Optimal control of a Vlasov-Poisson plasma
by fixed magnetic field coils

Patrik Knopf

University of Regensburg, 93040 Regensburg, Bavaria, Germany

Patrik.Knopf@ur.de
orcid.org/0000-0003-4115-4885

Jörg Weber

University of Bayreuth, 95440 Bayreuth, Bavaria, Germany

Joerg.Weber@uni-bayreuth.de

Please cite as: P. Knopf, and J. Weber, Appl Math Optim (2018)
https://doi.org/10.1007/s00245-018-9526-5

Abstract

We consider the Vlasov-Poisson system that is equipped with an external magnetic field to describe the time evolution of the distribution function of a plasma. An optimal control problem where the external magnetic field is the control itself has already been investigated in [14]. However, in real technical applications it will not be possible to choose the control field in such a general fashion as it will be induced by fixed field coils. In this paper we will use the fundamentals that were established in [14] to analyze an optimal control problem where the magnetic field is a superposition of the fields that are generated by \( N \) fixed magnetic field coils. Thereby, the aim is to control the plasma in such a way that its distribution function matches a desired distribution function at some certain final time \( T \) as closely as possible. This problem will be analyzed with respect to the following topics: Existence of a globally optimal solution, necessary conditions of first order for local optimality, derivation of an optimality system, sufficient conditions of second order for local optimality and uniqueness of the optimal control under certain conditions.

Keywords: Vlasov-Poisson equation, optimal control with PDE constraints, nonlinear partial differential equations, calculus of variations.

MSC Classification: 49J20, 35Q83, 82D10.

1 Introduction

The three dimensional Vlasov-Poisson system in the plasma physical case is given by the following nonlinear system of partial differential equations:

\[
\begin{align*}
\partial_t f + v \cdot \partial_x f - \partial_x \psi \cdot \partial_v f &= 0, \\
-\Delta \psi &= 4\pi \rho, \\
\lim_{|x| \to \infty} \psi(t, x) &= 0, \\
\rho(t, x) &= \int f(t, x, v) \, dv.
\end{align*}
\]  

(1)

Here \( f = f(t, x, v) \geq 0 \) denotes the distribution function of the particle ensemble that is a scalar function representing the density in phase space. The time evolution of \( f \) is described by the first line of (1) which is a first order partial differential equation that is referred to as the Vlasov equation. For any measurable set \( M \subset \mathbb{R}^6 \), \( \int_M f(t, x, v) \, d(x, v) \) represents the total charge of the particles that have space coordinates \( x \in \mathbb{R}^3 \) and velocity coordinates \( v \in \mathbb{R}^3 \) with \( (x, v) \in M \) at time \( t \geq 0 \). The function \( \psi \) is the electrostatic potential that is induced by the charge of the particles. It is given by Poisson’s equation \(-\Delta \psi = 4\pi \rho\) with an homogeneous boundary condition where \( \rho \) denotes the volume charge density. The self-consistent electric field is then given by \(-\partial_x \psi\). Note that both \( \psi \) and \(-\partial_x \psi\) depend linearly on \( f \). Hence the Vlasov-Poisson system is nonlinear due to the term \(-\partial_x \psi \cdot \partial_v f\) in the Vlasov equation.
the Vlasov equation. Assuming \( f \) to be sufficiently regular (e.g., \( f(t) := f(t, \cdot, \cdot) \in C^1_c(\mathbb{R}^6) \) for all \( t \geq 0 \)), we can solve Poisson’s equation explicitly and obtain
\[
\psi_f(t, x) = \int \int \frac{f(t, y, w)}{|x - y|} \, dw \, dy, \quad \partial_x \psi_f(t, x) = -\int \int \frac{x - y}{|x - y|^3} f(t, y, w) \, dw \, dy
\]
(2)

Considering \( f \mapsto \psi_f \) (or \( f \mapsto -\partial_x \psi_f \) respectively) to be a linear operator we can formally rewrite the Vlasov-Poisson system as
\[
\partial_t f + v \cdot \partial_x f - \partial_x \psi_f \cdot \partial_v f = 0.
\]
(3)
The function \( f \) can be considered as the state variable. Later, we will additionally introduce a control function. Combined with the condition
\[
f|_{t=0} = \hat{f}
\]
(4)
for some function \( \hat{f} \in C^1_c(\mathbb{R}^6) \) we obtain an initial value problem. A first local existence and uniqueness result to this initial value problem was proved by R. Kurth [15]. Later J. Batt [2] established a continuation criterion which claims that a local solution can be extended as long as its velocity support is under control. Finally, two different proofs for global existence of classical solutions were established independently and almost simultaneously, one by K. Pfaffelmoser [20] and one by P.-L. Lions and B. Perthame [17]. Later, a greatly simplified version of Pfaffelmoser’s proof was published by J. Schaeffer [22]. This means that the following result is established: Any nonnegative initial datum \( f \in C^1_c([0, \infty[ \times \mathbb{R}^6) \) of the Vlasov-Poisson system (1) satisfying the initial condition (4). Moreover, for every time \( t \in [0, \infty[ \), \( f(t) = f(t, \cdot, \cdot) \) is compactly supported in \( \mathbb{R}^6 \). Hence equation (2) and the reformulation of the Vlasov-Poisson system (3) are well-defined in the case \( f \in C^1_c(\mathbb{R}^6) \). For more information we recommend to consider the review article [21] by G. Rein that gives an overview on the most important results.

To control the distribution function \( f \) we add an external magnetic field \( B = B(t, x) \) to the Vlasov equation:
\[
\partial_t f + v \cdot \partial_x f - \partial_x \psi_f \cdot \partial_v f + (v \times B) \cdot \partial_v f = 0, \quad f|_{t=0} = \hat{f}.
\]
(5)
The cross product \( v \times B \) occurs since, unlike the electric field, the magnetic field interacts with the particles via Lorentz force. From a theoretical point of view, the magnetic field \( B \) can be thought of as the control itself. In [14], an optimal control problem for the state equation (5) was investigated: The basics for calculus of variations were established and a model problem of optimal control with a tracking type cost functional was studied. For the analysis, high regularity assumptions on the control fields have been necessary. This, however, is not desirable in terms of optimal control theory and brings several disadvantages. Especially, the derivation of optimality conditions is restricted by these strong regularity requirements.

Also, from a technical perspective, it will not be possible to create magnetic control fields ad libitum. Therefore, in this paper, we suppose that the exterior magnetic field is induced by a finite number of fixed field coils. We assume that the current density \( J_i \) and the induced magnetic field \( B_i \) of the \( i \)-th coil are given by
\[
J_i(t, x) = j_i(t) c_i(x) \quad \text{and} \quad B_i(t, x) = u_i(t) m_i(x).
\]
with functions \( c_i, m_i : \mathbb{R}^3 \to \mathbb{R}^3 \) and \( j_i, u_i : \mathbb{R} \to \mathbb{R} \). The function \( c_i \) corresponds to the geometry of the coil while \( j_i \) corresponds to the intensity of the current that flows through the coil. Without loss of generality, we may assume that \( c_i \) has compact support. For simplicity, we suppose that the field \( B_i \) is determined by the Biot-Savart law which is a magnetostatic approximation of Maxwell’s equations. This is a very common approach in physics and engineering science to describe the magnetic field induced by a field coil. We obtain
\[
-u_i(t) \nabla \times m_i(x) = -\nabla \times B_i(t, x) = -4\pi J_i(t, x) = -4\pi j_i(t) c_i(x)
\]
which is solved by
\[
u_i(t) = j_i(t) \quad \text{and} \quad m_i(x) = \int_{\mathbb{R}^3} \frac{c_i(y) \times (x - y)}{|x - y|^3} \, dy.
\]
Note that \( m_i \) is source-free (i.e., \( \text{div}_x(m_i) = 0 \)). Then, due to linear superposition, the total external magnetic field is given by
\[
B(u)(t, x) = \sum_{i=1}^{N} u_i(t) m_i(x).
\]
Now, we will take the vector \( u = (u_1, \ldots, u_N)^T \) to be the control in our model. The complete magnetic field will be described by the control-to-field operator \( u \mapsto B(u) \). Mathematically, this brings a lot of advantages compared to the model discussed in [14]. This is mainly because no stronger regularity assumptions than \( u \in L^2([0, T]; \mathbb{R}^N) \) are needed. Therefore, the set of admissible controls can be chosen as a box restricted subset of \( L^2([0, T]; \mathbb{R}^N) \) which gives us better possibilities for establishing optimality conditions compared to [14]. For instance, we can show that a locally optimal control satisfies a certain projection formula. From this formula we derive the optimality system which in turn can be used to obtain a uniqueness result on small time intervals. Altogether, the model we study in this paper is also more viable for possible numerical implementations.

To establish the fundamentals of calculus of variations and optimal control for our model, we can basically exploit the theory that was developed in [14]. In the next section we will sketch the most important essentials.

Finally, we want to mention some works on similar topics we could find in literature. The controllability of the Vlasov-Poisson system where the control appears as a right-hand side of the Vlasov equation was studied by O. Glass and D. Han Kwan [11, 12]. They also established a similar result for the relativistic Vlasov-Maxwell system [13]. The time evolution and magnetic confinement of Vlasov-Poisson plasmas under the influence of an exterior magnetic field was investigated by S. Caprino, G. Cavallaro and C. Marchioro [5, 6, 7]. For further results on magnetic confinement of plasmas modeled by Vlasov equations see, e.g., [18, 19, 23].

2 Notation, preliminaries and basics

Our notation is mostly standard or self-explaining. However, to avoid misunderstandings, we fix some of it here.

Let \( d \in \mathbb{N}, U \subset \mathbb{R}^d \) be any open subset and \( k \in \mathbb{N} \) be arbitrary. \( C^k(U) \) denotes the space of \( k \) times continuously differentiable functions on \( U \). \( C^k(U) \) denotes the space of \( C^k(U) \)-functions having compact support in \( U \) and \( C^k_b(U) \) denotes the space of \( C^k(U) \)-functions that are bounded with respect to the norm

\[
\|u\|_{C^k_b(U)} := \sup_{|\alpha| \leq k} \|D^\alpha u\|_\infty = \sup_{|\alpha| \leq k} \sup_{x \in U} |D^\alpha u(x)|.
\]

Note that \((C^k_b(U), \| \cdot \|_{C^k_b(U)})\) is a Banach space.

For any measurable subset \( U \subset \mathbb{R}^d \), \( 1 \leq p \leq \infty \) and \( k \in \mathbb{N} \), \( L^p(U) \) denotes the standard \( L^p \)-space on \( U \) and \( W^{k,p}(U) \) denotes the standard Sobolov space on \( U \), as, for instance, defined by E. Lieb and M. Loss in [16, s.2.1,6.7]. Endowed with their standard norms defined by

\[
\|u\|_{L^p(U)} = \left( \int_U |u(x)|^p \, dx \right)^{\frac{1}{p}} \quad \text{and} \quad \|u\|_{W^{k,p}(U)} = \left( \sum_{|\alpha| \leq k} \|\partial^\alpha u(x)\|_{L^p(U)}^p \right)^{\frac{1}{p}}
\]

the spaces \( L^p(U) \) and \( W^{k,p}(U) \) are Banach spaces. The space \( L^2(U) \) is even a Hilbert space with the inner product

\[
\langle u, v \rangle_{L^2(U)} = \int_U u(x) v(x) \, dx.
\]

By \( L^p(U; \mathbb{R}^n) \) and \( W^{k,p}(U; \mathbb{R}^n) \) (with \( n \in \mathbb{N} \)) we denote the spaces of vector valued functions \( u: U \to \mathbb{R}^n \) whose components lie in \( L^p(U) \) or in \( W^{k,p}(U) \) respectively. The standard norms on \( L^p(U; \mathbb{R}^n) \) and \( W^{k,p}(U; \mathbb{R}^n) \) are defined analogously where the symbol \( \cdot \) now stands for the euclidean norm in \( \mathbb{R}^n \). If \( U = \mathbb{R}^d \) we will often omit the argument "\( (\mathbb{R}^d)^n \) or "\( (\mathbb{R}^d; \mathbb{R}^n) \)."

We will also use Banach space valued \( L^p \)-spaces (so-called Bochner spaces) as defined by L. C. Evans [9, p. 301-305]. For any Banach space \( X \) and any real number \( T > 0 \), \( L^p(0, T; X) \) denotes the space of Banach space valued \( L^p \)-functions \([0, T] \ni t \mapsto u(t) \in X\). In this case, \( L^p(0, T; X) \) is also a Banach space with the standard norm

\[
\|u\|_{L^p(0, T; X)} = \left( \int_0^T \|u\|_X^p \, dt \right)^{\frac{1}{p}}.
\]
**Proposition 2.1.** For $T > 0$ and $\beta > 3$ we define the Banach space $\mathcal{W}$ by

$$\mathcal{W} := L^2(0, T; W^{2, \beta}(\mathbb{R}^3; \mathbb{R}^3)) \quad \text{with} \quad \| \cdot \|_{\mathcal{W}} := \| \cdot \|_{L^2(0, T; W^{2, \beta}(\mathbb{R}^3; \mathbb{R}^3))}.$$ 

For any $K > 0$, the closed ball

$$\mathcal{B}_K := \left\{ B \in \mathcal{W} \mid \| B \|_{\mathcal{W}} \leq K \right\}$$

is referred to as the **set of admissible fields** (with radius $K$). Then the following holds:

(a) $\mathcal{B}_K$ is a bounded, convex and closed subset of $\mathcal{W}$.

(b) $\mathcal{B}_K$ is a subset of $L^2(0, T; C^{1, \gamma}(\mathbb{R}^3; \mathbb{R}^3))$ with $\gamma = 1 - 3/\beta$ and there exists some constant $C > 0$ that depends only on $\beta$ such that for all $B \in \mathcal{B}_K$,

$$\| B(t) \|_{C^{1, \gamma}} \leq C \| B(t) \|_{W^{2, \beta}} \quad \text{for almost all } t \in [0, T].$$

(c) $\mathcal{B}_K$ is a relatively weakly sequentially compact subset of $\mathcal{W}$.

This result has been established in [14, Lem.3] for $\mathcal{V} := \mathcal{W} \cap L^2(0, T; H^1)$ instead of $\mathcal{W}$. However, in the following approach it will not be necessary that the magnetic fields are in $L^2(0, T; H^1)$ so we will just drop this condition. One can easily see that the proofs of [14] that we are referring to hold true in this case.

It was also proved in [14, Thm.2] that any admissible field $B$ has its corresponding unique solution of the initial value problem (5). However, as these fields are only $L^2$ in time, the solution is not classical but merely strong.

**Proposition 2.2.** Let $B \in \mathcal{B}_K$ be any admissible field and suppose that $\hat{f} \in C^2(\mathbb{R}^6)$. Then there exists a unique strong solution $f_B$ of the initial value problem (5) to the field $B$, i.e., the following items hold:

(a) $f_B \in W^{1,2}(0, T; C_b(\mathbb{R}^6)) \cap C([0, T]; C^1_b(\mathbb{R}^6)) \cap L^\infty(0, T; W^{2,\beta}(\mathbb{R}^6))$ with

$$\| f_B(t) \|_p = \| \hat{f} \|_p , \quad t \in [0, T], \quad 1 \leq p \leq \infty,$$

$$\| f_B \|_{W^{1,2}(0, T; C_b)} + \| f_B \|_{C([0, T]; C^1_b)} + \| f_B \|_{L^\infty(0, T; W^{2,\beta})} \leq C$$

for some constant $C > 0$ depending only on $\hat{f}$, $T$, $K$ and $\beta$.

(b) $f_B$ satisfies the Vlasov equation

$$\partial_t f_B + v \cdot \partial_x f_B - \partial_x \psi_{f_B} \cdot \partial_v f_B + (v \times B) \cdot \partial_v f_B = 0$$

almost everywhere on $[0, T] \times \mathbb{R}^6$.

(c) $f_B$ satisfies the initial condition $f_B|_{t=0} = \hat{f}$ everywhere on $\mathbb{R}^6$.

(d) There exists some radius $R > 0$ depending only on $\hat{f}$, $T$, $K$ and $\beta$ such that supp $f_B(t) \subset B_R(0)$ for all $t \in [0, T]$.

With this knowledge it is possible to define an operator that maps any admissible field onto its corresponding solution. The most important attributes of this field-to-state operator have also already been established in [14]. We will summarize them in the following proposition:

**Proposition 2.3.** For any field $B \in \mathcal{B}_K$, let $f_B$ denote its corresponding strong solution of the state equation (5). The operator

$$\mathcal{S} : \mathcal{B}_K \to C([0, T]; L^2(\mathbb{R}^6)), \quad \mathcal{S}(B) := f_B$$

is called the **field-to-state operator**. It has the following essential properties:
(a) The field-to-state operator is Lipschitz/Hölder continuous in the following sense: There exists constants \( C_L, C_H > 0 \) depending only on \( f, T, K \) and \( \beta \) such that for all \( B, \tilde{B} \in \mathbb{B}_K \),

\[
\| \mathcal{S}(B) - \mathcal{S}(\tilde{B}) \|_{C([0,T];C_0)} \leq C_L \| B - \tilde{B} \|_W,
\]

\[
\| \mathcal{S}(B) - \mathcal{S}(\tilde{B}) \|_{W^{1,2}(0,T;C_0)} + \| \mathcal{S}(B) - \mathcal{S}(\tilde{B}) \|_{C([0,T];C_0^1)} \leq C_H \| B - \tilde{B} \|_W^\gamma
\]

where \( \gamma \) is the constant from Proposition 2.1.

(b) The field-to-state operator has the following property: Let \((B_k) \subset \mathbb{B}_K\) be any sequence with \( B_k \to B \in \mathbb{B}_K \) if \( k \to \infty \). Then there is a subsequence \((B_{k_j})\) of \((B_k)\) such that

\[
\mathcal{S}(B_{k_j}) \to \mathcal{S}(B) \quad \text{in} \quad W^{1,2}(0,T;L^2(\mathbb{R}^6))
\]

for \( j \to \infty \). Since \( \mathbb{B}_K \) is weakly compact, any sequence \((B_k) \subset \mathbb{B}_K\) actually has a subsequence that converges weakly to some \( B \in \mathbb{B}_K \). Therefore, the weak closure in \( W^{1,2}(0,T;L^2(\mathbb{R}^6)) \) of the image \( \mathcal{S}(\mathbb{B}_K) \) is weakly (sequentially) compact. This means that \( \mathcal{S} \) is a weakly compact operator.

(c) The field-to-state operator is Fréchet differentiable on \( \mathbb{B}_K \). For any field \( B \in \mathbb{B}_K \) and any direction \( H \in W \), the Fréchet derivative \( \mathcal{S}'(B)[H] = f_B[H] \) is the unique strong solution of the initial value problem

\[
\partial_t f + v \cdot \partial_x f - \partial_x \psi f + \partial_x \psi f \cdot \partial_x f_B + (v \times B) \cdot \partial_x f + (v \times H) \cdot \partial_x f_B = 0, \quad f|_{t=0} = 0. \tag{6}
\]

This means that \( f_B[H] \) lies in \( L^\infty([0,T] \times \mathbb{R}) \) and \( H^1([0,T] \times \mathbb{R}) \subset C([0,T];L^2) \), satisfies (6) almost everywhere and there exists some radius \( \rho > 0 \) depending only on \( T, K, f \) and \( \beta \) such that \( \sup_{t \in [0,T]} f_B[H](t) \subset B_{\rho}(0) \) for all \( t \in [0,T] \).

Moreover, the Fréchet derivative \( \mathcal{S}'(B)[H] \) depends Hölder-continuously on \( B \in \mathbb{B}_K \) in the following sense:

There exists some constant \( C > 0 \) depending only on \( f, T, K \) and \( \beta \) such that for all \( B, \tilde{B} \in \mathbb{B}_K \) and \( H \in \mathcal{V} \),

\[
\sup_{\| H \|_W \leq 1} \| \mathcal{S}'(B)[H] - \mathcal{S}'(\tilde{B})[H] \|_{L^2([0,T];L^2)} \leq C \| B - \tilde{B} \|_W^\gamma.
\]

where \( \gamma \) is the constant from Proposition 2.1.

In [14] these properties could finally be used to analyze an optimal control problem with a tracking type cost functional where the field \( B \) was the control itself.

3 The set of admissible controls and the control-to-state operator

As motivated in the introduction, we assume that our magnetic field \( B \) is induced by \( N \) field coils. We assume that each coil generates a magnetic field of a certain shape \( m_i = m_i(x) \) and its intensity at time \( t \) is determined by a factor \( u_i(t) \). This means that the magnetic field of the \( i \)-th coil is given by \( B_i(t,x) = u_i(t) m_i(x) \) and thus

\[
B(u)(t,x) = \sum_{i=1}^N u_i(t) m_i(x).
\]

is the total external magnetic field. The intensity factor \( u_i \) is directly proportional to the current that flows through the \( i \)-th coil. Now the vector \((u_1, ..., u_N)^T\) is the control in this model and the magnetic field can be interpreted as the function value of the operator \( u \mapsto B(u) \). We suppose that \( m_i \in W^{2,\beta}(\mathbb{R}^3;\mathbb{R}^3) \) for every index \( i \in \{1, ..., N\} \) and, since real magnetic fields are always source-free, we also assume that \( \text{div}_x m_i = 0 \) on \( \mathbb{R}^3 \). The control \( u \) is supposed to lie in \( L^2([0,T];\mathbb{R}^N) \) in order to ensure that the field \( B(u) \) has the desired regularity. This is specified by the following definition:

**Definition 3.1.** Let \( N \) be a fixed positive integer and let \( M > 0 \) be a real number. For every \( i \in \{1, ..., N\} \) let \( m_i = (m_{i1}, m_{i2}, m_{i3})^T \) be a fixed vector-valued function in \( W^{2,\beta}(\mathbb{R}^3;\mathbb{R}^3) \subset C^{1,\gamma}(\mathbb{R}^3;\mathbb{R}^3) \) with \( \| m_i \|_{W^{2,\beta}} \leq M \) and \( \text{div}_x m_i = 0 \) on \( \mathbb{R}^3 \). Moreover let \( a = (a_1, ..., a_N)^T \) and \( b = (b_1, ..., b_N)^T \) be fixed functions in \( L^2([0,T];\mathbb{R}^N) \) with \( a_i \leq 0 \leq b_i \) almost everywhere on \([0,T]\) for all \( i \in \{1, ..., N\} \). We define

\[
U_i := \{ w \in L^2([0,T]) \mid a_i \leq w \leq b_i \ a.e. \ on \ [0,T] \}, \quad i = 1, ..., N \quad \text{and} \quad U := U_1 \times \ldots \times U_N.
\]
The set $\mathbb{U}$ is referred to as the set of admissible controls. Moreover we define the operator

$$B(\cdot) : L^2([0,T];\mathbb{R}^N) \to L^2(0,T;W^{2,\beta}(\mathbb{R}^3;\mathbb{R}^3)), u \mapsto B(u) \quad \text{where} \quad B(u)(t,x) := \sum_{i=1}^{N} u_i(t) m_i(x).$$

The operator $B(\cdot)$ is referred to as the control-to-field operator and

$$S : \mathbb{U} \to C([0,T];L^2(\mathbb{R}^6)), u \mapsto S(u) := \mathcal{S}(B(u))$$

is called the control-to-state operator.

Note that the set of admissible controls is not empty as the zero function lies in $\mathbb{U}$. This definition does only make sense if the fields that are generated by the control-to-field operator are admissible in the sense of Proposition 2.1. In this case the state $S(u) = f_B(u)$ is well-defined but we also have to investigate how it depends on the control $u$:

**Proposition 3.2.**

(a) The set $\mathbb{U}$ is a bounded, convex and closed subset of $L^2([0,T];\mathbb{R}^N)$ and thus it is relatively weakly sequentially compact.

(b) The operator $B(\cdot)$ is linear and continuous and there exists some constant $K > 0$ depending only on $N, a, b$ and $M$ such that $B(\mathbb{U}) \subset \mathbb{B}_{K/2} \subset \mathbb{B}_K$. This means that the control-to-field operator maps admissible controls onto admissible fields.

(c) The control-to-field operator $B(\cdot)$ is continuously Fréchet differentiable and its Fréchet derivative at the point $u \in L^2([0,T];\mathbb{R}^N)$ is given by

$$B'(u)[h] = B(h) \quad \text{for all} \quad h \in L^2([0,T];\mathbb{R}^N).$$

(d) The control-to-state operator $S$ is Fréchet differentiable on $\mathbb{U}$ and its Fréchet derivative at the point $u \in \mathbb{U}$ is given by

$$S'(u)[h] = \frac{dS(B(u))}{du}[h] = \mathcal{S}'(B(u))[B(h)] \quad \text{for all} \quad h \in L^2([0,T];\mathbb{R}^N).$$

Recall that $\mathcal{S}'(B(u))[B(h)]$ is determined by the initial value problem (6). The Fréchet derivative depends H"older-continuously on $u$, i.e., there exists some constant $C > 0$ depending only on $\hat{f}, T, K$ and $\beta$ such that

$$\|S'(u_1) - S'(u_2)\|_{L^2(0,T;L^2)} \leq C \|u_1 - u_2\|^\gamma_{L^2(0,T;\mathbb{R}^N)}$$

for all $u_1, u_2 \in \mathbb{U}$ and $h \in L^2([0,T];\mathbb{R}^N)$ where $\gamma$ is the constant from Proposition 2.1.

Sometimes, for $u \in \mathbb{U}$ and $h \in L^2([0,T];\mathbb{R}^N)$, we will also write $f_u := f_B(u)$ to denote $S(u)$ and $f_u'[h]$ to denote the Fréchet derivative $S'(u)[h]$.

**Proof** For any $i \in \{1, \ldots, N\}$ the set $\mathbb{U}_i \subset L^2([0,T])$ is evidently bounded, convex and closed. Thus weak compactness follows directly from the theorems of Banach-Alaoglu and Mazur. The same holds for the set $\mathbb{U} \subset L^2([0,T];\mathbb{R}^N)$ which proves (a). The operator $B(\cdot)$ is obviously linear and for all $u \in L^2([0,T];\mathbb{R}^N)$,

$$\|B(u)\|_W \leq \sum_{i=1}^{N} \|u_i\|_{L^2([0,T])} \|m_i\|_{W^{2,\beta}} \leq M \sum_{i=1}^{N} \|u_i\|_{L^2([0,T])}$$

$$\leq M \sqrt{N} \left( \sum_{i=1}^{N} \|u_i\|^2_{L^2([0,T])} \right)^{1/2} = M \sqrt{N} \|u\|_{L^2([0,T];\mathbb{R}^N)}.$$

Hence $B(\cdot)$ is continuous. Moreover this yields

$$\|B(u)\|_{L^2(0,T;W^{2,\beta})} < M \sqrt{N} \left( \|a\|_{L^2([0,T];\mathbb{R}^N)} + \|b\|_{L^2([0,T];\mathbb{R}^N)} \right) = K, \quad u \in \mathbb{U}$$

and thus $B(\mathbb{U}) \subset \mathbb{B}_{K/2}$. This proves (b) which directly implies (c). Finally (d) follows directly from (b), (c), Proposition 2.3(c) and the chain rule. \qed
4 The optimal control problem

We will now suppose that \( \lambda_i \geq 0 \) for all \( i \in \{1, ..., N\} \) and \( \bar{f}, f_d \in C^2_0(\mathbb{R}^6) \) such that \( \|\bar{f}\|_p = \|f_d\|_p \) for all \( p \in [1, \infty) \). The aim is to find a control \( u \in \mathbb{U} \) such that the distribution function \( f \) matches the desired distribution function \( f_d \) as closely as possible. This is modeled by the following minimization problem:

Minimize \[ J(f, u) = \frac{1}{2}\|\bar{f}(T) - f_d\|^2_{L^2(\mathbb{R}^6)} + \sum_{i=1}^{N} \frac{\lambda_i}{2}\|u_i\|^2_{L^2(0,T)} \]

s.t.

\[ \bullet \text{ } u \text{ is an admissible control, i.e., } u \in \mathbb{U} \]

\[ \bullet \text{ } f \text{ is a strong solution of the Vlasov-Poisson system} \]

\[ \partial_t f + v \cdot \partial_x f - \partial_x \phi \cdot \partial_x f + (v \times B(u)) \cdot \partial_x f = 0, \quad f |_{t=0} = \bar{f} \]

Recall that \( B(u) \) always lies in \( \mathbb{B}_{K/2} \subset \mathbb{B}_K \) if the control \( u \) is admissible. This means that the term „strong solution to the field \( B(u) \)“ is well defined. Using the control-to-state operator, this problem can be reduced to

Minimize \[ J(u) = \frac{1}{2}\|S_T(u) - f_d\|^2_{L^2(\mathbb{R}^6)} + \sum_{i=1}^{N} \frac{\lambda_i}{2}\|u_i\|^2_{L^2(0,T)} \]

s.t. \( u \in \mathbb{U} \).

(7)

(8)

where \( S_T(u) = f_u(T) \) denotes the control-to-state operator evaluated at time \( T \).

4.1 Existence of a globally optimal solution

Of course, this optimal control problem does only make sense if it has at least one solution. This is established by the following Theorem:

**Theorem 4.1.** The optimization problem (8) possesses a globally optimal solution \( \bar{u} \), i.e., for all \( u \in \mathbb{U} \), \( J(\bar{u}) \leq J(u) \). If \( \lambda_i > 0 \) for some index \( i \in \{1, ..., N\} \), it holds that

\[ \|\bar{u}_i\|_{L^2(0,T)} \leq \frac{2}{\sqrt{\lambda_i}} \|\bar{f}\|_{L^2(\mathbb{R}^6)}. \]

**Proof** \( J \) is bounded from below since \( J(u) \geq 0 \) for all \( u \in \mathbb{U} \). Hence the infimum \( \mathcal{M} := \inf_{u \in \mathbb{U}} J(u) \) exists and there also exists a minimizing sequence \( (u_k)_{k \in \mathbb{N}} \subset \mathbb{U} \) such that \( J(u_k) \to \mathcal{M} \) if \( k \to \infty \). As \( \mathbb{U} \) is weakly compact we obtain \( u_k \rightharpoonup \bar{u} \) in \( L^2(0,T;\mathbb{R}^N) \) for some weak limit \( \bar{u} \in \mathbb{U} \) after extraction of a subsequence. This especially means that \( \{u_k\} \rightharpoonup \bar{u}_i \) in \( L^2(0,T) \) for every \( i \in \{1, ..., N\} \) and obviously also \( B(u_k) \rightharpoonup B(\bar{u}) \in \mathbb{B}_K \) in \( L^2(0,T;W^{2,\beta}) \). Then, we can conclude from Proposition 2.3(c) that \( f_{u_k} \to f_{\bar{u}} \) in \( W^{1,2}(0,T;L^2) \) up to a subsequence. By the fundamental theorem of calculus this directly implies that \( f_{u_k}(T) \to f_{\bar{u}}(T) \) in \( L^2(\mathbb{R}^6) \). Hence we can deduce from the weak lower semicontinuity of the \( L^2 \)-norm that

\[ J(\bar{u}) = \frac{1}{2}\|\bar{f}(T) - f_d\|^2_{L^2} + \sum_{i=1}^{N} \frac{\lambda_i}{2}\|\bar{u}_i\|^2_{L^2} \leq \liminf_{k \to \infty} \left[ \frac{1}{2}\|f_{u_k}(T) - f_d\|^2_{L^2} + \sum_{i=1}^{N} \frac{\lambda_i}{2}\|u_k\|^2_{L^2} \right] \]

\[ \leq \liminf_{k \to \infty} \left[ \frac{1}{2}\|f_{u_k}(T) - f_d\|^2_{L^2} + \sum_{i=1}^{N} \frac{\lambda_i}{2}\|u_k\|^2_{L^2} \right] = \lim_{k \to \infty} J(u_k) = \mathcal{M}. \]

By the definition of infimum this yields \( J(\bar{u}) = \mathcal{M} \). Now suppose that there exists some \( i \in \{1, ..., N\} \) such that \( \lambda_i > 0 \) and \( \|\bar{u}_i\|_{L^2(0,T)} > (2/\sqrt{\lambda_i}) \|\bar{f}\|_{L^2(\mathbb{R}^6)} \). Then

\[ J(\bar{u}) \geq \frac{\lambda_i}{2}\|\bar{u}_i\|^2_{L^2(0,T)} > \frac{1}{3} \left( \frac{2}{\sqrt{\lambda_i}} \|\bar{f}\|_{L^2(\mathbb{R}^6)} \right)^2 = \frac{1}{2} \left( \|f_0(T)\|_{L^2} + \|f_d\|_{L^2} \right) \geq \frac{1}{2} \|f_0(T) - f_d\|^2_{L^2} = J(0) \]

where 0 denotes the null function \((0, ..., 0)^T \in \mathbb{U} \). This, however, is a contradiction to the global optimality of \( \bar{u} \) and thus the asserted inequality follows.

(8)

Of course, this theorem does not provide uniqueness but only existence of a globally optimal solution. Since the control-to-state operator \( u \to f_u \) is nonlinear we cannot expect the cost functional \( J \) to be convex. This means that the optimal control problem may also have several locally optimal solutions. In the following subsection, these locally optimal solutions will be characterized by necessary optimality conditions of first order.
4.2 Necessary conditions for local optimality

Since our set of admissible controls is a box-restricted subset of $L^2([0,T];\mathbb{R}^N)$ this provides better possibilities to establish necessary optimality conditions compared to the model in [14] where the magnetic field was the control itself. As the basic approach will be quite similar we will also have to discuss the costate equation. In this context we will need the constant $R_Z > 0$ from [14, Lem.6]. It was defined in such a way that for any admissible field $B \in \mathbb{B}_K$ the solution $Z_B = Z_B(s,t,x,v)$ of the characteristic system

$$
\dot{x} = v, \quad \dot{v} = -\partial_x \varphi_{f_B}(s,x) + v \times B(s,x)
$$

with $Z_B(t,t,x,v) = (x,v)$ satisfies $\|Z(s,t,\cdot)\|_{L^\infty(B_R(0))} \leq R_Z$ for all $s,t \in [0,T]$. Without loss of generality, we may assume that $R_Z \geq R$.

**Proposition 4.2.** Let $u \in \mathcal{U}$ be arbitrary and let $f_u = S(u)$ be its induced state that is given by the control-to-state operator. Suppose that $\chi \in C^2_c(\mathbb{R}^6;[0,1])$ with $\chi = 1$ on $B_{R_Z}(0)$ and $\text{supp} \chi \subset B_{2R_Z}(0)$. Then the costate equation

$$
\partial_t g + v \cdot \partial_x g - \partial_x \varphi_{f_u} \cdot \partial_v g + (v \times B(u)) \cdot \partial_v g = \Phi_{f_u,g} \chi, \quad g|_{t=T} = f_u(T) - f_d \tag{9}
$$

where

$$
\Phi_{\varphi,\gamma}(t,x) := -\int\int \frac{x - y}{|x - y|^3} \cdot \partial_v \varphi(t,y,w) \gamma(t,y,w) \, dy \, dw, \quad t \in [0,T], x \in \mathbb{R}^3
$$

has a unique strong solution

$$
g_u \in W^{1,2}(0,T;C_b) \cap C([0,T];C^1_b(\mathbb{R}^6)) \cap L^\infty(0,T;H^2(\mathbb{R}^6))
$$

with $\text{supp} g_u(t) \subset B_{R^*}(0)$, $t \in [0,T]$ for some constant $R^* > 0$ depending only on $\dot{f}, f_d, T, K$ and $\beta$. Note that $g_u|_{B_R(0)}$ does not depend on the choice of $\chi$. The operator

$$
\mathcal{A} : \mathcal{U} \to C([0,T];L^2(\mathbb{R}^6)), \quad u \mapsto \mathcal{A}(u) := g_u
$$

is called the **control-to-costate operator**. It depends Lipschitz/Hölder-continuously on $u$ in such a way that there exists some constant $C \geq 0$ depending only on $\dot{f}, f_d, T, K$ and $\beta$ such that

$$
\|\mathcal{A}(u_1) - \mathcal{A}(u_2)\|_{C([0,T];C^2_b)} \leq C \|u_1 - u_2\|_{L^2([0,T];\mathbb{R}^N)},
$$

$$
\|\mathcal{A}(u_1) - \mathcal{A}(u_2)\|_{W^{1,2}(0,T;C_b)} + \|\mathcal{A}(u_1) - \mathcal{A}(u_2)\|_{C([0,T];C^1_b)} \leq C \|u_1 - u_2\|_{L^2([0,T];\mathbb{R}^N)}
$$

for all $u_1, u_2 \in \mathcal{U}$.

**Proof** Since $u \in \mathcal{U}$ and thus $B(u) \in \mathbb{B}_K$, this result follows directly from [14, Thm.5] and the estimate

$$
\|B(u_1) - B(u_2)\|_{W} \leq C \|u_1 - u_2\|_{L^2([0,T];\mathbb{R}^N)}, \quad u_1, u_2 \in L^2([0,T];\mathbb{R}^N)
$$

that is a direct consequence of Lemma 3.2(b).

Of course the costate equation (9) does not appear out of thin air. In the proof of Theorem 4.3, this equation will be derived by Lagrangian technique. Using the costate, various equivalent necessary conditions for local optimality can be established. In the following, the most important ones are presented.

**Theorem 4.3.** Suppose that $\lambda_i > 0$ for every $i \in \{1,...,N\}$ and let $\bar{u} \in \mathcal{U}$ be any function. According to the definition of the control-to-state operator $f_{\bar{u}}$ denotes the unique strong solution of the state equation to the field $B(\bar{u}) \in \mathbb{B}_K$. Moreover let $\tilde{g}_{\bar{u}}$ denote the unique strong solution of the costate equation (9). We define the function

$$
p(\bar{u}) : [0,T] \to \mathbb{R}^N \text{ by } p(\bar{u}) = (p_1(\bar{u}),...,p_N(\bar{u}))^T \text{ with }
$$

$$
p_i(\bar{u})(t) := \int (v \times m_i(x)) \cdot \partial_x f_{\bar{u}}(t,x,v) \, d(x,v), \quad i = 1,...,N.
$$

For every $\bar{u} \in \mathcal{U}$, $p(\bar{u}) \in C([0,T];\mathbb{R}^N)$. 

8
The following items are equivalent:

(NC1) \( \bar{u} \) satisfies the variational inequality, i.e., for all \( u = (u_1, \ldots, u_N) \in \mathbb{U} \),

\[
\int_0^T \left( \lambda_i \bar{u}_i - p_i(\bar{u}) \right) (u_i - \bar{u}_i) \, dt \geq 0, \quad i = 1, \ldots, N.
\]

(NC2) \( \bar{u} \) is given implicitly by the projection formula, i.e., for almost all \( t \in [0, T] \) and any \( i \in \{1, \ldots, N\} \),

\[
\bar{u}_i(t) = \mathbb{P}_{[a_i(t), b_i(t)]} \left( \frac{1}{\lambda_i} p_i(\bar{u})(t) \right)
\]

where \( \mathbb{P}_{[a,b]} \) denotes a projection of \( \mathbb{R} \) onto the interval \([a, b]\) that is given by

\[
\mathbb{P}_{[a,b]}(\xi) = \min \{ \max \{\xi, a\}, b\}, \quad \xi \in \mathbb{R}.
\]

Now suppose that \( \bar{u} \) is a locally optimal solution of the optimization problem (8), i.e., there exists \( \delta > 0 \) such that \( J(\bar{u}) \leq J(u) \) for every \( u \in \mathbb{U} \) with \( \|\bar{u} - u\|_{L^2} < \delta \). Then \( \bar{u} \) satisfies the assertions (NC1) and (NC2). This means that these items are (equivalent) necessary conditions for local optimality.

Comment

(a) We can establish similar results if \( \lambda_i = 0 \). The item (NC1) stays true in this case if we just replace \( \lambda_i \) by zero. Instead of (NC2) we only have

\[
\bar{u}_i(t) = a_i(t) \quad \text{if} \quad p_i(\bar{u})(t) > 0 \quad \text{and} \quad \bar{u}_i(t) = b_i(t) \quad \text{if} \quad p_i(\bar{u})(t) < 0
\]

but \( \bar{u}_i(t) \) is undefined if \( p_i(\bar{u})(t) = 0 \).

(b) If \( a_i \) and \( b_i \) are continuous, so is \( \bar{u}_i \) due to item (NC2). If this holds for all \( i \in \{1, \ldots, N\} \) we know that \( \bar{u} \) is continuous and consequently \( B(\bar{u}) \in C([0, T]; C^1) \). In this case \( f_{\bar{u}} \) and \( g_{\bar{u}} \) are classical solutions of their respective systems.

Proof The assertion \( p_i(\bar{u}) \in C([0,T]) \) is obvious since \( f_{\bar{u}}, g_{\bar{u}} \in C([0,T]; C^1) \) and \( m_i \in C^{1,\gamma}(\mathbb{R}^3; \mathbb{R}^3) \) for all indices \( i \in \{1, \ldots, N\} \). First we will show that item (NC1) holds if \( \bar{u} \) is a locally optimal solution. Therefore we apply the Lagrangian technique: For \( u \in \mathbb{U} \) and \( f, g \in H^1([0,T] \times \mathbb{R}^6) \) with supp \( f(t) \subset B_R(0) \) for all \( t \in [0,T] \) we define the Lagrangian

\[
\mathcal{L}(f, u, g) := \frac{1}{2} \|f(T) - f_d\|_{L^2}^2 + \sum_{i=1}^N \frac{\lambda_i}{2} \|u_i\|^2_{L^2} - \int_{[0,T] \times \mathbb{R}^6} \left( \partial_t f + v \cdot \partial_x f - \partial_x \psi_f \cdot \partial_v f + (v \times B(u)) \cdot \partial_v f \right) g \, dt(x, v).
\]

It is possible to replace \([0,T] \times \mathbb{R}^6\) by \([0,T] \times B_R(0)\) because of the support condition on \( f \). As \( \text{div}_v(v \times B(u)) = 0 \), integration by parts yields

\[
\mathcal{L}(f, u, g) = \frac{1}{2} \|f(T) - f_d\|_{L^2}^2 + \sum_{i=1}^N \frac{\lambda_i}{2} \|u_i\|^2_{L^2} + \langle g(0), f(0) \rangle_{L^2} - \langle g(T), f(T) \rangle_{L^2}
\]

\[
+ \int_{[0,T] \times B_R(0)} \left( \partial_t g + v \cdot \partial_x g - \partial_x \psi_f \cdot \partial_v g + (v \times B(u)) \cdot \partial_v g \right) f \, dt(x, v).
\]

The Lagrangian is partially Fréchet differentiable with

\[
(\partial_f \mathcal{L})(f, u, g)[h] = \langle f(T) - f_d, h(T) \rangle_{L^2} + \langle g(T), h(T) \rangle_{L^2} + \langle g(0), h(0) \rangle_{L^2}
\]

\[
+ \int_{[0,T] \times B_R(0)} \left( \partial_t g + v \cdot \partial_x g - \partial_x \psi_f \cdot \partial_v g + (v \times B(u)) \cdot \partial_v g \right) h \, dt(x, v) - \int_{[0,T] \times B_R(0)} \Phi_{f,g}(t, x) \, h \, dt(x, v).
\]
for any \( h \in H^1([0, T] \times \mathbb{R}^6) \) and

\[
(\partial_u \mathcal{L})(f, u, g)[h] = \sum_{i=1}^{N} \lambda_i(u_i, h_i)_{L^2} - \int_{[0, T] \times \mathbb{R}^6} (v \times B(h)) \cdot \partial_v f g \, d(t, x, v)
\]

for any \( h \in L^2([0, T]; \mathbb{R}^N) \). Obviously \( J(\bar{u}) = \mathcal{L}(f, \bar{u}, g) \) and hence

\[
J'(\bar{u})[h] = (\partial_f \mathcal{L})(f, \bar{u}, g)[f'_u[h]] + (\partial_u \mathcal{L})(f, \bar{u}, g)[h]
\]

for all \( h \in L^2([0, T]; \mathbb{R}^N) \). Now, if \( \bar{u} \) is a local minimizer of \( J \) then \( J'(\bar{u})[h] \) is nonnegative for all directions \( h \in L^2([0, T]; \mathbb{R}^N) \) with \( u + h \in U \). Thus inserting \( g = g_u \) yields \((\partial_f \mathcal{L})(f, \bar{u}, g)[f'_u[h]] \geq 0 \) and thus

\[
0 \leq J'(\bar{u})h = (\partial_u \mathcal{L})(f, \bar{u}, g)[h] = \sum_{i=1}^{N} \int_{0}^{T} (\lambda_i \bar{u}_i(t) - p_i(\bar{u})(t)) h_i(t) \, dt
\]

(10)

for all \( h \in L^2([0, T]; \mathbb{R}^N) \) with \( u + h \in U \). For any fixed \( i \in \{1, \ldots, N\} \) we can choose \( h_j = 0 \) if \( j \neq i \) while \( h_i \) is still arbitrary. This finally implies that

\[
\int_{0}^{T} (\lambda_i \bar{u}_i - p_i(\bar{u})) \, h_i(t) \, dt \geq 0, \quad i = 1, \ldots, N
\]

for all \( h \in L^2([0, T]; \mathbb{R}^N) \) with \( u + h \in U \). For any \( u \in U \) we can now choose \( h := u - \bar{u} \in L^2([0, T]; \mathbb{R}^N) \) and hence we can conclude that for all \( u \in U \),

\[
\int_{0}^{T} (\lambda_i \bar{u}_i - p_i(\bar{u})) \, (u_i - \bar{u}_i) \, dt \geq 0, \quad i = 1, \ldots, N
\]

that is (NC1). The equivalence of (NC1) and (NC2) is a standard result (see, e.g., [24, pp. 67-71]).

If \( \bar{u} \in U \) is a locally optimal control we can also show that the triple \((f, g, u)\) satisfies a certain system of partial differential equation that will be referred to as the optimality system of the optimization problem. A strong solution of the optimality system is defined as follows:

**Definition 4.4.** Suppose that \( \lambda_i > 0 \) for all \( i \in \{1, \ldots, N\} \). The triple \((f, g, u)\) is called a strong solution of the optimality system if the following conditions hold:

(i) \( f, g \in W^{1,2}(0, T; C_b) \cap C([0, T]; C^1_b) \) and \( u \in L^2([0, T]; \mathbb{R}^N) \).

(ii) For any \( t \in [0, T] \), \( \text{supp} \, f(t) \subset B_R(0) \) and \( \text{supp} \, g(t) \subset B_R^*(0) \) where \( R, R^* > 0 \) are the constants from Theorem 2.2 and Theorem 4.2.

(iii) \( f, g \) and \( u \) satisfy the following system of equations almost everywhere:

\[
\begin{cases}
\partial_t f + v \cdot \partial_x f - \partial_x \psi f \cdot \partial_v f + (v \times B(u)) \cdot \partial_v f = 0, & f|_{t=0} = \hat{f}, \\
\partial_t g + v \cdot \partial_x g - \partial_x \psi f \cdot \partial_v g + (v \times B(u)) \cdot \partial_v g = \Phi_{f,g} \chi, & g|_{t=T} = f(T) - f_d, \\
u = (u_1, \ldots, u_N)^T & \text{with } u_i = P_{[a_i, b_i]} \left( \frac{1}{N} \int (v \times m_i) \cdot \partial_v f g \, d(x, v) \right),
\end{cases}
\]

(11)
We obtain the following condition for local optimality:

**Corollary 4.5.** Suppose that $\bar{u} \in U$ is a locally optimal solution of the optimization problem (8). Then the following holds:

\[(NC3) \quad (f_{\bar{u}}, g_{\bar{u}}, \bar{u}) \text{ is a strong solution of the optimality system (11).}\]

Moreover, condition (NC3) is equivalent to the necessary optimality conditions (NC1) and (NC2).

**Proof** It is obvious that condition (NC3) is equivalent to (NC2). Hence it is a necessary condition for local optimality that is equivalent to the conditions (NC1) and (NC2) according to Theorem 4.3.

Note that for any locally optimal control $\bar{u}$ the costate $g_{\bar{u}}$ can be considered as a Lagrangian multiplier with respect to the state condition on $f_{\bar{u}}$ that is the state equation. However, we can also find Lagrangian multipliers with respect to the control restrictions. This makes it possible to write the optimal control problem as a Karush-Kuhn-Tucker system.

**Corollary 4.6.** Suppose that $\bar{u} \in U$ is a locally optimal solution of the optimization problem (8). Then the following holds:

\[(NC4) \quad \bar{u} \text{ satisfies the Karush-Kuhn-Tucker conditions, i.e., there exist Lagrangian multipliers } \mu^a, \mu^b \text{ in } L^2([0,T]; \mathbb{R}^N) \text{ such that:}\]

(i) **Primal feasibility:** For all $i \in \{1, ..., N\}$,

$$a_i \leq \bar{u}_i \leq b_i \quad \text{almost everywhere on } [0,T].$$

(ii) **Dual feasibility:** For all $i \in \{1, ..., N\}$,

$$\mu_i^a, \mu_i^b \geq 0 \quad \text{almost everywhere on } [0,T].$$

(iii) **Complementary slackness:** For all $i \in \{1, ..., N\}$,

$$\mu_i^a(\bar{u}_i - a_i) = 0, \quad \mu_i^b(\bar{u}_i - b_i) = 0 \quad \text{almost everywhere on } [0,T].$$

(iv) **Stationarity:** For all $i \in \{1, ..., N\}$,

$$\lambda_i \bar{u}_i - p_i(\bar{u}) - \mu_i^a + \mu_i^b = 0 \quad \text{almost everywhere on } [0,T].$$

Moreover, condition (NC4) is equivalent to the necessary optimality conditions (NC1) and (NC2).

Since this is a standard result, we do not present a detailed proof. For a similar proof see [24, pp.71-73].

### 4.3 Uniqueness of the optimal solution on small time intervals

We can show that the solution of the optimality system is unique if the final time $T$ is small compared to the regularization parameters $\lambda_i$. From this, we can conclude that the optimal control is unique on sufficiently small time intervals. This approach to obtain uniqueness has already been used in literature in similar situations (see, e.g., [4, 10]). We also want to mention the paper [1] where a uniqueness result for the optimal control was established under the condition that the costate is bounded in a certain $L^q$-norm.

**Theorem 4.7.** Let $\lambda > 0$ be defined by $\lambda := \min\{\lambda_1, ..., \lambda_N\}$. Suppose that $\lambda \in [0,1]$ and let us assume that there exists a strong solution $(f, g, u)$ of the optimality system (11). Then this solution is unique if the quotient $\frac{T}{\lambda}$ is sufficiently small.
Proof Suppose that the triple \((\tilde{f}, \tilde{g}, \tilde{u})\) is another strong solution. Let \(C \geq 0\) denote some generic constant that may depend on \(T, N, M, a, b, f\) and \(f_d\). Since \(\text{supp} \ f(t)\), \(\text{supp} \ \tilde{f}(t) \subset B_R(0)\) we have for almost all \(t \in [0, T]\)

\[
|u(t) - \tilde{u}(t)| = \left| \mathbb{P}_{[a_i(t), b_i(t)]} \left( \frac{1}{\lambda_i} p_i(u)(t) \right) - \mathbb{P}_{[a_i(t), b_i(t)]} \left( \frac{1}{\lambda_i} p_i(\tilde{u})(t) \right) \right|
\]

\[
\leq \frac{1}{\lambda_i} \left| p_i(u)(t) - p_i(\tilde{u})(t) \right| ^{\lambda_i}
\]

\[
\leq \frac{C}{\lambda} \left\| g_u \right\|_{C(\{0,T]\};C_{\kappa})} \left\| f(u) - f(\tilde{u}) \right\|_\infty + \frac{C}{\lambda} \left\| f(\tilde{u}) \right\|_{C(\{0,T]\};C_{\lambda})} \left\| g_u(t) - g(\tilde{u}) \right\|_\infty
\]

and hence

\[
\left\| B(u)(t) - B(\tilde{u})(t) \right\|_\infty \leq \frac{C}{\lambda} \left\| f(u) - f(\tilde{u}) \right\|_\infty + \frac{C}{\lambda} \left\| g_u(t) - g(\tilde{u}) \right\|_\infty
\]

where now \(C = C(T)\) also depends on \(\|g_u\|_{C(\{0,T]\};C_{\kappa})}\) and \(\|f\|_{C(\{0,T]\};C_{\lambda})}\) and is monotonically increasing in \(T\). Note that, for any fixed \(T > 0\), the \(C(\{0,T]\};C_{\lambda})\)-norm of \(\tilde{u}\) is bounded uniformly for any \(\tilde{u} \in U\) by virtue of Proposition 2.2(a) so that the new constant \(C(T)\) can be assumed to be independent of \((\tilde{f}, \tilde{g}, \tilde{u})\).

Let now \(Z_u = Z_u(s, t, z)\) denote the solution of the characteristic system

\[
\dot{x} = v, \quad \dot{v} = -\partial_x \psi_f(s, x) + v \times B(u)(s, x)
\]

to the initial condition \(Z_u(t, t, z) = z\) and let \(Z_{\tilde{u}}\) be defined analogously. Now, let \(s, t \in [0, T]\) (without loss of generality, \(s \leq t\)) and \(z \in \mathbb{R}^6\) be arbitrary. Then

\[
|Z_u(s, t, z) - Z_{\tilde{u}}(s, t, z)| \leq \int_s^t |Z(\tau, t, z) - Z_{\tilde{u}}(\tau, t, z)| + \|\partial_x \psi_f(\tau)\|_\infty + R \left\| B(u)(\tau) - B(\tilde{u})(\tau) \right\|_\infty d\tau
\]

\[
\leq \int_s^t |Z(\tau, t, z) - Z_{\tilde{u}}(\tau, t, z)| + \frac{C}{T} \left\| f(\tau) - \tilde{f}(\tau) \right\|_\infty + \frac{C}{T} \left\| g(\tau) - \tilde{g}(\tau) \right\|_\infty d\tau.
\]

and thus by Gronwall’s lemma,

\[
|Z_u(s, t, z) - Z_{\tilde{u}}(s, t, z)| \leq \frac{C}{T} \int_s^t \left\| f(\tau) - \tilde{f}(\tau) \right\|_\infty + \left\| g(\tau) - \tilde{g}(\tau) \right\|_\infty d\tau.
\]

Analogously to the proof of [14, Thm. 8], we can conclude that

\[
\left\| f - \tilde{f} \right\|_{C(\{0,T]\};C_{\kappa})} \leq C_T \exp \left( C_T \frac{T}{\lambda} \right) \|g - \tilde{g}\|_{C(\{0,T]\};C_{\lambda})},
\]

(13)

and consequently

\[
\left\| f - \tilde{f} \right\|_{C(\{0,T]\};C_{\kappa})} \leq \left( C_T \frac{T}{\lambda} \right)^2 \exp \left( C_T \frac{T}{\lambda} \right) \|f - \tilde{f}\|_{C(\{0,T]\};C_{\lambda})}.
\]

If now \(\frac{T}{\lambda}\) is sufficiently small, we have \(\left( C_T \frac{T}{\lambda} \right)^2 \exp \left( C_T \frac{T}{\lambda} \right) < 1\) and it follows that \(\|f - \tilde{f}\|_{C(\{0,T]\};C_{\kappa})} = 0\). This means that \(f = \tilde{f}\) and then we obtain \(g = \tilde{g}\) by (14) and finally \(u = \tilde{u}\) by (12).

Finally, we can easily deduce uniqueness of the globally optimal solution if \(\frac{T}{\lambda}\) is small:

**Corollary 4.8.** Let \(\lambda\) denote the parameter from Theorem 4.7. Again, we assume that \(\lambda \in [0, 1]\) and that \(\frac{T}{\lambda}\) is sufficiently small. Then the following holds:

(a) The optimal control problem (8) has a unique globally optimal solution.
We already know from [14, Lem. 10] (which obviously holds true if the space is sufficiently small, we can conclude from Theorem 4.7 that \( \bar{u} \) is uniquely determined by the optimality system. Hence there is exactly one globally optimal solution of the optimal control problem and that is \( \bar{u} \). This proves (a).

If \( \bar{u} \) is the unique globally optimal solution it is obviously also locally optimal and satisfies the equivalent conditions for local optimality (NC1)-(NC4). To prove the reverse implication, we will now assume that there exists some control \( u \in U \) that satisfies the conditions (NC1)-(NC4). It remains to show that \( u \) is then the unique globally optimal solution. Recall that according to item (a) there is at least one unique globally optimal solution \( \bar{u} \) of the optimal control problem. As both \( u \) and \( \bar{u} \) satisfy the necessary condition (NC3) it must hold that \( u = \bar{u} \) if \( \frac{\lambda}{\bar{\xi}} \) is small enough. This proves the equivalence assertion of item (b). Therefore, in this case, (NC1)-(NC4) are necessary and sufficient conditions for global optimality as they are satisfied only by the unique globally optimal solution.

\[ \square \]

### 4.4 A sufficient condition for local optimality

In the previous section, we have showed that the necessary conditions (NC1)-(NC4) are also sufficient conditions if \( \frac{\lambda}{\bar{\xi}} \) is small enough. We will now establish a sufficient condition for (even strict) local optimality without direct restrictions on \( T \) and \( \lambda \). Therefore we will need the cost functional \( J \) to be twice continuously Fréchet differentiable. Unfortunately, we cannot prove that the control-to-state operator \( S \) is Fréchet differentiable, which is due to the following: Linearizing (6) once more, we see that the Vlasov equation which \( \mathcal{S}'(u)[h, \bar{h}] \) should satisfy contains terms \( \partial_v(\mathcal{S}'(u)[h]) \) and \( \partial_v(\mathcal{S}'(u)[\bar{h}]) \) as parts of a source term. Since \( \mathcal{S}'(u)[h] \) and \( \mathcal{S}'(u)[\bar{h}] \) are only of class \( H^1 \), the above derivatives are only of class \( L^2 \) which is not enough to show solvability of that new twice linearized Vlasov equation (cf. [14]). Also, standard approaches via the implicit function theorem fail due to the loss of regularity of solutions to the linearized equation (6).

However, it will do fine to have first-order Fréchet differentiability of the field-costate operator \( A \). This is established by the following Lemma:

**Proposition 4.9.** The field-costate operator \( A \) is Fréchet differentiable on \( U \). For any control \( u \in U \) and any direction \( h \in L^2([0, T]; \mathbb{R}^N) \), the Fréchet derivative \( A'(u)[h] = g_u[h] \) is the unique strong solution of the initial value problem

\[
\begin{align*}
\partial_t g + v \cdot \partial_x g - \partial_x \mathcal{A}'(u)[h] \cdot \partial_v g_u - \partial_x \mathcal{A}'(u)[h] \cdot \partial_v g + (v \cdot B(u)) \cdot \partial_v g + (v \cdot B(h)) \cdot \partial_v g_B &= \Phi_{f_u, g} \chi - \Phi_{g_u, f_u, h} \chi, \\
g|_{t=T} &= 0.
\end{align*}
\]

where \( \Phi \) is the operator from Proposition 4.2. This means that \( A'(u)[h] \) lies in \( (L^\infty \cap H^1)([0, T]; \mathbb{R}^N) \cap C([0, T]; L^2) \), satisfies (15) almost everywhere and there exists some radius \( g > 0 \) depending only on \( T, K, \bar{f} \) and \( \beta \) such that \( \text{supp} g_u[h][t] \subset B_0(0) \) for all \( t \in [0, T] \).

Moreover, the Fréchet derivative \( A'(u)[h] \) depends Hölder-continuously on \( u \in U \) in the following sense: There exists some constant \( C > 0 \) depending only on \( \bar{f}, T, K \) and \( \beta \) such that for all \( u, \bar{u} \in U \) and \( h \in L^2([0, T]; \mathbb{R}^N) \),

\[
\sup_{\|h\|_{L^2} \leq 1} \| A'(u)[h] - A'(\bar{u})[h] \|_{L^2(0, T; L^2)} \leq C \| u - \bar{u} \|_{L^2([0, T]; \mathbb{R}^N)}.
\]

**Proof** We already know from [14, Lem. 10] (which obviously holds true if the space \( \mathcal{V} = W \cap L^2(0, T; H^1) \) is replaced by \( W \)) that the operator

\[ g : B_K \to C([0, T]; L^2(\mathbb{R}^6)), \quad B \mapsto g_B \]

is Fréchet differentiable. The Fréchet derivative \( g'_B[H] \) is the unique strong solution of

\[
\begin{align*}
\partial_t g + v \cdot \partial_x g - \partial_x \mathcal{A}'(u)[h] \cdot \partial_v g_u - \partial_x \mathcal{A}'(u)[h] \cdot \partial_v g + (v \cdot B) \cdot \partial_v g + (v \cdot H) \cdot \partial_v g_B &= \Phi_{f, g} \chi - \Phi_{g, f, h} \chi, \\
g|_{t=T} &= 0.
\end{align*}
\]
This means that \( g_{\dot{B}}[H] \) lies in \( L^\infty \cap H^1([0,T] \times \mathbb{R}) \subset C([0,T]; L^2) \) satisfies (17) almost everywhere and there exists \( \varrho > 0 \) depending only on \( T, K, \dot{f} \) and \( \beta \) such that \( \text{supp} g_{\dot{B}}[H](t) \subset B_\varrho(0) \) for all \( t \in [0,T] \) and

\[
\sup_{\|H\|_{W} \leq 1} \|g_A'[H] - g_B'[H]\|_{L^2(0,T;L^2)} \leq C \|A - B\|_W, \quad A, B \in \mathbb{B}_K.
\]  

(18)

As the control-to-field operator \( u \mapsto B(u) \) is Fréchet differentiable according to Proposition 3.2(c), the chain rule implies that the control-costate operator \( u \mapsto \dot{g}_u = g_{\dot{B}(u)} \) is also Fréchet differentiable and the Fréchet derivative is given by \( g_u'[\dot{h}] = g_{\dot{B}(u)}'[B(\dot{h})] \). Hence, \( g_u'[h] \) is the unique strong solution of (15) with \( \text{supp} g_u'[h](t) \subset B_\varrho(0) \) for all \( t \in [0,T] \) and condition (16) holds.

From this result we can conclude that the cost functional is twice continuously Fréchet differentiable.

**Proposition 4.10.**

(a) The cost functional \( J \) of the optimization problem (8) is twice Fréchet differentiable on \( U \). The Fréchet derivative of second order at the point \( u \in U \) can be described as a bilinear operator

\[
J''(u) : L^2([0,T]; \mathbb{R}^N)^2 \rightarrow \mathbb{R},
\]

\[
J''(u)[h, \tilde{h}] = \sum_{i=1}^N \left\{ \lambda_i \langle h_i, \tilde{h} \rangle_{L^2([0,T])} - \int_{[0,T] \times \mathbb{R}^n} (v \times m_i) \cdot (\partial_v f_u g_u'[\tilde{h}] - \partial_v g_u f_u'\tilde{h}) h_i \, d(t,v) \right\}
\]

for all \( h, \tilde{h} \in L^2([0,T]; \mathbb{R}^N) \).

(b) There exists some constant \( C > 0 \) depending only on \( \dot{f}, f_d, N, M, a, b, T \) and \( \beta \) such that for all \( u, \tilde{u} \in U \),

\[
\|J''(u) - J''(\tilde{u})\| \leq C \|u - \tilde{u}\|_{L^2([0,T]; \mathbb{R}^N)}^2
\]

(19)

where

\[
\|J''(u)\| = \sup \left\{ \|J''(u)[h_1, h_2]\| : \|h_1\|_{L^2([0,T]; \mathbb{R}^N)} = 1, \|h_2\|_{L^2([0,T]; \mathbb{R}^N)} = 1 \right\}
\]

denotes the operator norm. This means that \( J \) is twice continuously differentiable.

**Proof.** Let \( C > 0 \) denote some generic constant depending only on \( \dot{f}, f_d, N, M, a, b, T \) and \( \beta \). In (10) we have already proved that

\[
J'(u)[h] = \sum_{i=1}^N \int_0^T \left[ \lambda_i u_i(t) - \left( \int (v \times m_i) \cdot \partial_v g_u d(x,v) \right) h_i(t) \right] \, dt
\]

for any \( u \in U \) and all \( h \in L^2([0,T]; \mathbb{R}^N) \). We will now prove that \( J'(u) \) is once more Fréchet differentiable with respect to \( u \). For \( \varepsilon > 0 \) we define the sets

\[
U_i^\varepsilon := \left\{ u \in L^2([0,T]; \mathbb{R}^N) : \|u\|_{L^2([0,T])} < a_i \|h_i\|_{L^2([0,T])} + \|b_i\|_{L^2([0,T])} + \varepsilon \right\}, \quad i = 1, ..., N,
\]

\[
U^\varepsilon := U_1^\varepsilon \times ... \times U_N^\varepsilon.
\]

Then, \( U^\varepsilon \) is open with \( U \subset U^\varepsilon \) and, similarly to the proof of Proposition 3.2(b), we can conclude that \( B(U^\varepsilon) \subset \mathbb{B}_K \) if \( \varepsilon \) is sufficiently small. Let now \( u \) be any function in \( U^\varepsilon \) and let \( \tilde{h} \in L^2([0,T]; \mathbb{R}^N) \) be arbitrary with \( u + \tilde{h} \in U^\varepsilon \). By Taylor expansion we obtain the decompositions

\[
f_{u + \tilde{h}} - f_u = f_u'[\tilde{h}] + f_u^{R}[\tilde{h}] \quad \text{and} \quad g_{u + \tilde{h}} - g_u = g_u'[\tilde{h}] + g_u^{R}[\tilde{h}]
\]

(20)

where \( f_u^{R}[\tilde{h}], g_u^{R}[\tilde{h}] \in C([0,T]; L^2) \). It follows from [14, Thm. 3, Lem. 10] and the boundedness of \( B(\cdot) \) that

\[
\|f_u^{R}[\tilde{h}]\|_{C([0,T]; L^2)} \leq C\|B(\tilde{h})\|_{W}^{1+\gamma} \leq C\|\tilde{h}\|_{L^1([0,T]; \mathbb{R}^N)}^{1+\gamma}, \quad \|g_u^{R}[\tilde{h}]\|_{C([0,T]; L^2)} \leq C\|B(\tilde{h})\|_{W}^{1+\gamma} \leq C\|\tilde{h}\|_{L^1([0,T]; \mathbb{R}^N)}^{1+\gamma}
\]

(21)
where $\gamma = 1 - \frac{3}{\beta} > 0$ denotes the constant from Proposition 2.1. Now, by integration by parts,

$$J'(u + \hat{h})[h] - J'(u)[h] = \sum_{i=1}^{N} \left\{ \lambda_i \langle h_i, \hat{h}_i \rangle_{L^2([0,T])} - \int_{[0,T] \times \mathbb{R}^6} (v \times m_i) \cdot (\partial_v f_u g^R_u[h] - \partial_v g_u f^R_u[h]) h_i \, d(t, x, v) \\
\quad \quad \quad - \int_{[0,T] \times \mathbb{R}^6} (v \times m_i) \cdot \left[ \partial_v f_u (g_{u+h} - g_u) - \partial_v g_u (f_{u+h} - f_u) \right] h_i \, d(t, x, v) \right\}.$$

Inserting (20) then yields

$$J'(u + \hat{h})[h] - J'(u)[h] = \sum_{i=1}^{N} \left\{ \lambda_i \langle h_i, \hat{h}_i \rangle_{L^2([0,T])} - \int_{[0,T] \times \mathbb{R}^6} (v \times m_i) \cdot (\partial_v f_u g^R_u[h] - \partial_v g_u f^R_u[h]) h_i \, d(t, x, v) \right\} + \mathcal{R}(h, \hat{h}) \quad (22)$$

where

$$\mathcal{R}(h, \hat{h}) := -\sum_{i=1}^{N} \int_{[0,T] \times \mathbb{R}^6} (v \times m_i) \cdot (\partial_v f_u g^R_u[h] - \partial_v g_u f^R_u[h]) h_i \, d(t, x, v) \quad (23)$$

$$\quad + \int_{[0,T] \times \mathbb{R}^6} (v \times m_i) \cdot \left[ \partial_v f_u (g_{u+h} - g_u) - \partial_v g_u (f_{u+h} - f_u) \right] h_i \, d(t, x, v).$$

The remainder $\mathcal{R}(h, \hat{h})$ can be bounded by

$$|\mathcal{R}(h, \hat{h})| \leq C \|m_i\|_{L^\infty} \|h\|_{L^2([0,T] ; \mathbb{R}^N)} \left( \|\partial_v f_u\|_{L^\infty} \|g^R_u[h]\|_{C([0,T] ; L^2)} + \|\partial_v g_u\|_{L^\infty} \|f^R_u[h]\|_{C([0,T] ; L^2)} \\
\quad \quad \quad + \|\partial_v f_u (g_{u+h} - g_u) - \partial_v g_u (f_{u+h} - f_u)\|_{L^2} \right).$$

From Proposition 3(a), Proposition 4.2 and (21) we deduce that $\mathcal{R}(\cdot, \hat{h})$ can be bounded in the operator norm by

$$||\mathcal{R}(\cdot, \hat{h})|| = \sup \{ ||\mathcal{R}(h, \hat{h})|| : ||h||_{L^2([0,T] ; \mathbb{R}^N)} \leq 1 \} \leq C ||\hat{h}||^{1+\gamma}_{L^2([0,T] ; \mathbb{R}^N)}.$$

Hence, we can conclude from equation (22) that $J$ is twice Fréchet differentiable on $\mathbb{U}^c$ (and thus especially on $\mathbb{U}$) where the second-order Fréchet derivative is given by

$$J''(u)[h, \hat{h}] = \sum_{i=1}^{N} \lambda_i \langle h_i, \hat{h}_i \rangle_{L^2([0,T])} - \int_{[0,T] \times \mathbb{R}^6} (v \times m_i) \cdot (\partial_v f_u g^R_u[h] - \partial_v g_u f^R_u[h]) h_i \, d(t, x, v) \quad (24)$$

since

$$\|J'(u + \hat{h})[\cdot] - J'(u)[\cdot] - J''(u)[\cdot, \hat{h}]\|_{L^2([0,T] ; \mathbb{R}^N)} \leq \|\mathcal{R}(\cdot, \hat{h})\|_{L^2([0,T] ; \mathbb{R}^N)} \leq C ||\hat{h}||^{1+\gamma}_{L^2([0,T] ; \mathbb{R}^N)} \to 0,$$

as $||\hat{h}||_{L^2([0,T] ; \mathbb{R}^N)} \to 0$ where $\| \cdot \|$ stands for the operator norm. We point out that the operator $\hat{h} \mapsto J''(u)[h, \hat{h}]$ is linear and bounded. For any $u, \tilde{u} \in \mathbb{U}$ and $h, \tilde{h} \in L^2([0,T] ; \mathbb{R}^N),$

$$|J''(u)[h, \hat{h}] - J''(u)[\tilde{u}, \tilde{h}]| \leq C \|h\|_{L^2([0,T] ; \mathbb{R}^N)} \left( \|\partial_v f_u - \partial_v \tilde{f}_u\|_{L^\infty} \|g^R_u[h]\|_{L^2} + \|\partial_v g_u - \partial_v \tilde{g}_u\|_{L^\infty} \|f^R_u[h]\|_{L^2} \\
\quad \quad \quad + \|\partial_v f_u (g_{u+h} - g_u) - \partial_v \tilde{g}_u (f_{u+h} - f_u)\|_{L^2} \right).$$

(23)
Note that for \( \|\hat{h}\|_{L^2([0,T];\mathbb{R}^N)} \leq 1 \) the terms \( \|g_u'[\hat{h}]\|_{L^2}, \|f'_u[\hat{h}]\|_{L^2}, \|\partial_c g_u\|_{L^\infty} \) and \( \|\partial_c f_u\|_{L^\infty} \) are all bounded from above by some constant depending only on \( \hat{f}, f_d, a, b, M \) and \( T \). This is due to \([14, \text{Cor. 2}]\) in combination with Proposition 3.2(b). Also recall that

\[
\begin{align*}
\frac{\partial_c f_u - \partial_c f_u}{\partial_c g_u - \partial_c g_u} & \leq C \|u - \bar{u}\|_{L^2([0,T];\mathbb{R}^N)} \\
\|f'_u[\hat{h}] - f'_u[\hat{h}]\|_{L^2} & \leq C \|u - \bar{u}\|_{L^2([0,T];\mathbb{R}^N)} \\
g'_u[\hat{h}] - g'_u[\hat{h}] & \leq C \|u - \bar{u}\|_{L^2([0,T];\mathbb{R}^N)}
\end{align*}
\]

if \( \|\hat{h}\|_{L^2([0,T];\mathbb{R}^N)} \leq 1 \) according to the Propositions 2.3(a), 4.2, 3.2(d) and 4.9. Hence, if both \( \|h\|_{L^2([0,T];\mathbb{R}^N)} \leq 1 \) and \( \|\hat{h}\|_{L^2([0,T];\mathbb{R}^N)} \leq 1 \), (23) implies that

\[
\|J''(u)[h, \bar{h}] - J''(\bar{u})[h, \bar{h}]\| \leq C \|u - \bar{u}\|_{L^2([0,T];\mathbb{R}^N)}^2.
\]

This proves (19) and completes the proof of Proposition 4.10.

\( \square \)

**Comment** Since \( J \) is twice continuously Fréchet differentiable, Schwarz’s theorem yields that \( J''(u) \) is symmetric, i.e., \( J''(u)[h, \bar{h}] = J''(u)[\bar{h}, h] \) for all \( u \in \mathbb{U} \) and \( h, \bar{h} \in L^2([0,T];\mathbb{R}^N) \).

Now that we know that the cost functional is twice continuously differentiable, we can easily establish a sufficient condition for strict local optimality: Let \( \bar{u} \in \mathbb{U} \) satisfy the variational inequality (NC1) and suppose that \( J''(\bar{u}) \) is positive definite, i.e., there exists \( c > 0 \) such that

\[
J''(\bar{u})[h, \bar{h}] \geq c \|h\|_{L^2([0,T];\mathbb{R}^N)}^2, \quad h \in L^2([0,T];\mathbb{R}^N).
\]

Then \( \bar{u} \) is a strict local minimizer of \( J \) on the set \( \mathbb{U} \). However this condition is far too restrictive and would resemble cracking a nut with a sledgehammer. (The authors’ attention was drawn to this fact by L. Blank [3]). Therefore, we will present a much weaker sufficient condition for strict local optimality where only directions \( h \) from a certain critical cone \( \mathcal{C}_{\bar{u}} \) have to be taken into account. For general semilinear (elliptic or parabolic) control problems this method was introduced by E. Casas, J. C. de los Reyes and F. Tröltzsch [8]. We will proceed similarly and define the cone of critical directions as follows:

**Definition 4.11.** We define the set

\[
\mathcal{C}_{\bar{u}} := \{ h \in L^2([0,T];\mathbb{R}^N) \mid h \text{ satisfies condition (25)} \}
\]

where condition (25) reads as follows:

\[
\begin{align*}
\text{For all } i \in \{1, \ldots, N\} \text{ and almost all } t \in [0,T] : \quad h_i(t) &= \begin{cases} 
\geq 0, & \bar{u}_i(t) = a_i(t) \\
\leq 0, & \bar{u}_i(t) = b_i(t) \\
= 0, & \lambda_i \bar{u}_i(t) - p_i(\bar{u})(t) \neq 0
\end{cases}
\end{align*}
\]

This set \( \mathcal{C}_{\bar{u}} \) is called the ***cone of critical directions***.

Now, we can establish the following sufficient condition for strict local optimality:

**Theorem 4.12.** Suppose that \( \bar{u} \in \mathbb{U} \) and let \( f_{\bar{u}} \) and \( g_{\bar{u}} \) be its induced state and costate. We assume that the variational inequality (NC1) holds for all \( u \in \mathbb{U} \), i.e.,

\[
\int_0^T (\lambda_i \bar{u}_i(t) - p_i(\bar{u})(t)) (u_i(t) - \bar{u}_i(t)) \, dt \geq 0, \quad i = 1, \ldots, N, \quad u \in \mathbb{U},
\]

and that for all critical directions \( h \in \mathcal{C}_{\bar{u}} \setminus \{0\} \),

\[
\sum_{i=1}^N \left( \lambda_i \|h_i\|_{L^2([0,T])}^2 - \int_{[0,T] \times \mathbb{R}^6} (v \times m_i) \cdot (\partial_c f_{\bar{u}} g_{\bar{u}}'[h] - \partial_c g_{\bar{u}} f_{\bar{u}}'[h]) h_i(t, x, v) \, d(t, x, v) \right) = J''(\bar{u})[h, h] > 0.
\]
In this case \( J \) satisfies a quadratic growth condition: There exist \( \delta > 0 \) and \( \varepsilon > 0 \) such that for all \( u \in U \) with \( \|u - \bar{u}\|_{L^2([0,T];\mathbb{R}^N)} < \delta \),

\[
J(u) \geq J(\bar{u}) + \frac{\varepsilon}{2}\|u - \bar{u}\|^2_{L^2([0,T];\mathbb{R}^N)}
\]

and hence \( \bar{u} \) is even a strict local minimizer of \( J \) on the set \( U \).

**Proof** We will proceed analogously to the corresponding proof in [8] and argue by contradiction. Therefore, we assume that \( \bar{u} \) does not satisfy the quadratic growth condition (27). Then, there exists a sequence \( (u_k)_{k \in \mathbb{N}} \subset U \setminus \{ \bar{u} \} \) such that

\[
u_k \to \bar{u} \text{ in } L^2([0,T];\mathbb{R}^N) \quad \text{and} \quad \forall k \in \mathbb{N} : \ J(\bar{u}) + \frac{1}{k}\|u_k - \bar{u}\|^2_{L^2([0,T];\mathbb{R}^N)} > J(u_k)
\]

For \( k \in \mathbb{N} \), we define

\[
d_k := \|u_k - \bar{u}\|_{L^2([0,T];\mathbb{R}^N)} \quad \text{and} \quad h_k := \frac{1}{d_k}(u_k - \bar{u}).
\]

Since \( \|h_k\|_{L^2([0,T];\mathbb{R}^N)} = 1 \), the Banach-Alaoglu theorem implies that there exists some function \( h \in L^2([0,T];\mathbb{R}^N) \) such that \( h_k \rightharpoonup h \) in \( L^2([0,T];\mathbb{R}^N) \) up to a subsequence. We split the argument into several steps:

**Step 1:** We will show that \( J'(\bar{u})[h] = 0 \). The mean value theorem yields

\[
J(u_k) = J(\bar{u}) + d_k J'(v_k)[h_k]
\]

where \( v_k \) is a point in \( L^2([0,T];\mathbb{R}^N) \) between \( \bar{u} \) and \( u_k \) and hence

\[
J'(v_k)[h_k] = \frac{1}{d_k}(J(u_k) - J(\bar{u})) \leq \frac{1}{k d_k}\|u_k - \bar{u}\|^2_{L^2([0,T];\mathbb{R}^N)} = \frac{1}{k}\|u_k - \bar{u}\|_{L^2([0,T];\mathbb{R}^N)}.
\]

Since \( u_k \to u \) in \( L^2([0,T];\mathbb{R}^N) \) we also have \( v_k \to \bar{u} \) in \( L^2([0,T];\mathbb{R}^N) \) as \( v_k \) was chosen between \( \bar{u} \) and \( u_k \). Because of continuity, this implies that \( f_{v_k} \to f_{\bar{u}} \) and \( g_{v_k} \to g_{\bar{u}} \) in \( C([0,T];C^1_b(\mathbb{R}^6)) \). Since for general \( u \in U \) and \( h \in L^2([0,T];\mathbb{R}^N) \) the derivative \( J'(u)[h] \) can be expressed by (10) one can easily conclude that

\[
J'(\bar{u})[h] = \lim_{k \to \infty} J'(v_k)[h_k] \leq \lim_{k \to \infty} \frac{1}{k}\|u_k - \bar{u}\|_{L^2([0,T];\mathbb{R}^N)} = 0.
\]

On the other hand, it follows from condition (NC1) that

\[
J'(\bar{u})[h] = \frac{1}{d_k}\sum_{i=1}^N \int_0^T \left( \lambda_i \bar{u}_i - p_i(\bar{u}) \right) (u_{k,i} - \bar{u}_i) \, dt \geq 0
\]

and \( J'(\bar{u})[h] = 0 \) immediately follows.

**Step 2:** We will show that \( h \in \mathcal{C}_{\bar{u}} \). First note that the set

\[
\mathcal{M}_{\bar{u}} := \left\{ h \in L^2([0,T];\mathbb{R}^N) \mid \begin{array}{l}
\text{For all } i \in \{1,\ldots,N\} \text{ and almost all } t \in [0,T] : \\
\lambda_i h_i(t) \geq 0 \text{ if } \bar{u}_i(t) = a_i(t) \land h_i(t) \leq 0 \text{ if } \bar{u}_i(t) = b_i(t) \\
\end{array} \right\} \subset L^2([0,T];\mathbb{R}^N)
\]

is obviously closed and convex. Thus it is weakly closed. One can easily see that for all \( k \in \mathbb{N} \), \( u_k - \bar{u} \) belongs to \( \mathcal{M}_{\bar{u}} \) and hence the same holds for \( h_k \). Consequently the weak limit \( h \) must belong to \( \mathcal{M}_{\bar{u}} \) as well. Moreover, the result of Step 1 yields

\[
\sum_{i=1}^N \int_0^T \left( \lambda_i \bar{u}_i(t) - p_i(\bar{u})(t) \right) h_i(t) \, dt = J'(\bar{u})[h] = 0.
\]

It is a well-known result that that the variational inequality (NC1) is equivalent to the so-called pointwise variational inequality (see [24, p. 68]). This means, it even holds that \( \lambda_i \bar{u}_i(t) - p_i(\bar{u})(t) \geq 0 \) for almost all \( t \in [0,T] \) and all \( i \in \{1,\ldots,N\} \). Together with (29) we can conclude that

\[
(\lambda_i \bar{u}_i(t) - p_i(\bar{u})(t)) \, h_i(t) = 0, \quad i = 1,\ldots,N
\]
for almost all \( t \in [0, T] \) which directly yields \( h_i(t) = 0 \) for almost all \( t \in [0, T] \) with \( \lambda_i \bar{u}_i(t) - p_i(\bar{u})(t) \neq 0 \). This proves \( h \in C_a \).

**Step 3:** We will now prove that \( h = 0 \). Because of (26), it suffices to show that \( J''(\bar{u})[h, h] \leq 0 \). By a second-order Taylor expansion we obtain that
\[
J(u_k) = J(\bar{u}) + d_k J'(\bar{u})[h_k] + \frac{d_k^2}{2} J''(w_k)[h_k, h_k]
\]
\[
= J(\bar{u}) + d_k J'(\bar{u})[h_k] + \frac{d_k^2}{2} J''(\bar{u})[h_k, h_k] + \frac{d_k^2}{2} \left[ J''(w_k)[h_k, h_k] - J''(\bar{u})[h_k, h_k] \right]
\]
for any \( k \in \mathbb{N} \) where \( w_k \) is a point in \( L^2([0, T]; \mathbb{R}^N) \) between \( u_k \) and \( \bar{u} \). This means that
\[
J''(\bar{u})[h_k, h_k] = \frac{2}{d_k^2} \left[ J(u_k) - J(\bar{u}) \right] - \frac{2}{d_k^2} J'(\bar{u})[h_k] - \frac{d_k^2}{2} \left[ J''(w_k)[h_k, h_k] - J''(\bar{u})[h_k, h_k] \right].
\] (30)

First note that, according to (10) and condition (NC1),
\[
\frac{2}{d_k^2} J'(\bar{u})[h_k] = \frac{2}{d_k^2} J'(\bar{u})[u_k - \bar{u}] = \frac{2}{d_k^2} \sum_{i=1}^N \int_0^T (\lambda_i \bar{u}_i - p_i(\bar{u}))(u_{k,i} - \bar{u}_i) \, dt \geq 0.
\] (31)

From (28) it follows that
\[
\frac{2}{d_k^2} \left[ J(u_k) - J(\bar{u}) \right] < \frac{2}{kd_k^2} \| u_k - \bar{u} \|^2_{L^2([0,T];\mathbb{R}^N)} = \frac{2}{k}.
\] (32)

Since \( \| h_k \|_{L^2([0,T];\mathbb{R}^N)} = 1 \), Proposition 4.10 (b) yields
\[
\left| J''(w_k)[h_k, h_k] - J''(\bar{u})[h_k, h_k] \right| \leq \| w_k - \bar{u} \|^2_{L^2([0,T];\mathbb{R}^N)} \leq d_k^2.
\] (33)

Now, we can use (31)-(33) to bound the right-hand side of (30) from above. We obtain
\[
J''(\bar{u})[h_k, h_k] < \frac{2}{k} + d_k^2 \to 0, \ k \to \infty, \quad \text{i.e.,} \quad \limsup_{k \to \infty} J''(\bar{u})[h_k, h_k] \leq 0.
\] (34)

From \( h_k \to h \) in \( L^2([0,T];\mathbb{R}^N) \) we can also deduce that every component \( h_{k,i} \) converges to \( h_i \) weakly in \( L^2([0,T]) \). As the norm is weakly lower semicontinuous, it follows that
\[
\| h_i \|^2_{L^2([0,T])} \leq \liminf_{k \to \infty} \| h_{k,i} \|^2_{L^2([0,T])}, \quad i = 1, ..., N.
\] (35)

For any \( i \in \{1, ..., N\} \), we will now consider the bilinear functional
\[
B_i(h, \tilde{h}) := \int_{[0,T] \times \mathbb{R}^d} (v \times m_i) \cdot \left( \partial_x f_a g_a[h] - \partial_x g_a f_a[h] \right) \tilde{h}_i \, dt \, dx, \quad h, \tilde{h} \in L^2([0,T];\mathbb{R}^N).
\]

Recall that the Fréchet derivatives \( f_a'[\cdot] \) and \( g_a'[\cdot] \) are linear and bounded operators. Thus, by the definition of weak convergence, \( f_a'[h_k] \to f_a'[h] \) and \( g_a'[h_k] \to g_a'[h] \) as \( k \to \infty \). Using Hölder’s inequality we obtain that \( B_i(h_k, h_k) - B_i(h_k) \to 0, \ k \to \infty \). Moreover, \( B_i(h_k, h_k) - B_i(h, h) \to 0 \) for \( k \to \infty \) follows directly from the weak convergence of \( (h_k) \) in \( L^2([0,T];\mathbb{R}^N) \). In summary, we have
\[
B_i(h_k, h_k) \to B_i(h, h), \quad k \to \infty.
\]

Consequently, using (34) we obtain that
\[
J''(\bar{u})[h, h] = \sum_{i=1}^N \left\{ \lambda_i \| h_i \|^2_{L^2([0,T])} - B_i(h, h) \right\} \leq \sum_{i=1}^N \left\{ \lambda_i \liminf_{k \to \infty} \| h_{k,i} \|^2_{L^2([0,T])} + \liminf_{k \to \infty} \left[ -B_i(h_k, h_k) \right] \right\}
\]
\[ \leq \limsup_{k \to \infty} J''(\bar{u})[h_k, h_k] \leq 0. \]

From (26) and the fact that \( h \in C \), we can finally conclude that \( h \) must be the null function.

**Step 4:** We will now show that \( \|h_k\|_{L^2([0,T];\mathbb{R}^N)} \) converges to \( h = 0 \) strongly in \( L^2([0,T];\mathbb{R}^N) \). The bound \( J''(\bar{u})[h_k, h_k] \leq 2/k + d_k^2 \) from (34) particularly implies that
\[
\sum_{i=1}^{N} \lambda_i \|h_k, i\|_{L^2([0,T])} \leq \frac{2}{k} + d_k^2 + \sum_{i=1}^{N} B_i(h_k, h_k) \to 0, \quad k \to \infty.
\]

Thus, \( \|h_k, i\|_{L^2([0,T])} \to 0 \) for all \( i \in \{1, \ldots, N\} \) which immediately yields \( \|h_k\|_{L^2([0,T];\mathbb{R}^N)} \to 0 \) as \( k \to \infty \).

However, this is a contradiction, because \( h_k \) was constructed in such a way that \( \|h_k\|_{L^2([0,T];\mathbb{R}^N)} = 1 \) for all \( k \in \mathbb{N} \). This means that our initial assumption was wrong and hence \( \bar{u} \) does satisfy the quadratic growth condition (27). The proof is complete. \( \square \)

**References**

1. A. Ahmad Ali, K. Deckelnick, and M. Hinze. Global minima for semilinear optimal control problems. *Comput. Optim. Appl.*, 65(1):261–288, 2016.
2. J. Batt. Global symmetric solutions of the initial value problem of stellar dynamics. *J. Differential Equations*, 25(3):342–364, 1977.
3. L. Blank. Personal communication, December 22, 2017 in Regensburg, Germany.
4. M.E. Bradley and S.M. Lenhart. Bilinear optimal control of a Kirchhoff plate via internal controllers. In *Modelling and optimization of distributed parameter systems (Warsaw, 1995)*, pages 233–240. Chapman & Hall, New York, 1996.
5. S. Caprino, G. Cavallaro, and C. Marchioro. Time evolution of a Vlasov-Poisson plasma with magnetic confinement. *Kinet. Relat. Models*, 5(4):729–742, 2012.
6. S. Caprino, G. Cavallaro, and C. Marchioro. On a magnetically confined plasma with infinite charge. *SIAM J. Math. Anal.*, 46(1):133–164, 2014.
7. S. Caprino, G. Cavallaro, and C. Marchioro. A Vlasov-Poisson plasma with unbounded mass and velocities confined in a cylinder by a magnetic mirror. *Kinet. Relat. Models*, 9(4):657–686, 2016.
8. E. Casas, J.C. de los Reyes, and F. Tröltzsch. Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints. *SIAM J. Optim.*, 19(2):616–643, 2008.
9. L.C. Evans. *Partial differential equations*, volume 19 of * Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.
10. K.R. Fister and J.C. Panetta. Optimal control applied to cell-cycle-specific cancer chemotherapy. *SIAM J. Appl. Math.*, 60(3):1059–1072, 2000.
11. O. Glass. On the controllability of the Vlasov-Poisson system. *J. Differential Equations*, 195(2):332–379, 2003.
12. O. Glass and D. Han-Kwan. On the controllability of the Vlasov-Poisson system in the presence of external force fields. *J. Differential Equations*, 252(10):5455–5491, 2012.
13. O. Glass and D.-H. Han-Kwan. On the controllability of the relativistic Vlasov-Maxwell system. *J. Math. Pures Appl.*, 103(3):695–740, 2015.
14. P. Knopf. Optimal control of a Vlasov-Poisson plasma by an external magnetic field. *Calc. Var.*, 57(134), 2018.
15. R. Kurth. Das Anfangswertproblem der Stellardynamik. *Z. Astrophys.*, 30:213–229, 1952.
16. E.H. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.
17. P.-L. Lions and B. Perthame. Propagation of moments and regularity for the 3-dimensional Vlasov-Poisson system. *Invent. Math.*, 105(2):415–430, 1991.
18. T.T. Nguyen, T.V. Nguyen, and W.A. Strauss. Erratum to: Global magnetic confinement for the 1.5D Vlasov-Maxwell system [MR3294216]. *Kinet. Relat. Models*, 8(3):615–616, 2015.
19. T.T. Nguyen, T.V. Nguyen, and W.A. Strauss. Global magnetic confinement for the 1.5D Vlasov-Maxwell system. *Kinet. Relat. Models*, 8(1):153–168, 2015.
20. K. Pfaffelmoser. Global classical solutions of the Vlasov-Poisson system in three dimensions for general initial data. *J. Differential Equations*, 95(2):281–303, 1992.
21. G. Rein. Collisionless kinetic equations from astrophysics—the Vlasov-Poisson system. In *Handbook of differential equations: evolutionary equations. Vol. III*, Handb. Differ. Equ., pages 383–476. Elsevier/North-Holland, Amsterdam, 2007.
22. J. Schaeffer. Global existence of smooth solutions to the Vlasov-Poisson system in three dimensions. *Comm. Partial Differential Equations*, 16(8-9):1313–1335, 1991.
[23] A. L. Skubachevskii. Vlasov-Poisson equations for a two-component plasma in a homogeneous magnetic field. *Uspekhi Mat. Nauk*, 69(2(416)):107–148, 2014.

[24] F. Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.