$

after an integration by parts in (τ, y). The integrand of the last integral is non-negative because of

$$\begin{aligned}
0 \le d_{\text{int}} &:= \frac{1}{2}(E_{\text{int}}\cdot\omega)^2 + \frac{1}{2}(B_{\text{int}} + \omega\wedge E_{\text{int}})^2 + 4\pi\int f\sqrt{1+|p|^2}(1+\widehat{p}\cdot\omega)\mathrm{d}p\\
&= \frac{1}{2}E_{\text{int},1}^2\omega_1^2 + \frac{1}{2}E_{\text{int},2}^2\omega_2^2 + \frac{1}{2}B_{\text{int}}^2 + B_{\text{int}}\omega_1 E_{\text{int},2} - B_{\text{int}}\omega_2 E_{\text{int},1} + \frac{1}{2}E_{\text{int},2}^2\omega_1^2\\
&\quad + \frac{1}{2}E_{\text{int},1}^2\omega_2^2 + 4\pi\int f\sqrt{1+|p|^2}\mathrm{d}p + \omega\cdot 4\pi\int fp\,\mathrm{d}p\\
&= \frac{1}{2}E_{\text{int},1}^2 + \frac{1}{2}E_{\text{int},2}^2 + \frac{1}{2}B_{\text{int}}^2 + 4\pi\int f\sqrt{1+|p|^2}\mathrm{d}p + \omega_1 B_{\text{int}}E_{\text{int},2} - \omega_2 B_{\text{int}}E_{\text{int},1} + \omega\cdot 4\pi\int fp\,\mathrm{d}p\\
&= e_{\text{int}} + \omega\cdot\left(-B_{\text{int}}E_{\text{int}}^{\perp} + 4\pi\int fp\,\mathrm{d}p\right);
\end{aligned}\tag{2.4}$$

note that $1 + \widehat{p}\cdot\omega \ge 1 - 1\cdot 1 = 0$ and $|\omega| = 1$. The left hand side of (2.3) has to be investigated. The external fields are bounded by C, hence

$$\left|\int_0^t\int_{|y-x|<t-\tau+R} E_{\text{ext}}\cdot j_f\,\mathrm{d}y\mathrm{d}\tau\right| \le C\int_0^t \|j_f(\tau)\|_{\mathrm{L}^1}\mathrm{d}\tau \le C\int_0^t\|\rho_f(\tau)\|_{\mathrm{L}^1}\mathrm{d}\tau \le C \tag{2.5}$$

since the L^1-norm of ρ_f is constant in time.

Now we can prove the assertions using (2.3), (2.4), and (2.5):

i) We have

$$\int_{|y-x|<R} e_{\mathrm{int}}\,\mathrm{d}y \le \int_{|y-x|<t+R} e_{\mathrm{int}}(0,y)\mathrm{d}y + C \le C(R+t)^2 + C \le C$$

since $t, R \le T$. Together with

$$e \le 2e_{\mathrm{int}} + |E_{\mathrm{ext}}|^2 + |B_{\mathrm{ext}}|^2 \le 2e_{\mathrm{int}} + C$$

we conclude

$$\int_{|y-x|<R} e\,\mathrm{d}y \le C + CR^2 \le C.$$

ii) Similarly,

$$\int_0^t \int_{|y-x|=t-\tau+R} d_{\mathrm{int}}\,dS_y \mathrm{d}\tau \le \sqrt{2}\int_{|y-x|<t+R} e_{\mathrm{int}}(0,y)\mathrm{d}y + C \le C$$

and

$$d := \frac{1}{2}(E\cdot\omega)^2 + \frac{1}{2}(B+\omega\wedge E)^2 + 4\pi\int f\sqrt{1+|p|^2}(1+\widehat{p}\cdot\omega)\mathrm{d}p \le 2d_{\mathrm{int}} + 2|E_{\mathrm{ext}}|^2 + |B_{\mathrm{ext}}|^2 \le 2d_{\mathrm{int}} + C$$

yield

$$\int_0^t \int_{|y-x|=t-\tau+R} d\,dS_y \mathrm{d}\tau \le C + Ct(t+R)^2 \le C.$$

iii) For $r > 0$ it holds that

$$\rho_f = 4\pi\int f\,\mathrm{d}p = 4\pi\int_{|p|<r} f\,\mathrm{d}p + 4\pi\int_{|p|\ge r} f\,\mathrm{d}p \le Cr^2 + 4\pi r^{-1}\int_{|p|\ge r} f\sqrt{1+|p|^2}\mathrm{d}p \le C\big(r^2 + r^{-1}e\big).$$

Now choose $r := e^{\frac{1}{3}} > 0$ to derive $\rho_f \le Ce^{\frac{2}{3}}$ (if $e = 0$ then also $\rho_f = 0$) and hence

$$\int_{|y-x|<R} \rho_f^{\frac{3}{2}}\,\mathrm{d}y \le C\int_{|y-x|<R} e\,\mathrm{d}y \le C.$$

iv) Similarly,

$$\int \frac{f}{\sqrt{1+|p|^2}}\,\mathrm{d}p \le C\int_{|p|<r}\frac{1}{\sqrt{1+|p|^2}}\,\mathrm{d}p + \frac{1}{1+r^2}\int_{|p|\ge r} f\sqrt{1+|p|^2}\mathrm{d}p \le C\int_0^r \frac{s}{\sqrt{1+s^2}}\,\mathrm{d}s + \frac{1}{r^2}e$$
$$\le C\big(r + r^{-2}e\big) \le Ce^{\frac{1}{3}}$$

for again $r := e^{\frac{1}{3}}$ which yields

$$\int_{|y-x|<R}\left(\int \frac{f}{\sqrt{1+|p|^2}}\,\mathrm{d}p\right)^3 \mathrm{d}y \le C\int_{|y-x|<R} e\,\mathrm{d}y \le C.$$

□

2.2.2. Estimates on the fields

The crucial problem is to estimate the fields properly. To this end, we use the representation formula stated in Lemma 2.3. Unfortunately, the estimates there can not be applied because, of course, we can not assume that $P(t)$ is already bounded.

Lemma 2.7. *We have*

$$\begin{aligned}|ES_1| + |ES_2| + |BS| &\le CP(t)\ln P(t) + C\int_0^t (\|E(\tau)\|_\infty + \|B(\tau)\|_\infty)\mathrm{d}\tau,\\ |ET_1| + |ET_2| + |BT| &\le CP(t)\ln P(t),\\ |ED|, |BD| &\le C\|U\|_{\mathrm{W}^{1,1}\left(0,T;\mathrm{C}_b^2\right)} \le C,\\ \left|E^0\right|, \left|B^0\right| &\le C.\end{aligned}$$

Proof. The estimates on the S-and T-terms are derived in much the same way as in [7], Section 2. Note that the energy estimates of Lemma 2.6, that had to be modified in our situation, are enough to carry out the proofs therein.

The estimate on the D-terms is derived straightforwardly, as well as the estimate on E^0, B^0, the latter parts only containing terms of the initial data and $U(0)$. □

Now we can finally prove:

Lemma 2.8. *The a priori bound $P(t) \le Q$ holds, where Q only depends on T, the C_b^1-norms of the initial data,* $\operatorname{supp}_p \mathring{f}$ *(which basically coincides with $P(0)$), L, and $\|U\|_{\mathcal{V}_L}$.*

Proof. Collecting all bounds on the fields we arrive at

$$\|E(t)\|_\infty + \|B(t)\|_\infty \le C + CP(t)\ln P(t) + C\int_0^t (\|E(\tau)\|_\infty + \|B(\tau)\|_\infty)\mathrm{d}\tau.$$

As in [7], this is enough to show that

$$P(t) \le C + C\int_0^t P(s)\ln P(s)\mathrm{d}s,$$

from which the assertion follows immediately. □

2.3. Existence of classical solutions

2.3.1. The iteration scheme

In the following we want to construct a solution of (CVM). We will only sketch the main ideas, since similar procedures have already been carried out in the literature, see for example [8], Section V.

We work with initial data $\mathring{f} \geq 0$ of class C^2_c, $\mathring{E}$, $\mathring{B}$ of class C^3_b, and control $U \in \mathcal{V}_L$ that satisfy (CC), *i.e.*, $\operatorname{div} \mathring{E} = \rho_{\mathring{f}}$. We have to approximate these functions, so let $\mathring{f}_k \to \mathring{f}$ in C^2_b, $\mathring{E}_k \to \mathring{E}$ and $\mathring{B}_k \to \mathring{B}$ in C^3_b with $\mathring{f}_k \in \mathrm{C}^\infty_c$, $\mathring{E}_k$, $\mathring{B}_k \in \mathrm{C}^\infty$, and furthermore $U_k \to U$ in $\mathcal{V}_L$ with $U_k \in \mathrm{C}^\infty$ (note that C^∞ is dense in $\mathcal{V}_L$).

The strategy to obtain a solution of (CVM) is the following: By iteration we construct densities f_k and fields E_k, B_k in such a way that these functions will converge in a proper sense and that we may pass to the limit in (CVM). However, it is more convenient to work with a modified system. As the previous section suggests, it is crucial to control the p-support of f. For this reason we first consider a cut-off system on $[0,T]$ where we modify the original Vlasov equation and use the second order Maxwell equations ((CC) and (LC) need not hold for the iterates):

$$\left.\begin{aligned}
&\partial_t f + \widehat{p}\cdot\partial_x f + \alpha(p)\big(E - \widehat{p}^\perp B\big)\cdot\partial_p f = 0,\\
&\partial_t^2 E - \Delta_x E = -\partial_t j_f - \partial_t U - \partial_x \rho_f + \partial_x \int_0^t \operatorname{div}_x U \,\mathrm{d}\tau,\\
&\partial_t^2 B - \Delta_x B = \operatorname{curl}_x j_f + \operatorname{curl}_x U,\\
&(f,E,B)(0) = \Big(\mathring{f}, \mathring{E}, \mathring{B}\Big),\\
&\partial_t E(0) = -\nabla_x^\perp \mathring{B} - j_{\mathring{f}} - U(0),\\
&\partial_t B(0) = -\operatorname{curl}_x \mathring{E}.
\end{aligned}\right\} \tag{αVM}$$

Here, let the cut-off function α be of class $\mathrm{C}^\infty_c\big(\mathbb{R}^2\big)$ with $\alpha(p) = 1$ for $|p| \leq 2Q$. The property of the constant Q will imply that a solution of (αVM) is also a solution of (CVM).

We start the iteration with $f_0(t,x,p) := \mathring{f}_0(x,p)$, $E_0(t,x) := \mathring{E}_0(x)$, $B_0(t,x,p) := \mathring{B}_0(x)$. The induction hypothesis is that f_k, E_k, and B_k are of class C^∞ and that the fields are bounded. Given f_{k-1}, E_{k-1}, and B_{k-1}, we firstly define f_k as the solution of

$$\begin{aligned}
\partial_t f_k + \widehat{p}\cdot\partial_x f_k + \alpha(p)\big(E_{k-1} - \widehat{p}^\perp B_{k-1}\big)\cdot\partial_p f_k &= 0,\\
f_k(0) &= \mathring{f}_k,
\end{aligned}$$

namely

$$f_k(t,x,p) = \mathring{f}_k(X_k(0,t,x,p), P_k(0,t,x,p))$$

with the characteristics $X_k = X_k(s,t,x,p)$, $P_k = P_k(s,t,x,p)$ defined by

$$\begin{aligned}
\dot{X}_k &= \widehat{P}_k, & X_k(t,t,x,p) &= x,\\
\dot{P}_k &= \alpha(P_k)\Big(E_{k-1} - \widehat{P}_k^\perp B_{k-1}\Big)(s, X_k), & P_k(t,t,x,p) &= p,
\end{aligned}$$

the dot referring to differentiation with respect to the first variable s. We conclude that X_k and P_k are of class C^∞ in all four variables by the induction hypothesis. This yields that even $f_k \in \mathrm{C}^\infty$. Since α is compactly supported the p-support of f_k is controlled by a constant C. Hence, ρ_{f_k} and j_{f_k} are well-defined as $\mathrm{C}^\infty \cap \mathrm{C}^1_b$-functions.

Secondly, we define E_k and B_k as the solution of

$$\begin{aligned}
&\partial_t^2 E_k - \Delta_x E_k = -\partial_t j_{f_k} - \partial_t U_k - \partial_x \rho_{f_k} + \partial_x \int_0^t \operatorname{div}_x U_k \,\mathrm{d}\tau,\\
&\partial_t^2 B_k - \Delta_x B_k = \operatorname{curl}_x j_{f_k} + \operatorname{curl}_x U_k,\\
&(E_k, B_k)(0) = \Big(\mathring{E}_k, \mathring{B}_k\Big),
\end{aligned}$$

$$\partial_t E_k(0) = -\nabla_x^\perp \mathring{B}_k - j_{\mathring{f}_k} - U_k(0),$$
$$\partial_t B_k(0) = -\operatorname{curl}_x \mathring{E}_k.$$

Indeed, we can solve these wave equations by applying the solution formula for the wave equation. Since the right-hand sides of the above equations are of class C^∞ and bounded, so are also E_k and B_k. Applying Lemmas 2.1, 2.3, and 2.5 then shows that the iterates are bounded in C_b^1.

As for the second derivatives, we differentiate (αVM) and have, for example,

$$\left.\begin{aligned}
&\partial_t\partial_{x_i} f_k + \widehat{p}\cdot\partial_x\partial_{x_i} f_k \\
&\quad +\alpha K_{k-1}\cdot\partial_p\partial_{x_i} f_k = -\alpha\partial_{x_i} K_{k-1}\cdot\partial_p f_k,\\
&\partial_t^2\partial_{x_i} E_k - \Delta_x\partial_{x_i} E_k = -\partial_t j_{\partial_{x_i} f_k} - \partial_t\partial_{x_i} U_k - \partial_x \rho_{\partial_{x_i} f_k} + \partial_x \int_0^t \operatorname{div}_x \partial_{x_i} U_k \,\mathrm{d}\tau,\\
&\partial_t^2\partial_{x_i} B_k - \Delta_x\partial_{x_i} B_k = \operatorname{curl}_x j_{\partial_{x_i} f_k} + \operatorname{curl}_x \partial_{x_i} U_k,\\
&(\partial_{x_i} f_k, \partial_{x_i} E_k, \partial_{x_i} B_k)(0) = \left(\partial_{x_i}\mathring{f}_k, \partial_{x_i}\mathring{E}_k, \partial_{x_i}\mathring{B}_k\right),\\
&\partial_t\partial_{x_i} E_k(0) = -\nabla_x^\perp\partial_{x_i}\mathring{B}_k - j_{\partial_{x_i}\mathring{f}_k} - \partial_{x_i} U_k(0),\\
&\partial_t\partial_{x_i} B_k(0) = -\operatorname{curl}_x \partial_{x_i}\mathring{E}_k
\end{aligned}\right\} \tag{2.6}$$

and then apply the estimates of Lemmas 2.3 and 2.5. Note that for this we need four space derivatives in the definition of $\mathcal{V}_L$ so that $\|\partial_x U_k\|_{\mathrm{W}^{2,1}\left(0,T;\mathrm{C}_b^3\right)}$ is bounded. Likewise, one proceeds with the other second order derivatives. Altogether, the iterates are bounded in C_b^2.

After that, considering the difference of the iterates of the k-th step and the l-th step, Lemmas 2.1, 2.3, and 2.5 yield that the iteration sequences are even Cauchy sequences in C_b^1, so that they converge to some (f, E, B) in the C_b^1-norm.

For later considerations it will be convenient that the density and the fields are even C_b^2. Since all second derivatives are bounded in $\mathrm{L}^\infty\left([0,T]\times\mathbb{R}^j\right)$ ($j = 4$ or 2 respectively) they converge, after extracting a suitable subsequence, in the weak-*-sense. Of course, these limits have to be the respective weak derivatives of f, E, and B. The remaining part is to show that the weak derivatives just obtained are in fact classical ones. For this sake, have a look at the representation formula for $\partial_{x_i}\partial_{x_j} B_k$; use system (2.6) and Lemma 2.5:

$$\begin{aligned}
&\partial_{x_i}\partial_{x_j} B_k - \partial_{x_i}\overline{B}_k^0\\
&= \int_0^t \int_{|x-y|<t-\tau} \int \frac{bt}{(t-\tau)\sqrt{(t-\tau)^2 - |x-y|^2}} \partial_{x_i}\partial_{x_j} f_k \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau\\
&\quad + \int_0^t \int_{|x-y|<t-\tau} \int \frac{(\alpha\partial_p(bs) + bs\nabla\alpha)\cdot\partial_{x_j} f_k \partial_{x_i} K_{k-1}}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau\\
&\quad + \int_0^t \int_{|x-y|<t-\tau} \int \frac{(\alpha\partial_p(bs) + bs\nabla\alpha)\cdot K_{k-1}\partial_{x_i}\partial_{x_j} f_k}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau\\
&\quad - \int_0^t \int_{|x-y|<t-\tau} \int \frac{(bs)\alpha\partial_{x_i}\partial_{x_j} K_{k-1}\cdot\partial_p f_k}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau\\
&\quad - \int_0^t \int_{|x-y|<t-\tau} \int \frac{(bs)\alpha\partial_{x_j} K_{k-1}\cdot\partial_{x_i}\partial_p f_k}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau + \frac{1}{2\pi}\int_0^t \int_{|x-y|<t-\tau} \frac{\operatorname{curl}_x \partial_{x_i} U_k}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}y\mathrm{d}\tau.
\end{aligned}$$

Here, $\overline{B}_k^0$ is the 'B^0' of system (2.6) and converges to the respective expression without indices.

We are allowed to pass to the limit in the integral expressions because all kernels are integrable, (f_k, E_k, B_k) converge in C_b^1, the second derivatives weak-* in L^∞, and U_k in $\mathcal{V}_L$. Hence, we can omit the indices in the equation above or equivalently

$$\begin{aligned}
&\partial_{x_i}\partial_{x_j} B - \partial_{x_i}\overline{B}^0 \\
&= \int_0^t \int_{|x-y|<t-\tau} \int \frac{bt}{(t-\tau)\sqrt{(t-\tau)^2 - |x-y|^2}} \partial_{x_i}\partial_{x_j} f \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau \\
&\quad + \int_0^t \int_{|x-y|<t-\tau} \int \frac{(\alpha\partial_p(bs) + bs\nabla\alpha)\cdot\partial_{x_i}\big(K\partial_{x_j} f\big)}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau \\
&\quad - \int_0^t \int_{|x-y|<t-\tau} \int \frac{(bs)\alpha\partial_{x_i}\big(\partial_{x_j}K\cdot\partial_p f\big)}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}p\mathrm{d}y\mathrm{d}\tau + \frac{1}{2\pi}\int_0^t \int_{|x-y|<t-\tau} \frac{\operatorname{curl}_x \partial_{x_i} U}{\sqrt{(t-\tau)^2 - |x-y|^2}} \,\mathrm{d}y\mathrm{d}\tau
\end{aligned}$$

and conclude that $\partial_{x_i}\partial_{x_j} B$ is continuous which is an immediate consequence of $U \in \mathcal{V}_L$ and the following lemma:

Lemma 2.9. *Denote $M := \{(s,z) \in [0,T]\times\mathbb{R}^n \mid 0 \le s \le T,\ |z| < s\}$ and let $h \in \mathrm{C}^0([0,T]\times\mathbb{R}^{n+m})$ with uniform support in $p \in \mathbb{R}^m$, i.e., $\operatorname{supp}_p h \subset B_r$ for some $r > 0$, and let $w \in \mathrm{C}^1(M \times B_r)$ and $\gamma \in \{t, x_1, \dots x_n\}$. Furthermore, let one of the following options hold:*

i) $h \in \mathrm{W}^{1,\infty}([0,T]\times\mathbb{R}^{n+m})$ and $w \in \mathrm{L}^1(M\times B_r)$,
ii) $h \in \mathrm{W}^{1,1}(0,T;\mathrm{L}^\infty(\mathbb{R}^{n+m}))$ if $\gamma = t$ or $h \in \mathrm{L}^\infty\big(0,T;\mathrm{W}^{1,\infty}(\mathbb{R}^{n+m})\big)$ if $\gamma = x_i$ respectively, and

$$\int_{s-d<|z|<s} \int_{B_r} |w(s,z,p)|\mathrm{d}p\mathrm{d}z \to 0$$

for $d \to 0$ uniformly in $s \in [0,T]$.

Then

$$\begin{aligned}
H_\gamma(t,x) &:= \int_0^t \int_{|x-y|<t-\tau} \int (\partial_\gamma h)(\tau,y,p) w(t-\tau, y-x, p)\mathrm{d}p\mathrm{d}y\mathrm{d}\tau \\
&= \int_0^t \int_{|z|<s} \int (\partial_\gamma h)(t-s, x+z, p) w(s,z,p)\mathrm{d}p\mathrm{d}z\mathrm{d}s
\end{aligned}$$

is continuous in $(t,x) \in [0,T]\times\mathbb{R}^n$.

Proof. Let $\gamma = x_i$ and $\epsilon > 0$ be given. For $(t,x) \in [0,T]\times\mathbb{R}^n$ and $d > 0$ define

$$I_d(t,x) := \int_0^t \int_{s-d<|z|<s} \int (\partial_{x_i} h)(t-s, x+z, p) w(s,z,p)\mathrm{d}p\mathrm{d}z\mathrm{d}s$$

and estimate in case i)

$$|I_d(t,x)| \le \|\partial_{x_i} h\|_\infty \int_0^T \int_{s-d<|z|<s} \int_{B_r} |w(s,z,p)|\mathrm{d}p\mathrm{d}z\mathrm{d}s \to 0$$

and in case ii)

$$|I_d(t,x)| \leq \int_0^T \|\partial_{x_i} h(s)\|_\infty \mathrm{d}s \left\| s \mapsto \int_{s-d<|z|<s} \int_{B_r} |w(s,z,p)| \mathrm{d}p\mathrm{d}z \right\|_\infty \to 0$$

for $d \to 0$ uniformly in (t,x). Thus, we can choose d so that $|I_d(t,x)| < \frac{\epsilon}{4}$ for all (t,x). For now fixed d consider the remaining integral and integrate by parts

$$\begin{aligned}
J_d(t,x) &:= \int_0^t \int_{|z|<s-d} \int (\partial_{x_i} h)(t-s, x+z, p) w(s,z,p) \mathrm{d}p\mathrm{d}z\mathrm{d}s \\
&= \int_0^t \int_{|z|<s-d} \int (\partial_{z_i} h)(t-s, x+z, p) w(s,z,p) \mathrm{d}p\mathrm{d}z\mathrm{d}s \\
&= -\int_0^t \int_{|z|<s-d} \int h(t-s, x+z, p) \partial_{z_i} w(s,z,p) \mathrm{d}p\mathrm{d}z\mathrm{d}s \\
&\quad + \int_0^t \int_{|z|=s-d} \int h(t-s, x+z, p) w(s,z,p) \frac{1}{\sqrt{2}} \mathrm{d}p dS_z \mathrm{d}s + \int_{|z|<t-d} \int h(0, x+z, p) w(t,z,p) \mathrm{d}p\mathrm{d}z.
\end{aligned}$$

This is allowed because the integration domain is away from the possibly singular set $|z| = s$. For that very reason J_d is obviously continuous by the standard theorem for parameter integrals, so if $(\delta t, \delta x)$ is small enough (with $t + \delta t \in [0,T]$) we have

$$|J_d(t+\delta t, x+\delta x) - J_d(t,x)| < \frac{\epsilon}{2}.$$

Finally with $H_\gamma = I_d + J_d$ we conclude

$$|H_\gamma(t+\delta t, x+\delta x) - H_\gamma(t,x)| \leq |I_d(t+\delta t, x+\delta x)| + |I_d(t,x)| + |J_d(t+\delta t, x+\delta x) - J_d(t,x)| < \epsilon.$$

Analogously, one proves the assertion for $\gamma = t$. □

This lemma is applicable since f has uniform support in p, $\partial_x f$, $\partial_p f$, and $\partial_x K$ are of class $\mathrm{W}^{1,\infty}$, $|bs|$, $|bt| \leq C(r)$, and by Remark 2.2. Next, we have a representation formula for $\partial_t \partial_{x_j} B_k$ according to Lemma 2.5. Analogously we conclude that $\partial_t \partial_{x_j} B$ is continuous. For this, note that the terms without an $\int_0^t$-integral are easy to handle since there only initial values appear.

The procedure for E is nearly the same. The only critical point is to ensure that

$$\int_0^t \int_{|x-y|<t-\tau} \frac{\partial_t^2 \partial_{x_j} U}{\sqrt{(t-\tau)^2 - |x-y|^2}} \mathrm{d}y\mathrm{d}\tau$$

is continuous for $U \in \mathcal{V}_L$. To this end, we can apply Lemma 2.9 with $h = \partial_t \partial_{x_j} U \chi$ where $\chi = \chi(p) \in \mathrm{C}_c^\infty(\mathbb{R}^2)$ with $\int \chi \,\mathrm{d}p = 1$. Note that $\partial_t \partial_{x_j} U$ is continuous and of class $\mathrm{W}^{1,1}(0,T;\mathrm{L}^\infty)$ by $U \in \mathcal{V}_L$, and that

$$\int_{s-d<|z|<s} \frac{1}{\sqrt{s^2 - |z|^2}} \mathrm{d}z = 2\pi \sqrt{2sd - d^2} 1_{s \geq d} \leq 2\pi \sqrt{T} \sqrt{d},$$

where $1_{s \geq d}$ denotes the indicator function of the set $\{s \mid s \geq d\}$. So there only remain the ∂_t^2-derivatives of E and B. By the known convergence, we can pass to the limit in (αVM) so that the Vlasov equation holds everywhere and the Maxwell equations almost everywhere. With this knowledge and the just proven fact that the second space derivatives of the fields are continuous, we conclude that also the ∂_t^2-derivatives are continuous.

Now the fact that all weak derivatives are continuous instantly implies that they are classical ones. Therefore, the fields are of class C^2. Thus, the characteristics

$$\dot{X} = \widehat{P},\ \dot{P} = \alpha(P)\Big(E - \widehat{P}^\perp B\Big)(s, X),\ (X, P)(t, t, x, p) = (x, p)$$

are well defined and of class C^2 in (t, x, p). Hence,

$$f(t, x, p) = \mathring{f}((X, P)(0, t, x, p))$$

is also of class C^2.

Therefore, we are able to pass to the limit in (αVM), but actually (CVM) is to be solved: Obviously, (αVM) coincides with (CVM2nd) as long as f vanishes for $|p| \geq Q$. But this property is guaranteed by Lemma 2.8. Therefore, (f, E, B) is a solution of (CVM2nd) and hence of (CVM) by equivalence.

We collect some properties of (f, E, B):

Theorem 2.10. *Let $T, L > 0$, $\mathring{f} \geq 0$ of class C_c^2, $\mathring{E}$, $\mathring{B}$ of class C_b^3, and $U \in \mathcal{V}_L$ that satisfy $\operatorname{div} \mathring{E} = \rho_{\mathring{f}}$. Then there is a solution (f, E, B) of (CVM) on $[0, T]$ with:*

i) f, E, and B are of class C^2,
ii) f vanishes for $|p| \geq Q$ or $|x| \geq R + T$ (where Q only depends on T, the initial data (their C_b^1-norms and $P(0)$), and $\|U\|_{\mathcal{V}_L}$, and where $\operatorname{supp}_x \mathring{f} \subset B_R$),
iii) E, B vanish for $|x| \geq \widetilde{R} + L + R + T$ if their initial data are compactly supported, i.e., $\operatorname{supp} \mathring{E}$, $\operatorname{supp} \mathring{B} \subset B_{\widetilde{R}}$,
iv) the C_b^2-norms of the solution are estimated by a constant only depending on T, the initial data (their C_b^2-norms and $P(0)$), L, and $\|U\|_{\mathcal{V}_L}$.

Proof. For ii) note that $\left|\dot{X}\right| \leq 1$, for iii) recall the representation formula of the fields, and iv) holds because it holds for all iterates, they converge in C_b^1 and their second derivatives weakly-* in L^∞. □

2.3.2. Uniqueness

We prove uniqueness of the solution.

Theorem 2.11. *The obtained solution (f, E, B) of (CVM) is unique in $\mathrm{C}^1 \times \left(\mathrm{C}^2\right)^2$.*

Proof. The proof is standard. Consider the difference of two solutions and apply Lemmas 2.1 and 2.3 to show that the difference (measured in the C_b^0-norm) vanishes after a Gronwall argument. □

Moreover, it is possible to show that the solution is unique in an even larger class. Here, the constructed solution satisfies the conditions if $\mathring{E}$ and $\mathring{B}$ are compactly supported.

Theorem 2.12. *A solution (f, E, B) of (CVM) with the properties*

i) f, E, and B are of class $\mathrm{W}^{1,\infty} \cap \mathrm{H}^1$,
ii) $\operatorname{supp} f \subset [0, T] \times B_r^2$ for some $r > 0$,

is unique (here, 'solution' means that (CVM) holds pointwise almost everywhere).

Proof. Let $\left(\widetilde{f}, \widetilde{E}, \widetilde{B}\right)$ (with the above properties) solve (CVM) too and define $\overline{f} \coloneqq \widetilde{f} - f$ and so on. Then we have the system

$$\begin{aligned}
\partial_t \overline{f} + \widehat{p} \cdot \partial_x \overline{f} + \left(\widetilde{E} - \widehat{p}^\perp \widetilde{B}\right) \cdot \partial_p \overline{f} &= -\left(\overline{E} - \widehat{p}^\perp \overline{B}\right) \cdot \partial_p f, \\
\partial_t \overline{E} + \nabla_x^\perp \overline{B} &= -j_{\overline{f}}, \\
\partial_t \overline{B} + \operatorname{curl}_x \overline{E} &= 0, \\
\left(\overline{f}, \overline{E}, \overline{B}\right)(0) &= 0.
\end{aligned}$$

Note that initial values make sense because of $\mathrm{H}^1 \subset \mathrm{H}^1\left(0, T; \mathrm{L}^2\right) \hookrightarrow \mathrm{C}^0\left(0, T; \mathrm{L}^2\right)$. We have

$$\begin{aligned}
\frac{1}{2}\left\|\overline{f}(t)\right\|_{\mathrm{L}^2}^2 &= \int_0^t \int\int \overline{f} \partial_t \overline{f} \,\mathrm{d}p\mathrm{d}x\mathrm{d}\tau = \int_0^t \int\int \overline{f}\left(-\widehat{p} \cdot \partial_x \overline{f} - \left(\widetilde{E} - \widehat{p}^\perp \widetilde{B}\right) \cdot \partial_p \overline{f} - \left(\overline{E} - \widehat{p}^\perp \overline{B}\right) \cdot \partial_p f\right) \mathrm{d}p\mathrm{d}x\mathrm{d}\tau \\
&= \int_0^t \int\int \left(-\frac{1}{2} \operatorname{div}_x\left(\widehat{p}\overline{f}^2\right) - \frac{1}{2} \operatorname{div}_p\left(\left(\widetilde{E} - \widehat{p}^\perp \widetilde{B}\right)\overline{f}^2\right) - \overline{f}\left(\overline{E} - \widehat{p}^\perp \overline{B}\right) \cdot \partial_p f\right) \mathrm{d}p\mathrm{d}x\mathrm{d}\tau \\
&= -\int_0^t \int\int \overline{f}\left(\overline{E} - \widehat{p}^\perp \overline{B}\right) \cdot \partial_p f \,\mathrm{d}p\mathrm{d}x\mathrm{d}\tau \\
&\leq \|f\|_{\mathrm{W}^{1,\infty}} \int_0^t \left\|\overline{f}(\tau)\right\|_{\mathrm{L}^2}\left(\left\|\overline{E}(\tau)\right\|_{\mathrm{L}^2} + \left\|\overline{B}(\tau)\right\|_{\mathrm{L}^2}\right) \mathrm{d}\tau,
\end{aligned}$$

which implies

$$\left\|\overline{f}(t)\right\|_{\mathrm{L}^2} \leq \|f\|_{\mathrm{W}^{1,\infty}} \int_0^t \left(\left\|\overline{E}(\tau)\right\|_{\mathrm{L}^2} + \left\|\overline{B}(\tau)\right\|_{\mathrm{L}^2}\right) \mathrm{d}\tau$$

via the quadratic version of Gronwall's inequality, *cf.* [2], Theorem 5. Similarly,

$$\begin{aligned}
\frac{1}{2}\left\|\overline{B}(t)\right\|_{\mathrm{L}^2}^2 &= \int_0^t \int \overline{B} \partial_t \overline{B} \,\mathrm{d}x\mathrm{d}\tau = \int_0^t \int \overline{B}\left(-\partial_{x_1} \overline{E}_2 + \partial_{x_2} \overline{E}_1\right) \mathrm{d}x\mathrm{d}\tau = \int_0^t \int \left(\overline{E}_2 \partial_{x_1} \overline{B} - \overline{E}_1 \partial_{x_2} \overline{B}\right) \mathrm{d}x\mathrm{d}\tau \\
&= \int_0^t \int \left(-\overline{E} \cdot \partial_t \overline{E} - \overline{E} \cdot j_{\overline{f}}\right) \mathrm{d}x\mathrm{d}\tau.
\end{aligned}$$

Note that in the integration by parts no surface terms appear because of E, $B \in \mathrm{H}^1$. This computation leads to

$$\begin{aligned}
\frac{1}{2}\left(\left\|\overline{E}(t)\right\|_{\mathrm{L}^2}^2 + \left\|\overline{B}(t)\right\|_{\mathrm{L}^2}^2\right) &= \int_0^t \int -\overline{E} \cdot j_{\overline{f}} \,\mathrm{d}x\mathrm{d}\tau \leq \int_0^t \left\|\overline{E}(\tau)\right\|_{\mathrm{L}^2} \left\|j_{\overline{f}}(\tau)\right\|_{\mathrm{L}^2} \mathrm{d}\tau \\
&\leq C(r) \int_0^t \left(\left\|\overline{E}(\tau)\right\|_{\mathrm{L}^2} + \left\|\overline{B}(\tau)\right\|_{\mathrm{L}^2}\right) \left\|\overline{f}(\tau)\right\|_{\mathrm{L}^2} \mathrm{d}\tau.
\end{aligned}$$

Here, the last inequality holds because $\overline{f}$ vanishes as soon as $|p| > r$. Now again, the quadratic Gronwall lemma implies

$$\left\|\overline{E}(t)\right\|_{\mathrm{L}^2} + \left\|\overline{B}(t)\right\|_{\mathrm{L}^2} \leq C(r) \int_0^t \left\|\overline{f}(\tau)\right\|_{\mathrm{L}^2} \mathrm{d}\tau \leq C(r, T)\|f\|_{\mathrm{W}^{1,\infty}} \int_0^t \left(\left\|\overline{E}(\tau)\right\|_{\mathrm{L}^2} + \left\|\overline{B}(\tau)\right\|_{\mathrm{L}^2}\right) \mathrm{d}\tau.$$

This yields $\left(\overline{E}, \overline{B}\right) = 0$ and hence also $\overline{f} = 0$. $\square$

3. The control-to-state operator

From now on the initial data always stay fixed with $0 \le \mathring{f} \in \mathrm{C}_c^2$ and $\mathring{E}$, $\mathring{B} \in \mathrm{C}_c^3$, and $\operatorname{div} \mathring{E} = \rho_{\mathring{f}}$. As a result of the last section we may define the control-to-state operator *via*

$$\begin{aligned} S\colon \mathcal{V}_L &\to \mathrm{C}_b^2\big([0,T]\times\mathbb{R}^4\big)\times \mathrm{C}_b^2\big([0,T]\times\mathbb{R}^2;\mathbb{R}^2\big)\times \mathrm{C}_b^2\big([0,T]\times\mathbb{R}^2\big),\\ U &\mapsto (f,E,B). \end{aligned}$$

The goal is to show that S is differentiable with respect to suitable norms.

3.1. Lipschitz continuity

First we show that S is Lipschitz continuous; to be more precise, locally Lipschitz continuous. Let U, $\delta U \in \mathcal{V}_L$ and denote $(f,E,B)=S(U)$, $(\overline{f},\overline{E},\overline{B}) = S(U+\delta U)$, and $\Big(\widetilde{f},\widetilde{E},\widetilde{B}\Big) = S(U+\delta U)-S(U)$. We arrive at the system

$$\begin{aligned} \partial_t \widetilde{f} + \widehat{p}\cdot\partial_x \widetilde{f} + \big(E-\widehat{p}^\perp B\big)\cdot\partial_p\widetilde{f} &= -\Big(\widetilde{E}-\widehat{p}^\perp\widetilde{B}\Big)\cdot\partial_p\overline{f},\\ \partial_t\widetilde{E} + \nabla_x^\perp\widetilde{B} &= -j_{\widetilde{f}} - \delta U,\\ \partial_t\widetilde{B} + \operatorname{curl}_x\widetilde{E} &= 0,\\ \Big(\widetilde{f},\widetilde{E},\widetilde{B}\Big)(0) &= 0, \end{aligned}$$

which is equivalent to the system with second order Maxwell equations because of Lemmas 1.1 and 1.2.

Note that the x- and p-support of the density and the C_b^1-norm of the solution is controlled by a constant dependent on T, the initial data, L, and the $\mathcal{V}_L$-norm of the control, see Theorem 2.10. Therefore, we can perform the same estimates also on the 'bar'-solution with a constant dependent on T, the initial data, L, and $\|U\|_{\mathcal{V}_L}$ because, for instance, for $\|\delta U\|_{\mathcal{V}_L}\le 1$ we have $\|U+\delta U\|_{\mathcal{V}_L}\le \|U\|_{\mathcal{V}_L}+1$. Hence, we will only show the locally Lipschitz continuity of S.

Indeed, using again the estimates of Lemmas 2.1, 2.3, and 2.5, we see that

$$\Big\|\Big(\widetilde{f},\widetilde{E},\widetilde{B}\Big)\Big\|_{\mathrm{C}_b^1} \le C\|\delta U\|_{\mathcal{V}_L}.$$

Thus, we have proved:

Lemma 3.1. *The control-to-state map $S\colon \mathcal{V}_L \to \mathrm{C}_b^1\big([0,T]\times\mathbb{R}^4\big)\times\mathrm{C}_b^1\big([0,T]\times\mathbb{R}^2\big)^3$ is locally Lipschitz continuous.*

3.2. Solvability of a linearized system

To show even differentiability of S we will have to analyze a linearized system of the form

$$\left.\begin{aligned} \partial_t f + \widehat{p}\cdot\partial_x f + G\cdot\partial_p f &= \big(E-\widehat{p}^\perp B\big)\cdot g + a,\\ \partial_t E + \nabla_x^\perp B &= -j_f - h,\\ \partial_t B + \operatorname{curl}_x E &= 0,\\ (f,E,B)(0) &= 0 \end{aligned}\right\} \tag{LVM}$$

with already given functions $a\in \mathrm{L}^1\big(0,T;\mathrm{L}^2\big)$, $G\in\mathrm{C}_b^2$ with $\operatorname{div}_p G = 0$, $g\in\mathrm{C}_b^1$ with $g=\partial_p\widetilde{g}$ for some $\widetilde{g}\in\mathrm{C}_b^2$ and $g(t,x,p)=0$ for $|x|\ge r$ or $|p|\ge r$ for some $r>0$, and $h\in\mathcal{V}_L$. We call (f,E,B) a solution of (LVM) if f,

E, and B are of class $\mathrm{C}^0 \cap \mathrm{H}^1$, the equalities hold pointwise almost everywhere, and f vanishes for $|p| \geq R$ for some $R > 0$.

A crucial estimate is the following:

Lemma 3.2. *Let (f, E, B) be a solution of (LVM). Then*

$$\|f(t)\|_{\mathrm{L}^2} + \|E(t)\|_{\mathrm{L}^2} + \|B(t)\|_{\mathrm{L}^2} \leq C(R, \|g\|_\infty, T) \int_0^t (\|a(\tau)\|_{\mathrm{L}^2} + \|h(\tau)\|_{\mathrm{L}^2}) \mathrm{d}\tau.$$

Proof. The proof is very similar to that of Theorem 2.12 and is omitted. □

We approximate G, $\widetilde{g}$, and h with smooth functions G_k, $\widetilde{g}_k$, and h_k which are converging to G, $\widetilde{g}$, and h in C_b^2 and $\mathcal{V}_L$ respectively, and define $g_k := \partial_p \widetilde{g}_k$.

To show solvability of (LVM) for $a = 0$ we proceed similarly as before. Define $f_0 = E_{0,1} = E_{0,2} = B_0 = 0$ and solve in the k-th step

$$\begin{gathered} \partial_t f_k + \widehat{p} \cdot \partial_x f_k + G_k \cdot \partial_p f_k = \left(E_{k-1} - \widehat{p}^\perp B_{k-1}\right) \cdot g_k, \\ f_k(0) = 0 \end{gathered}$$

by defining

$$f_k(t,x,p) = \int_0^t \left(\left(E_{k-1} - \widehat{p}^\perp B_{k-1}\right) \cdot g_k\right)(X_k(0,t,x,p), P_k(0,t,x,p)) \mathrm{d}\tau$$

with the characteristics

$$\begin{aligned} \dot{X}_k &= \widehat{P}_k, & X_k(t,t,x,p) &= x, \\ \dot{P}_k &= G_k(s, X_k, P_k), & P_k(t,t,x,p) &= p, \end{aligned}$$

and then solving

$$\begin{aligned} \partial_t^2 E_k - \Delta_x E_k &= -\partial_t j_{f_k} - \partial_t h_k - \partial_x \rho_{f_k} + \partial_x \int_0^t \operatorname{div}_x h_k \, \mathrm{d}\tau, \\ \partial_t^2 B_k - \Delta_x B_k &= \operatorname{curl}_x j_{f_k} + \operatorname{curl}_x h_k, \\ (E_k, B_k)(0) &= 0, \\ \partial_t E_k(0) &= -U_k(0), \\ \partial_t B_k(0) &= 0. \end{aligned}$$

All iterates are again of class C^∞. Furthermore, the characteristics are independent of the solution sequence (f_k, E_k, B_k). Thus, we instantly have $\left|\dot{P}_k\right| \leq C$, so $|P_k - p| \leq CT$. Having a look at the formula for f_k we conclude that f_k vanishes as soon as

$$|p| \geq 2r + CT =: Q \tag{3.1}$$

since then the integrand vanishes as a result of

$$|P_k(s,t,x,p)| \geq |p| - |P_k - p| \geq 2r + CT - CT = 2r.$$

The same can be done for the x-coordinate starting with $\left|\dot{X}_k\right| \le 1$; hence $f_k(t,x,p) = 0$ for $|x| \ge 2r + T$. The assertions of Section 2.1 are directly applicable. We do not have to insert some α because of the already known bound on the p-support of f_k. Therefore, (LC) holds for the iterated system, and we can thus switch between first order and second order Maxwell equations; note that $\big(E_{k-1} - \widehat{p}^\perp B_{k-1}\big) \cdot g_k = \mathrm{div}_p\big(\big(E_{k-1} - \widehat{p}^\perp B_{k-1}\big)\widetilde{g}_k\big)$.

We proceed like in Section 2.3: The iterates are bounded in C^1_b and are Cauchy with respect to the C^0_b-norm. However, after that there appears a difference: Unfortunately, we can not show the Cauchy property with respect to the C^1_b-norm. For this we would first have to bound second derivatives of f_k which would require control of second derivatives of g_k. This, on the other hand, would require a smoother g. But for the later application we will not have more regularity of g than C^1_b.

Thus, we have to proceed differently: Since f_k, E_k, and B_k are bounded in the C^1_b-norm, their first derivatives converge, after extracting a suitable subsequence, to the respective derivatives of f, E, and B in L^∞ in the weak-*-sense. Because of

$$\begin{aligned}
&\left|\int_0^T \int\int (G_k \cdot \partial_p f_k \varphi - G \cdot \partial_p f \varphi) \mathrm{d}p\mathrm{d}x\mathrm{d}\tau\right| \\
&\le \int_0^T \int\int |G_k - G||\partial_p f_k||\varphi| \mathrm{d}p\mathrm{d}x\mathrm{d}\tau + \left|\int_0^T \int\int G(\partial_p f_k - \partial_p f)\varphi \,\mathrm{d}p\mathrm{d}x\mathrm{d}\tau\right| \\
&\le C\|G_k - G\|_\infty \|\varphi\|_{\mathrm{L}^1} + \left|\int_0^T \int\int G(\partial_p f_k - \partial_p f)\varphi \,\mathrm{d}p\mathrm{d}x\mathrm{d}\tau\right| \to 0
\end{aligned}$$

for $k \to \infty$ for any test function φ, (f,E,B) satisfies (LVM) pointwise almost everywhere; the other terms are obviously easier to handle. Altogether we have found a solution of (LVM) of class $\mathrm{C}^0 \cap \mathrm{W}^{1,\infty}$. Furthermore, it is also of class H^1 because all sequence elements have compact support with respect to x, p or x respectively uniformly in t and k; for the fields recall the representation formula.

For uniqueness, let (f_1, E_1, B_1) be a solution of (LVM) too and define $f_2 := f - f_1$ and so on which yields

$$\begin{aligned}
\partial_t f_2 + \widehat{p} \cdot \partial_x f_2 + G \cdot \partial_p f_2 &= \big(E_2 - \widehat{p}^\perp B_2\big) \cdot g, \\
\partial_t E_2 + \nabla_x^\perp B_2 &= -j_{f_2}, \\
\partial_t B_2 + \mathrm{curl}_x\, E_2 &= 0, \\
(f_2, E_2, B_2)(0) &= 0.
\end{aligned}$$

Applying Lemma 3.2 this instantly implies that f_2, E_2, and B_2 vanish.

3.3. Differentiability

We want to study the differentiability of $S\colon \mathcal{V}_L \to \mathrm{C}^0\big(0,T;\mathrm{L}^2\big(\mathbb{R}^4\big)\big) \times \mathrm{C}^0\big(0,T;\mathrm{L}^2\big(\mathbb{R}^2\big)\big)^3$. Let $U \in \mathcal{V}_L$ and let $\delta U \in \mathcal{V}_L$ be some perturbation. In the following denote $(f,E,B) = S(U)$ and $\big(\overline{f},\overline{E},\overline{B}\big) = S(U + \delta U)$. The candidate for the linearization is $S'(U)\delta U = (\delta f, \delta E, \delta B)$ where the right hand side satisfies

$$\begin{aligned}
\partial_t \delta f + \widehat{p} \cdot \partial_x \delta f + \big(E - \widehat{p}^\perp B\big) \cdot \partial_p \delta f &= -\big(\delta E - \widehat{p}^\perp \delta B\big) \cdot \partial_p f, \\
\partial_t \delta E + \nabla_x^\perp \delta B &= -j_{\delta f} - \delta U, \\
\partial_t \delta B + \mathrm{curl}_x\, \delta E &= 0, \\
(\delta f, \delta E, \delta B)(0) &= 0.
\end{aligned}$$

Indeed, this system can be solved because of $G \coloneqq E - \widehat{p}^{\perp} B \in \mathrm{C}_b^2$ (note that $\operatorname{div}_p G = 0$), $g \coloneqq -\partial_p f \in \mathrm{C}_b^1$, and $h \coloneqq \delta U \in \mathcal{V}_L$. First we note that $S'(U)$ is linear and that by Lemma 3.2

$$\|(\delta f, \delta E, \delta B)\|_{\mathrm{C}^0(0,T;\mathrm{L}^2)} \le C \int_0^T \|\delta U(t)\|_{\mathrm{L}^2} \mathrm{d}t \le C \|\delta U\|_{\mathcal{V}_L} \tag{3.2}$$

which says that $S'(U)$ is bounded. The last inequality holds because of $\operatorname{supp} \delta U(t) \subset B_L$.

The next step is to show that $S(U + \delta U) - S(U) - S'(U)\delta U$ is 'small'. Defining $\widetilde{f} \coloneqq \overline{f} - f - \delta f$ and so on and subtracting the respective equations yield

$$\begin{aligned}
\partial_t \widetilde{f} + \widehat{p} \cdot \partial_x \widetilde{f} + \big(E - \widehat{p}^{\perp} B\big) \cdot \partial_p \widetilde{f} &= -\Big(\widetilde{E} - \widehat{p}^{\perp} \widetilde{B}\Big) \cdot \partial_p f - \big(\overline{E} - E - \widehat{p}^{\perp}\big(\overline{B} - B\big)\big) \cdot \partial_p \big(\overline{f} - f\big), \\
\partial_t \widetilde{E} + \nabla_x^{\perp} \widetilde{B} &= -j_{\widetilde{f}}, \\
\partial_t \widetilde{B} + \operatorname{curl}_x \widetilde{E} &= 0, \\
\Big(\widetilde{f}, \widetilde{E}, \widetilde{B}\Big)(0) &= 0.
\end{aligned}$$

Applying Lemma 3.2 we conclude

$$\left\| \Big(\widetilde{f}, \widetilde{E}, \widetilde{B}\Big) \right\|_{\mathrm{C}^0(0,T;\mathrm{L}^2)} \le C \int_0^T \|a(t)\|_{\mathrm{L}^2} \mathrm{d}t$$

where

$$a \coloneqq -\big(\overline{E} - E - \widehat{p}^{\perp}\big(\overline{B} - B\big)\big) \cdot \partial_p \big(\overline{f} - f\big).$$

Here we have to exploit the Lipschitz property of S. Lemma 3.1 yields

$$\|a(t)\|_{\mathrm{L}^2} \le C\big(\big\|\overline{E} - E\big\|_\infty + \big\|\overline{B} - B\big\|_\infty\big)\big\|\overline{f} - f\big\|_{\mathrm{C}_b^1} \le C \|\delta U\|_{\mathcal{V}_L}^2.$$

Note that for the first inequality the fact was used that $\overline{f}$ and f have compact support in x and p uniformly in t and independent of $\|\delta U\|_{\mathcal{V}_L}$ for, for instance, $\|\delta U\|_{\mathcal{V}_L} \le 1$ (recall Theorem 2.10 and the reasoning in Sect. 3.1).

Finally we arrive at

$$\left\| \Big(\widetilde{f}, \widetilde{E}, \widetilde{B}\Big) \right\|_{\mathrm{C}^0(0,T;\mathrm{L}^2)} \le C \|\delta U\|_{\mathcal{V}_L}^2 \tag{3.3}$$

which proves part of i) of the following theorem:

Theorem 3.3. *The following maps are continuously Fréchet-differentiable with locally Lipschitz derivative:*

i) $S \colon \mathcal{V}_L \to W \coloneqq \mathrm{C}^0\big(0,T;\mathrm{L}^2\big(\mathbb{R}^4\big)\big) \times \mathrm{C}^0\big(0,T;\mathrm{L}^2\big(\mathbb{R}^2\big)\big)^3$,
ii) $\Phi \coloneqq \rho \circ S_1 \colon \mathcal{V}_L \to \mathrm{C}^0\big(0,T;\mathrm{L}^2\big(\mathbb{R}^2\big)\big)$, $U \mapsto \rho_f$,
iii) $\overline{\Phi} \coloneqq \rho \circ S_1 \colon \mathcal{V}_L \to \mathrm{C}^0\big(0,T;\mathrm{L}^1\big(\mathbb{R}^2\big)\big)$, $U \mapsto \rho_f$.

Proof. For part ii) define

$$\Phi'(U)\delta U \coloneqq \rho_{\delta f}. \tag{3.4}$$

Now it is crucial to bound the p-support of $\overline{f}$, f, and δf by a constant $C > 0$ only depending on T, the initial data, L, and $\|U\|_{\mathcal{V}_L}$. We first consider δf. The control of the p-support in (3.1) holds for all iterates and hence for δf. The constant there only depends on T, $\|G\|_\infty = \left\|E - \widehat{p}^\perp B\right\|_\infty$, the p-support of $\partial_p f$, and L. Because of Theorem 2.10 the absolute values of the fields E and B and the p-support of f are controlled by some constant only depending on T, the initial data, L, and $\|U\|_{\mathcal{V}_L}$. Hence, we have together with (3.2)

$$\|\rho_{\delta f}(t)\|_{\mathrm{L}^2} = \left(\int \left|\int \delta f \,\mathrm{d}p\right|^2 \mathrm{d}x\right)^{\frac{1}{2}} \leq C\left(\int\int |\delta f|^2 \mathrm{d}p\mathrm{d}x\right)^{\frac{1}{2}} \leq C\|\delta U\|_{\mathcal{V}_L}$$

which implies that $\Phi'(U)$ is bounded. Furthermore, the p-supports of $\overline{f}$ and f only depend on T, the initial data, L, and $\|U\|_{\mathcal{V}_L}$ (for again $\|\delta U\|_{\mathcal{V}_L} \leq 1$ for example). Hence, the same assertion holds for $\widetilde{f} = \overline{f} - f - \delta f$ and therefore with (3.3)

$$\left\|\rho_{\widetilde{f}}(t)\right\|_{\mathrm{L}^2} = \left(\int \left|\int \widetilde{f} \,\mathrm{d}p\right|^2 \mathrm{d}x\right)^{\frac{1}{2}} \leq C\left(\int\int \left|\widetilde{f}\right|^2 \mathrm{d}p\mathrm{d}x\right)^{\frac{1}{2}} \leq C\|\delta U\|_{\mathcal{V}_L}^2.$$

Together with the equality

$$\Phi(U + \delta U) - \Phi(U) - \Phi'(U)\delta U = \rho_{\overline{f}} - \rho_f - \rho_{\delta f} = \rho_{\widetilde{f}}$$

this instantly yields that $\Phi'(U)$ is indeed the Fréchet-derivative of Φ in U. Part iii) is an instant consequence of ii) and the support assertions discussed above. The derivative of $\overline{\Phi}$ is given by (3.4) as before.

To show continuity of S', let $\delta V \in \mathcal{V}_L$ with $\|\delta V\|_{\mathcal{V}_L} \leq 1$. We have to investigate

$$\left(\check{f}, \check{E}, \check{B}\right) := \left(f^1, E^1, B^1\right) - \left(f^0, E^0, B^0\right) := S'(U + \delta U)\delta V - S'(U)\delta V.$$

Applying the previously given formula for S' we arrive at

$$\begin{aligned}
\partial_t \check{f} + \widehat{p} \cdot \partial_x \check{f} + \left(\overline{E} - \widehat{p}^\perp \overline{B}\right) \cdot \partial_p \check{f} &= -\left(\check{E} - \widehat{p}^\perp \check{B}\right) \cdot \partial_p \overline{f} - \left(E^0 - \widehat{p}^\perp B^0\right) \cdot \partial_p\left(\overline{f} - f\right)\\
&\quad - \left(\overline{E} - E - \widehat{p}^\perp\left(\overline{B} - B\right)\right) \cdot \partial_p f^0,\\
\partial_t \check{E} + \nabla_x^\perp \check{B} &= -j_{\check{f}},\\
\partial_t \check{B} + \operatorname{curl}_x \check{E} &= 0,\\
\left(\check{f}, \check{E}, \check{B}\right)(0) &= 0.
\end{aligned}$$

We know that the p-support of f^0 and the absolute values of E^0 and B^0 are controlled by a constant only depending on T, the initial data, L, $\|U\|_{\mathcal{V}_L}$, and $\|\delta V\|_{\mathcal{V}_L}$ (the latter can be neglected, of course). The dependence on some terms in f, E, and B can be eliminated like in the beginning of this proof. Hence, proceeding as before and using Lemma 3.2 and the locally Lipschitz continuity of S, we conclude

$$\left\|\left(\check{f}, \check{E}, \check{B}\right)\right\|_W \leq C\|\delta U\|_{\mathcal{V}_L}$$

where C only depends on T, the initial data, L, and $\|U\|_{\mathcal{V}_L}$. This leads to

$$\|S'(U + \delta U) - S'(U)\|_{\mathcal{L}(\mathcal{V}_L, W)} \leq C\|\delta U\|_{\mathcal{V}_L}$$

which says that S' is even locally Lipschitz continuous.

Using the assertions for the p-support of f^0 and f^1 (controlled by a constant only depending on T, the initial data, L, and $\|U\|_{\mathcal{V}_L}$ if $\|\delta U\|_{\mathcal{V}_L} \leq 1$) we conclude

$$\left\|\rho_{\check{f}}\right\|_{\mathrm{C}^0(0,T;\mathrm{L}^2)}, \left\|\rho_{\check{f}}\right\|_{\mathrm{C}^0(0,T;\mathrm{L}^1)} \leq C\|\check{f}\|_{\mathrm{C}^0(0,T;\mathrm{L}^2)} \leq C\|\delta U\|_{\mathcal{V}_L}$$

as before. This implies that Φ' and $\overline{\Phi}'$ are locally Lipschitz continuous. □

4. Optimal control problem

Now we consider some optimal control problems. We want to minimize some objective function that depends on the external control U and the state (f, E, B). The control and the state are coupled *via* (CVM) so that (CVM) appears as a constraint.

We first give thought to a problem with general controls and a general objective function. Then we proceed with optimizing problems where the objective function is explicitly given and where the control set is restricted to such controls that are realizable in applications concerning the control of a plasma.

4.1. General problem

4.1.1. Control space

Until now we have worked with the control space

$$\mathcal{V}_L = \left\{U \in \mathrm{W}^{2,1}\left(0,T;\mathrm{C}_b^4\left(\mathbb{R}^2;\mathbb{R}^2\right)\right) \mid U(t,x) = 0 \text{ for } |x| \geq L\right\}.$$

To apply standard optimization techniques it is necessary that the control space is reflexive. Hence, we choose

$$\mathcal{U}_L := \left\{U \in \mathrm{H}^2\left(0,T;\mathrm{W}^{5,\gamma}\left(\mathbb{R}^2;\mathbb{R}^2\right)\right) \mid U(t,x) = 0 \text{ for } |x| \geq L\right\},$$

where $\gamma > 2$ is fixed, equipped with the $\mathrm{H}^2\left(0,T;\mathrm{W}^{5,\gamma}\right)$-norm. By Sobolev's embedding theorems, $\mathcal{U}_L$ is continuously embedded in $\mathcal{V}_L$.

In accordance with Theorems 2.10 and 3.3, we have already proved that there is a continuously differentiable control-to-state operator

$$\begin{aligned} S\colon \mathcal{V}_L &\to \left(\mathrm{C}_b^2\left([0,T]\times\mathbb{R}^4\right)\times\mathrm{C}_b^2\left([0,T]\times\mathbb{R}^2;\mathbb{R}^2\right)\times\mathrm{C}_b^2\left([0,T]\times\mathbb{R}^2\right), \|\cdot\|_{\mathrm{C}^0(0,T;\mathrm{L}^2)}\right),\\ U &\mapsto (f,E,B), \end{aligned}$$

such that (CVM) holds for (f, E, B) and control U. Furthermore, the map $U \mapsto \rho_f$ is continuously differentiable with respect to the $\mathrm{C}^0\left(0,T;\mathrm{L}^2\right)$- and $\mathrm{C}^0\left(0,T;\mathrm{L}^1\right)$-norm in the image space. Moreover, the C_b^2-norm and the x- and p-support of (f, E, B) are controlled by a constant only depending on T, L, the initial data, and $\|U\|_{\mathcal{V}_L}$.

By $\mathcal{U}_L \hookrightarrow \mathcal{V}_L$, these assertions also hold with $\mathcal{U}_L$ instead of $\mathcal{V}_L$.

We are aware that the regularity assumed for $U \in \mathcal{U}_L$ is quite high. However, to derive the assertions mentioned above, $U \in \mathcal{V}_L$ was really necessary.

4.1.2. Existence of minimizers

We consider the general problem

$$\left.\begin{aligned} \min_{(f,E,B)\in\left(\mathrm{C}^2\cap\mathrm{H}^1\right)^3, U\in\mathcal{U}_L} \quad &\phi(f,E,B,U)\\ \text{s.t. } (f,E,B) &= S(U). \end{aligned}\right\} \tag{GP}$$

We have to specify some assumptions on ϕ:

Condition 4.1. i) $\phi\colon \left(\mathrm{C}^2\cap \mathrm{H}^1\right)^3 \times \mathcal{U}_L \to \mathbb{R}\cup\{\infty\}$ and $\phi \not\equiv \infty$,

ii) ϕ is coercive in $U\in\mathcal{U}_L$, *i.e.*, in general: Let X, Y be normed spaces; $\psi\colon X\times Y\to\mathbb{R}$ is said to be coercive in $y\in Y$ iff for all sequences $(y_k)\subset Y$ with $\|y_k\|_Y\to\infty$, $k\to\infty$, then also $\psi(x_k,y_k)\to\infty$, $k\to\infty$, for any sequence $(x_k)\subset X$,

iii) ϕ is weakly lower semicontinuous, *i.e.*: if $(f_k,E_k,B_k)\rightharpoonup(f,E,B)$ in H^1 and $U_k\rightharpoonup U$ in $\mathcal{U}_L$, then $\phi(f,E,B,U)\le\liminf_{k\to\infty}\phi(f_k,E_k,B_k,U_k)$.

These assumptions allow us to prove existence of a (not necessarily unique) minimizer. We will first prove a lemma that will be useful later:

Lemma 4.2. *Let $(U_k)\subset\mathcal{V}_L$ be bounded and $(f_k,E_k,B_k)=S(U_k)$. Then, after extracting a suitable subsequence, it holds that:*

i) The sequences (f_k), (E_k), and (B_k) converge weakly in H^1, weakly- in $\mathrm{W}^{1,\infty}$, and strongly in L^2 to some f, E, and B.*

ii) There is $r>0$ such that f, E, B, and, for all $k\in\mathbb{N}$, f_k, E_k, and B_k vanish if $|x|\ge r$ or $|p|\ge r$.

iii) If additionally $U_k\to U$ in the sense of distributions for some $U\in\mathcal{V}_L$ for $k\to\infty$, then $(f,E,B)=S(U)$ and f, E, and B are of class C^2_b.

Proof. By Theorem 2.10, on the one hand, (f_k,E_k,B_k) is bounded in the C^1_b-norm. On the other hand, f_k vanishes as soon as $|p|$ is large enough uniformly in k. Moreover, f_k, E_k, and B_k vanish as soon as $|x|$ is large enough. Hence, (f_k,E_k,B_k) is also bounded in H^1 and in $\mathrm{H}^1\left(0,T;\mathrm{L}^2\right)$. Together with the boundedness in C^1_b, (f_k,E_k,B_k) converge, after extracting a suitable subsequence, to some (f,E,B), namely weakly in H^1, and weakly-* in $\mathrm{W}^{1,\infty}$. This proves ii) and part of i).

For the remaining part of i) (strong convergence in L^2) we have to exploit some compactness. This compactness is guaranteed by the theorem of Rellich-Kondrachov. By the reasoning above, (f_k,E_k,B_k) are bounded in H^1 and in fact, only a bounded subset of the x- and p-space matters. Hence, (a subsequence of) (f_k,E_k,B_k) converges strongly in L^2 to the limit (f,E,B).

For iii), we have to pass to the limit in (CVM). First, the initial conditions are preserved in the limit since $\mathrm{H}^1\hookrightarrow\mathrm{H}^1\left(0,T;\mathrm{L}^2\right)\hookrightarrow\mathrm{C}^0\left(0,T;\mathrm{L}^2\right)$. Furthermore, the Vlasov and Maxwell equations hold pointwise almost everywhere for the limit functions: The only difficult part is the nonlinear term in the Vlasov equation. To handle this, we have to make use of the strong convergence in L^2 obtained above. We find for each $\varphi\in\mathrm{C}^\infty_c\left(]0,T[\times\mathbb{R}^4\right)$ that

$$\begin{aligned}&\left|\int_0^T\int\int\left(\left(E_k-\widehat{p}^\perp B_k\right)\cdot\partial_p f_k-\left(E-\widehat{p}^\perp B\right)\cdot\partial_p f\right)\varphi\,\mathrm{d}p\mathrm{d}x\mathrm{d}t\right|\\&\le\left|\int_0^T\int\int\left(E-\widehat{p}^\perp B\right)\cdot\left(\partial_p f_k-\partial_p f\right)\varphi\,\mathrm{d}p\mathrm{d}x\mathrm{d}t\right|+\|\partial_p f_k\|_\infty\int_0^T\int\int\left(|E_k-E|+|B_k-B|\right)|\varphi|\mathrm{d}p\mathrm{d}x\mathrm{d}t.\end{aligned}$$

Both terms converge to 0 for $k\to\infty$ since $f_k\rightharpoonup f$ in H^1, $E_k\to E$, $B_k\to B$ in L^2, and f_k is bounded in C^1_b. Therefore, altogether, (CVM) holds pointwise almost everywhere. Now we can apply Theorem 2.12 to conclude (f,E,B) equals $S(U)$ and is hence of class C^2_b. □

Theorem 4.3. *Let ϕ satisfy Condition 4.1. Then there is a minimizer of* (GP).

Proof. We consider a minimizing sequence (f_k,E_k,B_k,U_k) with $(f_k,E_k,B_k)=S(U_k)$ and

$$\lim_{k\to\infty}\phi(f_k,E_k,B_k,U_k)=m:=\inf_{U\in\mathcal{U}_L,(f,E,B)=S(U)}\phi(f,E,B,U)\in\mathbb{R}\cup\{-\infty\}.$$

By coercivity in U, cf. Condition 4.1 ii), (U_k) is bounded in $\mathcal{U}_L$ and therefore in $\mathcal{V}_L$. Hence, we may extract a weakly convergent subsequence (also denoted by U_k) since $\mathrm{H}^2\big(0,T;\mathrm{W}^{5,\gamma}\big)$ is reflexive. The weak limit U is the candidate for being an optimal control. Of course, by weak convergence, U vanishes for $|x| \geq L$; hence $U \in \mathcal{U}_L$. Because of $\mathcal{U}_L \hookrightarrow \mathrm{L}^1$ we also get $U_k \rightharpoonup U$ in L^1 and hence $U_k \to U$ in the sense of distributions. Lemma 4.2 yields $(f_k, E_k, B_k) \rightharpoonup (f, E, B)$ in H^1 (after extracting a suitable subsequence) and $(f, E, B) = S(U)$. Together with the weak lower semicontinuity of ϕ, see Condition 4.1 iii), we instantly get $\phi(f, E, B, U) = m$ which proves optimality. □

In order to be able of examining some problem that is somehow application-oriented, we first have to think about possible problems concerning the conditions on the objective function ϕ. Especially the coercivity in U will make some trouble since the $\mathcal{U}_L$-norm is pretty strong. One can try to guarantee these conditions in various ways, for example if $\phi(f, E, B, U) = \psi(f, E, B) + \|U\|_{\mathcal{U}_L}^2$; the objective function contains some cost term of the control in the full $\mathcal{U}_L$-norm. But typically in applications, such a strong cost term makes no sense. Furthermore, first order optimality conditions would contain a differential equation of very high order, which is hard to solve.

On the other hand, we can not simply use a less regular control space. Firstly, we need $\mathcal{U}_L \hookrightarrow \mathcal{V}_L$ to ensure that the control-to-state operator is differentiable; this will be useful later. Secondly, $\mathcal{U}_L$ needs to be reflexive to extract (in some sense) converging subsequences from a minimizing sequence. Here we should remark that we also could demand $\mathrm{W}^{2,p}$-regularity in time for $p > 1$ instead of H^2-regularity which would allow more controls if $1 < p < 2$. However, working in a H^2-setting (at least in time) is more convenient.

4.2. An optimization problem with realizable external currents

4.2.1. Motivation

As the previous considerations suggest, it would be nice if we somehow eliminated the variability of the control with respect to the space coordinate. This can be achieved by only considering controls of the form

$$U(t, x) = \sum_{j=1}^{N} u_j(t) z_j(x)$$

where the functions $0 \not\equiv z_j \in \mathrm{C}_b^4\big(\mathbb{R}^2; \mathbb{R}^2\big)$ with z_j vanishing for $|x| \geq r_j > 0$ are fixed, and we only vary the functions $u_j \in \mathrm{H}^2([0, T])$.

From a physical point of view, this model describes an ensemble of N coils with ‘size’ r_j, that stay fixed in time. Obviously, U is an element of $\mathcal{V}_L$ if we set $L = \max\{r_j \mid j = 1, \dots N\}$. Each coil generates a current z_j at full capacity that is tangential to the plane and that extends infinitely in the third space dimension. We control the system by turning these coils on whereby the capacity u_j is suitably adjusted as a function of time. Hence, we will have to consider an additional constraint $|u_j| \leq 1$. Physically, the consideration only of controls of the above form is no substantial restriction at all because only such control fields are realizable in applications.

A similar approach was done by P. Knopf and the author [11].

4.2.2. Formulation

The problem to be considered is the following:

$$\left.\begin{aligned} &\min_{\substack{(f,E,B)\in\left(\mathrm{C}^2\cap\mathrm{H}^1\right)^3,\\ u\in\mathrm{H}^2([0,T])^N}} \frac{1}{2}\|\rho_f - \rho_d\|_{\mathrm{L}^2([0,T]\times\mathbb{R}^2)}^2 + \frac{\beta}{2}\sum_{j=1}^{N} c_j\left(\|u_j\|_{\mathrm{L}^2([0,T])}^2 + \beta_1\|\partial_t u_j\|_{\mathrm{L}^2([0,T])}^2 + \beta_2\big\|\partial_t^2 u_j\big\|_{\mathrm{L}^2([0,T])}^2\right)\\ &\qquad\text{s.t. } (f, E, B) = S\left(\sum_{j=1}^{N} u_j z_j\right), |u_j| \leq 1 \end{aligned}\right\} \tag{P}$$

where $c_j := \|z_j\|^2_{\mathrm{L}^2(\mathbb{R}^2;\mathbb{R}^2)}$. We give some comments on the objective function:

- The charge density shall be as close as possible to some given desired density $\rho_d \in \mathrm{L}^2\big([0,T]\times\mathbb{R}^2\big)$. One could consider the L^2-norm of some $f - f_d$ instead but the space coordinates of the particles are of actual interest rather than their momenta.
- Furthermore, the cost term containing the control shall be as small as possible. We have to use the full H^2-norm (an equivalent norm, to be more precise) of the u_j in the regularization term so that the objective function is coercive in $u \in \mathrm{H}^2([0,T])^N$. However, the L^2-norms of the u_j itself are more interesting than the ones of their derivatives. Hence, it is suitable to choose $0 < \beta_1, \beta_2 \ll 1$.
- The parameter $\beta > 0$ indicates which of the two aims mentioned above shall rather be achieved.

4.2.3. Existence of minimizers

Section 4.1.2 is useful for showing existence of minimizers of (P).

Theorem 4.4. *There is a minimizer of* (P).

Proof. The objective function, abbreviated by $\phi = \phi(f,E,B,u) = \phi_1(f) + \phi_2(u)$ (let ϕ_1 be the term with $\rho_f - \rho_d$ and ϕ_2 the remaining sum), is coercive in $u \in \mathrm{H}^2([0,T])^N$ because of

$$\phi(f,E,B,u) \geq \frac{\beta}{2}\min\{1,\beta_1,\beta_2\}\min\{c_j \mid j = 1,\dots,N\}\|u\|^2_{(\mathrm{H}^2)^N},$$

where $\|u\|^2_{(\mathrm{H}^2)^N} = \sum_{j=1}^N \|u_j\|^2_{\mathrm{H}^2([0,T])}$. Hence, considering a minimizing sequence $\big(f_k,E_k,B_k,u^k\big)$ (we use upper indices for u^k to avoid confusion with the components) with $(f_k,E_k,B_k) = S\Big(\sum_{j=1}^N u_j^k z_j\Big)$ and $\big|u_j^k\big| \leq 1$, we conclude that $\big(u^k\big)$ is bounded in $\big(\mathrm{H}^2\big)^N$; hence $u^k \rightharpoonup u$ in $\big(\mathrm{H}^2\big)^N$ for some $u \in \big(\mathrm{H}^2\big)^N$ for $k \to \infty$, possibly after extracting a suitable subsequence. The constraint $|u_j| \leq 1$ is obviously preserved by weak convergence. Furthermore, the sequence $(U_k) := \Big(\sum_{j=1}^N u_j^k z_j\Big)$ is bounded in $\mathcal{V}_L$ because of $\mathrm{H}^2([0,T]) \hookrightarrow \mathrm{W}^{2,1}([0,T])$.

Clearly, $U_k \to U := \sum_{j=1}^N u_j z_j$ in the sense of distributions by $u_j^k \rightharpoonup u_j$ in H^2. Therefore, Lemma 4.2 is applicable and delivers some f, E, and B so that (CVM) is preserved in the limit. The remaining part is to show that U is indeed an optimal control. Firstly, $u^k \rightharpoonup u$ in $\big(\mathrm{H}^2\big)^N$ instantly implies $\phi_2(u) \leq \liminf_{k\to\infty}\phi_2\big(u^k\big)$. Secondly, by Lemma 4.2, all f_k and f have compact support with respect to p uniformly in k, and $f_k \to f$ in L^2. These properties yield $\rho_{f_k} \to \rho_f$ in L^2 by Hölder's inequality and therefore $\phi_1(f) = \lim_{k\to\infty}\phi_1(f_k)$. This finally proves the desired optimality. □

4.2.4. Differentiability of the objective function

Next we study the differentiability of the objective function.

Theorem 4.5. *i) The solution map*

$$\Xi\colon \big(\mathrm{H}^2([0,T])\big)^N \to \mathrm{C}^0\big(0,T;\mathrm{L}^2\big(\mathbb{R}^4\big)\big) \times \mathrm{C}^0\big(0,T;\mathrm{L}^2\big(\mathbb{R}^2\big)\big)^3,$$
$$u \mapsto (f,E,B) = S\left(\sum_{j=1}^N u_j z_j\right)$$

is continuously Fréchet-differentiable and $\Xi'(u)\delta u = (\delta f, \delta E, \delta B)$ *satisfies*

$$\begin{aligned}\partial_t \delta f + \widehat{p} \cdot \partial_x \delta f + \big(E - \widehat{p}^{\perp} B\big) \cdot \partial_p \delta f &= -\big(\delta E - \widehat{p}^{\perp} \delta B\big) \cdot \partial_p f, \\ \partial_t \delta E + \nabla_x^{\perp} \delta B &= -j_{\delta f} - \delta U, \\ \partial_t \delta B + \operatorname{curl}_x \delta E &= 0, \\ (\delta f, \delta E, \delta B)(0) &= 0\end{aligned}$$

where $\delta U = \sum_{j=1}^N \delta u_j z_j$.

ii) The maps

$$\begin{aligned}\Psi \colon \big(\mathrm{H}^2([0,T])\big)^N &\to \mathrm{C}^0\big(0,T;\mathrm{L}^2(\mathbb{R}^2)\big), \\ u &\mapsto \rho_f\end{aligned}$$

and

$$\begin{aligned}\overline{\Psi} \colon \big(\mathrm{H}^2([0,T])\big)^N &\to \mathrm{C}^0\big(0,T;\mathrm{L}^1(\mathbb{R}^2)\big), \\ u &\mapsto \rho_f\end{aligned}$$

are continuously Fréchet-differentiable and $\Psi'(u)\delta u = \rho_{\delta f}$ *with* δf *from above.*

iii) The objective function

$$\begin{aligned}&\overline{\phi} \colon \big(\mathrm{H}^2([0,T])\big)^N \to \mathbb{R}, \\ &u \mapsto \frac{1}{2}\|\rho_f - \rho_d\|_{\mathrm{L}^2}^2 + \frac{\beta}{2}\sum_{j=1}^N c_j\Big(\|u_j\|_{\mathrm{L}^2}^2 + \beta_1\|\partial_t u_j\|_{\mathrm{L}^2}^2 + \beta_2\big\|\partial_t^2 u_j\big\|_{\mathrm{L}^2}^2\Big)\end{aligned}$$

is continuously Fréchet-differentiable and

$$\overline{\phi}'(u)\delta u = \langle \rho_f - \rho_d, \rho_{\delta f}\rangle_{\mathrm{L}^2} + \beta \sum_{j=1}^N c_j\big(\langle u_j, \delta u_j\rangle_{\mathrm{L}^2} + \beta_1\langle \partial_t u_j, \partial_t \delta u_j\rangle_{\mathrm{L}^2} + \beta_2\big\langle \partial_t^2 u_j, \partial_t^2 \delta u_j\big\rangle_{\mathrm{L}^2}\big)$$

with δf *from above.*

Proof. Clearly, $u \mapsto \sum_{j=1}^N u_j z_j$ is differentiable by linearity and boundedness. Hence, all assertions follow immediately by Theorem 3.3 and the chain rule. □

4.2.5. Optimality conditions

Now we want to deduce first order optimality conditions for a (local) minimizer of (P). First we write (P) in the equivalent form

$$\left.\begin{aligned}&\min_{u\in \mathrm{H}^2([0,T])^N} \frac{1}{2}\|\Psi(u) - \rho_d\|_{\mathrm{L}^2([0,T]\times\mathbb{R}^2)}^2 + \frac{\beta}{2}\sum_{j=1}^N c_j\Big(\|u_j\|_{\mathrm{L}^2([0,T])}^2 + \beta_1\|\partial_t u_j\|_{\mathrm{L}^2([0,T])}^2 + \beta_2\big\|\partial_t^2 u_j\big\|_{\mathrm{L}^2([0,T])}^2\Big) \\ &\qquad \text{s.t. } -u_j + 1 \geq 0, u_j + 1 \geq 0.\end{aligned}\right\} \quad \text{(P')}$$

Here, the objective function $\overline{\phi} = \overline{\phi}(u) = \phi(\Xi(u), u)$ is a function of only the control.

The constraints will lead to corresponding Lagrange multipliers. In general, to prove their existence, some condition on the constraints is necessary. On this account we verify the constraint qualification of Zowe and Kurcyusz, see [16], which is based on a fundamental work of Robinson [15]. We rewrite the constraints: $g(u) \in K$, where $g(u) = (-u+1, u+1) \in K$, K denoting the cone of component-wise positive functions in $\mathrm{C}^0([0,T])^{2N}$. The constraint qualification we have to verify is

$$g'(u)\big(\mathrm{H}^2([0,T])\big)^N - \{k - \lambda g(u) \mid k \in K, \lambda \geq 0\} = \mathrm{C}^0([0,T])^{2N}.$$

In other words, for given $(w^+, w^-) \in \big(\mathrm{C}^0([0,T])\big)^{2N}$ we have to find $\delta u \in \mathrm{H}^2([0,T])^N$, $\lambda \in \mathbb{R}_{\geq 0}$, and $k = (\theta^+, \theta^-) \in \big(\mathrm{C}^0([0,T])\big)^{2N}$ with θ_j^+, $\theta_j^- \geq 0$, satisfying

$$(-\delta u, \delta u) - \big(\theta^+, \theta^-\big) + \lambda(-u+1, u+1) = \big(w^+, w^-\big). \tag{4.1}$$

We abbreviate

$$\vartheta^+ \coloneqq \max_{i=1,\ldots,N} \left\|w_i^+\right\|_\infty, \vartheta^- \coloneqq \max_{i=1,\ldots,N} \left\|w_i^-\right\|_\infty.$$

Now let

$$\lambda \coloneqq \frac{\vartheta^+ + \vartheta^-}{2} + 1,\ \theta_j^+ \coloneqq \vartheta^+ - u_j + 1 - w_j^+,\ \theta_j^- \coloneqq \vartheta^- + u_j + 1 - w_j^-,\ \delta u_j \coloneqq -\frac{1}{2}\big(\vartheta^+(u_j+1) + \vartheta^-(u_j-1)\big).$$

Obviously, $\lambda \geq 0$ and δu_j is of class H^2. Furthermore, θ_j^+, $\theta_j^- \in \mathrm{C}^0([0,T])$ and are ≥ 0 by choice of ϑ^+, ϑ^-, and feasibility of u. Thereby (4.1) can easily be verified.

Thus, we deduce the following KKT-conditions for a minimizer of (P'). We denote by $M([0,T]) \cong \mathrm{C}^0([0,T])^*$ the set of regular Borel measures on $[0,T]$.

Theorem 4.6. *Let $\overline{u}$ be a local minimizer of (P'). Then there are Lagrange multipliers λ_j^+ (corresponding to the constraint $u_j \leq 1$), $\lambda_j^- \in M([0,T])$ (corresponding to $u_j \geq -1$), $j = 1, \ldots, N$, satisfying:*

i) (Primal feasibility): $|\overline{u}_j| \leq 1$.
ii) (Dual feasibility): $\lambda_j^+, \lambda_j^- \geq 0$, i.e., $\lambda_j^+ v, \lambda_j^- v \geq 0$ for all $v \in \mathrm{C}^0([0,T])$ with $v \geq 0$.
iii) (Complementary slackness): $\lambda_j^+(\overline{u}_j - 1) = 0$, $\lambda_j^-(\overline{u}_j + 1) = 0$.
iv) (Stationarity): For all $\delta u \in \big(\mathrm{H}^2([0,T])\big)^N$ it holds that

$$\Big\langle \rho_{\overline{f}} - \rho_d, \rho_{\delta f}\Big\rangle_{\mathrm{L}^2} + \beta \sum_{j=1}^N c_j \big(\langle \overline{u}_j, \delta u_j\rangle_{\mathrm{L}^2} + \beta_1 \langle \partial_t \overline{u}_j, \partial_t \delta u_j\rangle_{\mathrm{L}^2} + \beta_2 \langle \partial_t^2 \overline{u}_j, \partial_t^2 \delta u_j\rangle_{\mathrm{L}^2}\big) = \sum_{j=1}^N \big(\lambda_j^- - \lambda_j^+\big)\delta u_j$$

where δf is obtained by solving

$$\begin{aligned} \partial_t \delta f + \widehat{p} \cdot \partial_x \delta f + \big(\overline{E} - \widehat{p}^\perp \overline{B}\big) \cdot \partial_p \delta f &= -\big(\delta E - \widehat{p}^\perp \delta B\big) \cdot \partial_p \overline{f}, \\ \partial_t \delta E + \nabla_x^\perp \delta B &= -j_{\delta f} - \delta U, \\ \partial_t \delta B + \mathrm{curl}_x\, \delta E &= 0, \\ (\delta f, \delta E, \delta B)(0) &= 0 \end{aligned}$$

with $\delta U = \sum_{j=1}^N \delta u_j z_j$ and $\big(\overline{f}, \overline{E}, \overline{B}\big) = \Xi(\overline{u})$.

4.2.6. Adjoint equation

Considering the optimality conditions above, we note that we have to compute $\overline{\phi}'$ and thus the whole derivative Ξ' at an optimal point $\overline{u}$. However, there is a more efficient way, the adjoint approach, that is to say firstly solve the adjoint equation

$$\partial_y F(\Xi(u),u)^* q = -\partial_y \phi(\Xi(u),u)$$

for the adjoint state q and secondly compute

$$\overline{\phi}'(u) = \partial_u F(\Xi(u),u)^* q + \partial_u \phi(\Xi(u),u). \tag{4.2}$$

Here, $y=(f,E,B)$ denotes the state and $F(y,u)=0$ the PDE system.

In order to apply these considerations to our problem we have to define F suitably. Here, 'suitably' means that the differentiability of F and the differentiability of the control-to-state operator Ξ have to fit together. In other words, $F(y,u)$ should be differentiable with respect to the $\mathrm{C}^0\big(0,T;\mathrm{L}^2\big)$-norm in the state variable $y=(f,E,B)$. In the following let

$$M_R := \big\{(f,E,B)\in \mathrm{C}^2_c\big([0,T]\times\mathbb{R}^4\big)\times \mathrm{C}^2_c\big([0,T]\times\mathbb{R}^2;\mathbb{R}^2\big)\times \mathrm{C}^2_c\big([0,T]\times\mathbb{R}^2\big) \mid f(t,x,p)=0 \text{ for all } |p|\geq R\big\}$$

for some $R>0$, and let M_R be equipped with the $\mathrm{C}^0\big(0,T;\mathrm{L}^2\big)$-norm. Here, the index '$c$' means 'compactly supported with respect to x and p' (or x respectively). Furthermore, let

$$Z := \mathrm{H}^1\big([0,T]\times\mathbb{R}^4\big)^* \times \Big(\mathrm{H}^1\big([0,T]\times\mathbb{R}^2\big)^*\Big)^3 \times \mathrm{L}^2\big(\mathbb{R}^4\big)^* \times \Big(\mathrm{L}^2\big(\mathbb{R}^2\big)^*\Big)^3.$$

Now define $F_R\colon M_R\times\big(\mathrm{H}^2([0,T])\big)^N \to Z$ *via*

$$\begin{aligned}
&F_R((f,E,B),u)(g,h_1,h_2,h_3,a_1,a_2,a_3,a_4)\\
&=\Big(-\int_0^T\int\int\big(\partial_t g+\widehat{p}\cdot\partial_x g+\big(E-\widehat{p}^\perp B\big)\cdot\partial_p g\big)f\,\mathrm{d}p\mathrm{d}x\mathrm{d}t+\langle g(T),f(T)\rangle_{\mathrm{L}^2}-\langle g(0),f(0)\rangle_{\mathrm{L}^2},\\
&\quad\int_0^T\int(-E_1\partial_t h_1+B\partial_{x_2}h_1+j_{f,1}h_1+U_1h_1)\mathrm{d}x\mathrm{d}t+\langle h_1(T),E_1(T)\rangle_{\mathrm{L}^2}-\langle h_1(0),E_1(0)\rangle_{\mathrm{L}^2},\\
&\quad\int_0^T\int(-E_2\partial_t h_2-B\partial_{x_1}h_2+j_{f,2}h_2+U_2h_2)\mathrm{d}x\mathrm{d}t+\langle h_2(T),E_2(T)\rangle_{\mathrm{L}^2}-\langle h_2(0),E_2(0)\rangle_{\mathrm{L}^2},\\
&\quad\int_0^T\int(-B\partial_t h_3-E_2\partial_{x_1}h_3+E_1\partial_{x_2}h_3)\mathrm{d}x\mathrm{d}t+\langle h_3(T),B(T)\rangle_{\mathrm{L}^2}-\langle h_3(0),B(0)\rangle_{\mathrm{L}^2},\\
&\quad\int\int\Big(f(0)-\mathring{f}\Big)a_1\,\mathrm{d}p\mathrm{d}x,\int\Big(E_1(0)-\mathring{E}_1\Big)a_2\,\mathrm{d}x,\int\Big(E_2(0)-\mathring{E}\Big)a_3\,\mathrm{d}x,\int\Big(B(0)-\mathring{B}\Big)a_4\,\mathrm{d}x\Big)
\end{aligned}$$

where $U=\sum_{j=1}^N u_j z_j$. After several integrations by parts, it is obvious that (f,E,B) solves (CVM) with control U iff $F_R((f,E,B),u)=0$ for any $R>0$ with $\operatorname{supp}_p f\subset B_R$. Since no derivatives of the state $y=(f,E,B)$ appear above and the state is of class C^0_b, $\partial_y F_R$ exists and is given by

$$\begin{aligned}
&\partial_y F_R((f,E,B),u)(\delta f,\delta E,\delta B)(g,h_1,h_2,h_3,a_1,a_2,a_3,a_4)\\
&=\Bigg(-\int_0^T\int\int\big(\big(\partial_t g+\widehat{p}\cdot\partial_x g+\big(E-\widehat{p}^\perp B\big)\cdot\partial_p g\big)\delta f+\big(\delta E-\widehat{p}^\perp\delta B\big)f\cdot\partial_p g\big)\mathrm{d}p\mathrm{d}x\mathrm{d}t\\
&+\langle g(T),\delta f(T)\rangle_{\mathrm{L}^2}-\langle g(0),\delta f(0)\rangle_{\mathrm{L}^2},\\
&\quad\int_0^T\int(-\delta E_1\partial_t h_1+\delta B\partial_{x_2}h_1+j_{\delta f,1}h_1)\mathrm{d}x\mathrm{d}t+\langle h_1(T),\delta E_1(T)\rangle_{\mathrm{L}^2}-\langle h_1(0),\delta E_1(0)\rangle_{\mathrm{L}^2},\\
&\quad\int_0^T\int(-\delta E_2\partial_t h_2-\delta B\partial_{x_1}h_2+j_{\delta f,2}h_2)\mathrm{d}x\mathrm{d}t+\langle h_2(T),\delta E_2(T)\rangle_{\mathrm{L}^2}-\langle h_2(0),\delta E_2(0)\rangle_{\mathrm{L}^2},\\
&\quad\int_0^T\int(-\delta B\partial_t h_3-\delta E_2\partial_{x_1}h_3+\delta E_1\partial_{x_2}h_3)\mathrm{d}x\mathrm{d}t+\langle h_3(T),\delta B(T)\rangle_{\mathrm{L}^2}-\langle h_3(0),\delta B(0)\rangle_{\mathrm{L}^2},\\
&\quad\int\int\delta f(0)a_1\,\mathrm{d}p\mathrm{d}x,\int\delta E_1(0)a_2\,\mathrm{d}x,\int\delta E_2(0)a_3\,\mathrm{d}x,\int\delta B(0)a_4\,\mathrm{d}x\Bigg)
\end{aligned}$$

for $(\delta f,\delta E,\delta B)\in M_R$. Note that it is important that f vanishes for $|p|\geq R$ so that for $i=1,2$ the linear map

$$(f,E,B)\mapsto\left(h\mapsto\int_0^T\int j_{f,i}h\,\mathrm{d}x\mathrm{d}t\right)\in\mathrm{H}^1\big([0,T]\times\mathbb{R}^2\big)^*$$

is bounded due to

$$\left|\int_0^T\int j_{f,i}h\,\mathrm{d}x\mathrm{d}t\right|\leq C(T,R)\|f\|_{\mathrm{C}^0(0,T;\mathrm{L}^2)}\|h\|_{\mathrm{H}^1}$$

and hence differentiable.

On the other hand, we have

$$\partial_y\phi((f,E,B),u)(\delta f,\delta E,\delta B)=\langle\rho_f-\rho_d,\rho_{\delta f}\rangle_{\mathrm{L}^2}.$$

Here again, the support condition given in the definition of M_R is important to estimate

$$\left|\int_0^T\int(\rho_f-\rho_d)\rho_{\delta f}\,\mathrm{d}x\mathrm{d}t\right|\leq C(T,R)\|\rho_f-\rho_d\|_{\mathrm{L}^2}\|\delta f\|_{\mathrm{C}^0(0,T;\mathrm{L}^2)}$$

and

$$\int_0^T\int\rho_{\delta f}^2\,\mathrm{d}x\mathrm{d}t\leq C(T,R)\|\delta f\|^2_{\mathrm{C}^0(0,T;\mathrm{L}^2)}.$$

Now we search for an adjoint state

$$q=(g,h_1,h_2,h_3,a_1,a_2,a_3,a_4)\in Z^*\cong\mathrm{H}^1\big([0,T]\times\mathbb{R}^4\big)\times\big(\mathrm{H}^1\big([0,T]\times\mathbb{R}^2\big)\big)^3\times\mathrm{L}^2\big(\mathbb{R}^4\big)\times\big(\mathrm{L}^2\big(\mathbb{R}^2\big)\big)^3$$

satisfying the adjoint system. In other words, after integrating by parts once,

$$
\begin{aligned}
&-\int_0^T\int\int\big(\partial_t g+\widehat{p}\cdot\partial_x g+\big(E-\widehat{p}^{\perp}B\big)\cdot\partial_p g-4\pi(\widehat{p}_1h_1+\widehat{p}_2h_2)\big)\delta f\,\mathrm{d}p\mathrm{d}x\mathrm{d}t\\
&+\int_0^T\int\Big(-\partial_t h_1+\partial_{x_2}h_3+\int g\partial_{p_1}f\,\mathrm{d}p\Big)\delta E_1\,\mathrm{d}x\mathrm{d}t+\int_0^T\int\Big(-\partial_t h_2-\partial_{x_1}h_3+\int g\partial_{p_2}f\,\mathrm{d}p\Big)\delta E_2\,\mathrm{d}x\mathrm{d}t\\
&+\int_0^T\int\Big(-\partial_t h_3+\partial_{x_2}h_1-\partial_{x_1}h_2-\int g\widehat{p}^{\perp}\cdot\partial_p f\,\mathrm{d}p\Big)\delta B\,\mathrm{d}x\mathrm{d}t+\langle g(T),\delta f(T)\rangle_{\mathrm{L}^2}-\langle g(0)-a_1,\delta f(0)\rangle_{\mathrm{L}^2}\\
&+\langle h_1(T),\delta E_1(T)\rangle_{\mathrm{L}^2}-\langle h_1(0)-a_2,\delta E_1(0)\rangle_{\mathrm{L}^2}+\langle h_2(T),\delta E_2(T)\rangle_{\mathrm{L}^2}-\langle h_2(0)-a_3,\delta E_2(0)\rangle_{\mathrm{L}^2}\\
&+\langle h_3(T),\delta B(T)\rangle_{\mathrm{L}^2}-\langle h_3(0)-a_4,\delta B(0)\rangle_{\mathrm{L}^2}\\
=&-\int_0^T\int\int 4\pi(\rho_f-\rho_d)\delta f\,\mathrm{d}p\mathrm{d}x\mathrm{d}t \qquad (4.3)
\end{aligned}
$$

for all $(\delta f,\delta E,\delta B)\in M_R$. Therefore, the adjoint state solves the adjoint system

$$
\left.\begin{aligned}
\partial_t g+\widehat{p}\cdot\partial_x g+\big(E-\widehat{p}^{\perp}B\big)\cdot\partial_p g&=4\pi(\widehat{p}_1h_1+\widehat{p}_2h_2)+4\pi(\rho_f-\rho_d),\\
\partial_t h_1-\partial_{x_2}h_3&=\int g\partial_{p_1}f\,\mathrm{d}p',\\
\partial_t h_2+\partial_{x_1}h_3&=\int g\partial_{p_2}f\,\mathrm{d}p',\\
\partial_t h_3-\partial_{x_2}h_1+\partial_{x_1}h_2&=-\int g\widehat{p'}^{\perp}\cdot\partial_p f\,\mathrm{d}p',\\
(g,h_1,h_2,h_3)(T)&=0
\end{aligned}\right\} \qquad (\mathrm{Ad})
$$

for $|p|<R$. Since $R>0$ (with $\mathrm{supp}_p\, f\subset B_R$) is arbitrary, it is natural to demand (Ad) holds globally on $[0,T]\times\mathbb{R}^4$. Conversely, if (Ad) holds for all p, then (4.3) holds for all $(\delta f,\delta E,\delta B)\in M_R$ for any $R>0$ if we simply set $a_1=g(0)$, $(a_2,a_3,a_4)=(h_1,h_2,h_3)(0)$. The latter equations are unsubstantial and can be ignored.

In accordance with (4.2), we compute the derivative of $\overline{\phi}$ *via*

$$\overline{\phi}'(u)\delta u=\int_0^T\int(\delta U_1h_1+\delta U_2h_2)\mathrm{d}x\mathrm{d}t+\beta\sum_{j=1}^N c_j\big(\langle u_j,\delta u_j\rangle_{\mathrm{L}^2}+\beta_1\langle\partial_t u_j,\partial_t\delta u_j\rangle_{\mathrm{L}^2}+\beta_2\big\langle\partial_t^2u_j,\partial_t^2\delta u_j\big\rangle_{\mathrm{L}^2}\big)$$

where $\delta U=\sum_{j=1}^N\delta u_jz_j$.

System (Ad) has to be investigated. It is a final value problem which can easily be turned into an initial value problem *via* $\widetilde{g}(t,x,p)=g(T-t,-x,-p)$ and $\widetilde{h}(t,x)=h(T-t,-x)$, so that the left-hand sides of the differential equations in (Ad) do not change. In other words, the hyperbolic system (Ad) is time reversible.

To show unique solvability of (Ad), one can proceed similar to the dealing with (LVM). Yet there are some differences, which we will briefly sketch. Firstly, the source terms in the Maxwell equations are not the current densities induced by g but some other moments of g. Additionally, even in the fourth equation of (Ad) a source term appears. Hence, we have to prove analogues of Lemmas 2.3 and 2.5 with more general source terms. Secondly, the right-hand side of the Vlasov equation (and hence a solution g) does not have compact support with respect to p. But this will not cause any problems since in a representation formula for h there will appear a factor $\partial_p f$ (or first derivatives of $\partial_p f$). Because of the known fact that f is compactly supported with respect to p uniformly in t, x, we do not have to demand that g has this property. In Section 2.1 we had to assume this property for the density since the integral defining the current density induced by this density contains the factor $\widehat{p}$ which is obviously not compactly supported in p.

References

[1] R.J. Di Perna and P.L. Lions, Global weak solutions of Vlasov-Maxwell systems. *Commun. Pure Appl. Math.* **42** (1989) 729–757.

[2] S.S. Dragomir, Some Gronwall Type Inequalities and Applications. Nova Science Publishers (2003).

[3] O. Glass and D. Han-Kwan, On the controllability of the relativistic Vlasov-Maxwell system. *J. de Math. Pures Appl.* **103** (2015) 695–740.

[4] R.T. Glassey and J. Schaeffer, On the 'one and one-half dimensional' relativistic Vlasov–Maxwell system. *Math. Methods Appl. Sci.* **13** (1990) 169–179.

[5] R.T. Glassey and J. Schaeffer, The "two and one–half dimensional" relativistic Vlasov Maxwell system. *Commun. Math. Phys.* **185** (1997) 257–284.

[6] R.T. Glassey and J. Schaeffer, The relativistic Vlasov-Maxwell system in two space dimensions: part I. *Arch. Ration. Mech. Anal.* **141** (1998) 331–354.

[7] R.T. Glassey and J. Schaeffer, The relativistic Vlasov-Maxwell system in two space dimensions: part II. *Arch. Ration. Mech. Anal.* **141** (1998) 355–374.

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

[9] T. Kato, The Cauchy problem for quasi-linear symmetric hyperbolic systems. *Arch. Ration. Mech. Anal.* **58** (1975) 181–205.

[10] P. Knopf, Optimal control of a Vlasov–Poisson plasma by an external magnetic field. *Cal. Variat. Part. Differ. Equ.* **57** (2018) 134.

[11] P. Knopf and J. Weber, Optimal control of a Vlasov–Poisson plasma by fixed magnetic field coils. *Appl. Math. Optim.* **81** (2020) 961–988.

[12] M. Kunzinger, G. Rein, R. Steinbauer and G. Teschl, On classical solutions of the relativistic Vlasov-Klein-Gordon system. *Electron. J. Differ. Equ.* **2005** (2005) 1–17.

[13] T.T. Nguyen, T.V. Nguyen and W.A. Strauss, Global magnetic confinement for the 1.5D Vlasov-Maxwell system. *Kinet. Relat. Models* **8** (2015) 53.

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

[15] S. Robinson, Stability theory for systems of inequalities, part II: differentiable nonlinear systems. *SIAM J. Numer. Anal.* **13** (1976) 497–513.

[16] J. Zowe and S. Kurcyusz, Regularity and stability for the mathematical programming problem in Banach spaces. *Appl. Math. Optim.* **5** (1979) 49–62.