

OPTIMAL CONTROL OF THE INHOMOGENEOUS RELATIVISTIC MAXWELL NEWTON LORENTZ EQUATIONS

C. MEYER, S. M. SCHNEPP, AND O. THOMA

ABSTRACT. This note is concerned with an optimal control problem governed by the relativistic Maxwell-Newton-Lorentz equations, which describes the motion of charges particles in electro-magnetic fields and consists of a hyperbolic PDE system coupled with a nonlinear ODE. An external magnetic field acts as control variable. Additional control constraints are incorporated by introducing a scalar magnetic potential which leads to an additional state equation in form of a very weak elliptic PDE. Existence and uniqueness for the state equation is shown and the existence of a global optimal control is established. Moreover, first-order necessary optimality conditions in form of Karush-Kuhn-Tucker conditions are derived. A numerical test illustrates the theoretical findings.

Key words. Optimal control, Maxwell's equation, Abraham model, Dirichlet control, state constraints.

AMS subject classification. 49J20, 49J15, 49K20, 49K15, 35Q61

1 Introduction

In this paper we discuss an optimal control problem governed by the relativistic Maxwell-Newton-Lorentz equations. This system of equations consists of Maxwell's equations, i.e., a hyperbolic PDE system, and a nonlinear ODE. It models the relativistic motion of charged particles in electromagnetic fields and is therefore used for the simulation of particle accelerators [1, 18, 21, 30]. The control variable is an additional exterior magnetic field, which, in practice, could be realized by exterior (dipole, quadrupole etc.) magnets surrounding the accelerator tube [39, 45]. The aim of the optimization is to steer the particle beam to a given desired track and/or end-time position. Beside the Maxwell-Newton-Lorentz system, the optimization problem is subject to several additional constraints. First, the particle beam should stay inside the accelerator tube, which is realized by pointwise constraints on the particle position and constitutes a pointwise state constraint from a mathematical point of view. Moreover, as a stationary magnetic field, the control has to satisfy certain constraints, e.g. its divergence has to vanish. In order to guarantee these constraints, we introduce a scalar magnetic potential, whose boundary data serve as new control variable. This gives rise to a Poisson equation for the exterior magnetic field entering the system of state equations. Physically, the new control variable can be interpreted as a surface current on the boundary of the computational domain. In this way we obtain a Dirichlet boundary control problem.

Let us put our work into perspective. Optimal control of Maxwell's equations and coupled systems involving these have been subject to intensive research in the recent past. We only mention the work of Tröltzsch et al. [16, 33–35, 44] and Yousept [46–50]. However, most of these contributions deal with stationary or time harmonic Maxwell's equation. In [35] the so-called evolution Maxwell equation in form of a (degenerate) parabolic PDE is considered. In contrast to this, we deal with a

first-order hyperbolic system for the electric and the magnetic fields. Optimal control of magneto-hydrodynamic processes was investigated in [22]. These processes are modeled by a coupled system consisting of Maxwell's equation and the Navier-Stokes equations. However, [22] also focuses on the stationary case. Up to our best knowledge, the non-standard coupling of the (hyperbolic) Maxwell's equation and the ODE for the relativistic motion of charged particles have not been treated so far in the context of optimal control, neither from an analytical nor from a numerical point of view. The mathematical treatment of the Maxwell-Newton-Lorentz system itself however has been investigated by several authors before. Concerning the analysis we mention [3, 17, 26, 42] and the references therein. Regarding its numerical treatment we refer to [18, 21, 30]. The analytical and numerical investigations presented in this paper will partly rely on these findings. As mentioned before the control constraints on the external magnetic field are realized by introducing a scalar potential which leads to a boundary control problem of Dirichlet type. Optimal control problems of this type have been intensely investigated in the recent past, see e.g. [11, 15, 29, 31, 37]. We choose $L^2(\Gamma)$ as control space, so that the associated Poisson equation is treated in very weak form, which is a well-established procedure, cf. e.g. [31]. Another challenging aspect of the optimal control under consideration are the pointwise state constraints on the particle position. Lagrange-multipliers associated with constraints of this type, in general, lack in regularity and are only measures, see e.g. [9, 10] for the case of PDEs and [23] and the references therein for the case of ODEs. Numerically, such constraints are frequently treated by regularization and relaxation methods, especially in the PDE case, cf. e.g. [24, 32, 41]. We also follow this approach and apply an interior point method to realize the state constraints.

The paper is organized as follows: in the following section we introduce the physical model, i.e., the Maxwell-Newton-Lorentz system. This model is not directly amenable for a mathematically rigorous treatment mainly due to two reasons, which are addressed at the end of Section 2. We therefore slightly modify the model in Section 3 by replacing the point charge with a distributed volume charge density. In addition the scalar magnetic potential is introduced in this section which allows us to formulate the optimal control problem, first in a formal way. After stating our standing assumptions in Section 3.1, Section 3.2 is then devoted to a mathematically sound and rigorous statement of the optimal control problem, including the function spaces for all optimization variables as well as the notion of solutions for the differential equations involved in the state system. We start the analysis of the optimal control problem by discussing the state equation in Section 4. Then we turn to the optimal control problem and show the existence of globally optimal controls in Section 5. The analytical part of the paper ends with the derivation of first-order-necessary optimality conditions involving Lagrange multipliers in Section 6. The final Section 7 is devoted to the numerical treatment of the optimal control problem. After describing the discretization of the state system and the optimization algorithm, we present an exemplary numerical result for the end time tracking of a single-particle beam.

2 Statement of the physical model

In this section we introduce the physical model underlying the optimal control problem. The precise mathematical model will be stated in Section 3.2.

To keep the discussion concise we will restrict to the motion of only one particle in the accelerator. The adaptation of the model to a finite number of particles is

straightforward, see Remark 2.2 below. Our model is based on the classical inhomogeneous Maxwell's equations with the boundary conditions of a perfect conductor. In strong form these equations read:

$$\epsilon \frac{\partial}{\partial t} E(x, t) - \mu^{-1} \operatorname{curl} B(x, t) = j(x, t) \quad \text{in } \Omega \times [0, T] \quad (2.1a)$$

$$\frac{\partial}{\partial t} B(x, t) + \operatorname{curl} E(x, t) = 0 \quad \text{in } \Omega \times [0, T] \quad (2.1b)$$

$$\operatorname{div} E(x, t) = \frac{1}{\epsilon} \rho(x, t), \quad \operatorname{div} B(x, t) = 0 \quad \text{in } \Omega \times [0, T] \quad (2.1c)$$

$$E(x, 0) = E_0(x), \quad B(x, 0) = B_0(x) \quad \text{in } \Omega \quad (2.1d)$$

$$E \times n = 0, \quad B \cdot n = 0 \quad \text{on } \Gamma \times [0, T]. \quad (2.1e)$$

Herein, E and B denote the electric and magnetic field, respectively, and Ω is the domain occupied by the interior of the accelerator channel. Its boundary $\partial\Omega$ is denoted by Γ , and n is the outward unit normal on Γ . Moreover, ϵ is the permittivity of free space, while μ denotes the permeability, which are assumed to be constant in Ω . Finally, ρ and j denote the charge density and the electric current.

Remark 2.1. *Provided the conservation of charge holds, the two Gauss laws in (2.1c) as well as the boundary condition on B intrinsically follow from Faraday's and Ampere's laws in (2.1a) and (2.1b) so that (2.1) is not overdetermined.*

In our case, the charge density is generated by a single point charge and therefore given by

$$\rho(x, t) := q\delta(|x - r(t)|_2) \quad \text{in } \Omega \times [0, T], \quad (2.2)$$

where $q > 0$ is the constant particle charge, r denotes the particle position, and $|\cdot|_2$ is the Euclidean norm of a vector. Furthermore, $\delta : \mathbb{R} \rightarrow \{0, \infty\}$ is the Dirac delta distribution. The current $j(x, t)$ arising on the right hand side in (2.1a) is generated by the motion of the particle and thus given by

$$j(x, t) := -q\delta(|x - r(t)|_2)v(p(t)) \quad \text{in } \Omega \times [0, T], \quad (2.3)$$

where p denotes the relativistic momentum of the particle. Moreover, we set

$$v(p(t)) := (m_0^q \gamma(p(t)))^{-1} p(t) \quad (2.4)$$

with the mass at rest m_0^q and the Lorentz factor

$$\gamma(p(t)) := \sqrt{1 + \frac{\|p(t)\|^2}{(m_0^q c)^2}},$$

where $c > 0$ denotes the speed of light. Note that $v(p)$ is nothing else than the velocity of the particle. It is easily verified that ρ and j chosen in this way satisfy the conservation of charge.

We summarize the constants of the model in Table 2.1.

Physical constants	Name of quantity
c	speed of light
ϵ	permittivity
μ	permeability
m_0^q	rest mass
q	particle charge

TABLE 2.1. Overview of arising constants

In addition to (2.4) we introduce the abbreviation

$$\beta(p(t)) := c^{-1} v(p(t)), \quad (2.5)$$

which prove helpful in the sequel.

The motion of the particle in electromagnetic fields is governed by the relativistic Newton-Lorentz equations given by the formulae

$$\dot{p}(t) = q \left[e(r(t)) + E(r(t), t) + \beta(p(t)) \times (b(r(t)) + B(r(t), t)) \right] \quad \text{in } [0, T] \quad (2.6a)$$

$$\dot{r}(t) = v(p(t)) \quad \text{in } [0, T] \quad (2.6b)$$

$$p(0) = p_0 \quad \text{and} \quad r(0) = r_0 \quad (2.6c)$$

with initial particle position and momentum $p_0, r_0 \in \mathbb{R}^3$. Furthermore, e and b denote the external electric and magnetic fields, respectively. These fields are generated by exterior capacitors and magnets in order to steer the particle beam. They are assumed to fulfill the homogeneous Maxwell's equations in Ω . As we only consider magnets for manipulating the beam, we assume e to equal zero. Therefore, the external magnetic field b has to satisfy the conditions

$$\operatorname{div} b = 0, \operatorname{curl} b = 0 \quad \text{and} \quad \partial_t b = 0 \quad \text{in } \Omega. \quad (2.7)$$

This external magnetic field b will serve as control in the following.

To summarize the overall model reads as follows:

$$\epsilon \frac{\partial}{\partial t} E(x, t) - \mu^{-1} \operatorname{curl} B(x, t) = -q \delta(|x - r(t)|_2) v(p(t)) \quad \text{in } \Omega \times [0, T] \quad (2.8a)$$

$$\frac{\partial}{\partial t} B(x, t) + \operatorname{curl} E(x, t) = 0 \quad \text{in } \Omega \times [0, T] \quad (2.8b)$$

$$\operatorname{div} E(x, t) = \frac{1}{\epsilon} q \delta(|x - r(t)|_2), \quad \operatorname{div} B(x, t) = 0 \quad \text{in } \Omega \times [0, T] \quad (2.8c)$$

$$\dot{p}(t) = q \left(E(r(t), t) + \beta(p(t)) \times (b(r(t)) + B(r(t), t)) \right) \quad \text{in } [0, T] \quad (2.8d)$$

$$\dot{r}(t) = v(p(t)) \quad \text{in } [0, T] \quad (2.8e)$$

$$E(x, 0) = E_0(x), \quad B(x, 0) = B_0(x), \quad r(0) = r_0, \quad p(0) = p_0, \quad \text{in } \Omega \quad (2.8f)$$

$$E \times n = 0, \quad B \cdot n = 0 \quad \text{on } \Gamma \times [0, T]. \quad (2.8g)$$

Remark 2.2. *In case of an entire bunch of n particles the electric current is given by $-\sum_{i=1}^n q_i \delta(|x - r_i(t)|_2) v(p_i(t))$, while the charge density becomes $\sum_{i=1}^n q_i \delta(|x - r_i(t)|_2)$. The rest of the system remains unchanged, except that we had n equations of the form (2.8d), (2.8e) for each of the n particles, cf. e.g. [42, Section 11]. It is therefore straightforward to adapt the analysis presented in the following to the situation of n particles.*

The model equations in (2.8) feature two critical aspects. First, the particle must not leave the computational domain Ω , i.e. the interior of the accelerator, since otherwise the right hand side in (2.8d) is not well defined. This issue will be resolved by adding an additional state constraints to the optimal control problem. From an application driven point of view this constraint is meaningful, too. Secondly, the pointwise evaluation of the electric and the magnetic fields precisely at the point $x = r(t)$ in (2.8d) is, in general, not well defined, since solutions of Maxwell's equations with j given by (2.3) are singular at this point. We will overcome this difficulty by introducing the so-called Abraham model, which is addressed in the next section. For further details on the Abraham model, we refer to [42, Section 2.4].

3 The optimal control problem

This section is devoted to the optimal control problem. Having established the Abraham model, we introduce a scalar potential to cope with the additional conditions in the external magnetic field in (2.7). Then, we state the complete optimal control problem including the objective functional and the additional state constraints on the particle position. The rest of this section is concerned with the standing assumptions and the mathematically rigorous statement of the optimal control problem.

As described above, the pointwise evaluation in (2.8d) is, in general, not well defined. To resolve this issue, we replace the Dirac delta distribution by a smeared out version. For this purpose we fix a function $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that

$$\begin{aligned} \varphi \in C^{2,1}(\mathbb{R}^3), \quad \text{supp}(\varphi) \subseteq B_R(0), \quad \varphi(x) \geq 0 \quad \forall x \in \mathbb{R}^3 \\ \int_{\mathbb{R}^3} \varphi(x) dx = 1, \quad \varphi(x) = \varphi(y) \quad \text{if } |x|_2 = |y|_2 \end{aligned} \quad (3.1)$$

(i.e., φ is rotationally symmetric). The pointwise evaluations in (2.8d) are then approximated by

$$\begin{aligned} E(r(t), t) + \beta(p(t)) \times (b(r(t)) + B(r(t), t)) \\ \approx \int_{\Omega} \varphi(x - r(t)) \left[E(x, t) + \beta(p(t)) \times (b(x) + B(x, t)) \right] dx. \end{aligned} \quad (3.2)$$

Accordingly, the charge distribution and the current density are replaced by

$$\rho(x, t) = q \varphi(x - r(t)) \quad \text{and} \quad j(x, t) = -q \varphi(x - r(t)) v(p(t)). \quad (3.3)$$

One readily verifies that the conservation of charge is also fulfilled by this choice for ρ and j .

To incorporate the conditions on the external magnetic field in (2.7), we introduce a scalar magnetic potential as solution of the following Poisson's equation with Dirichlet boundary data

$$-\Delta \eta = 0 \quad \text{in } \Omega, \quad \eta = u \quad \text{on } \Gamma. \quad (3.4)$$

Under the assumption that Ω is a simply connected domain, the gradient $b := \nabla \eta$ is a conservative vector field so that

$$\text{div } b = \text{div } (\nabla \eta) = \Delta \eta = 0, \quad \text{curl } b = \text{curl } (\nabla \eta) = 0, \quad \partial_t b = 0,$$

i.e. (2.7), is fulfilled almost everywhere. The Dirichlet data u in (3.4) will serve as the new control variable in the following. Employing (3.4) and integration by parts, one rewrites the integral involving b in (3.2) by

$$\begin{aligned} \int_{\Omega} \varphi(x - r(t)) \beta(p(t)) \times b(x) dx &= -q \int_{\Omega} \eta \nabla \varphi(x - r(t)) \times \beta(p(t)) dx \\ &\quad + q \int_{\Gamma} u \varphi(x - r(t)) \beta(p(t)) \times n ds. \end{aligned} \quad (3.5)$$

Summing up all components of the physical model, the optimal control problem under consideration reads

$$\text{minimize } \mathcal{J}(r, u) := \int_0^T J_1(r(t)) dt + J_2(r(T)) + \frac{\alpha}{2} \int_{\Gamma} u^2 d\zeta \quad (\tilde{\text{P}})$$

subject to Maxwell's equations

$$\epsilon \frac{\partial}{\partial t} E(x, t) - \mu^{-1} \operatorname{curl} B(x, t) = -q\varphi(x - r(t))v(p(t)) \quad \text{in } \Omega \times [0, T] \quad (3.6a)$$

$$\frac{\partial}{\partial t} B(x, t) + \operatorname{curl} E(x, t) = 0 \quad \text{in } \Omega \times [0, T] \quad (3.6b)$$

$$\operatorname{div} E(x, t) = \frac{1}{\epsilon} q\varphi(x - r(t)), \quad \operatorname{div} B(x, t) = 0 \quad \text{in } \Omega \times [0, T] \quad (3.6c)$$

$$E(x, 0) = E_0(x), \quad B(x, 0) = B_0(x) \quad \text{in } \Omega \quad (3.6d)$$

$$E \times n = 0, \quad B \cdot n = 0 \quad \text{on } \Gamma \times [0, T], \quad (3.6e)$$

the relativistic Newton-Lorentz equations

$$\begin{aligned} \dot{p}(t) = & q \int_{\Omega} \varphi(x - r(t)) \left[E(x, t) + \beta(p(t)) \times B(x, t) \right] dx \\ & - q \int_{\Omega} \eta \nabla \varphi(x - r(t)) \times \beta(p(t)) dx \end{aligned} \quad (3.7a)$$

$$\begin{aligned} & + q \int_{\Gamma} u \varphi(x - r(t)) \beta(p(t)) \times n ds \quad \text{in } [0, T] \\ \dot{r}(t) = & v(p(t)) \quad \text{in } [0, T] \end{aligned} \quad (3.7b)$$

$$r(0) = r_0, \quad p(0) = p_0, \quad (3.7c)$$

Poisson's equation

$$-\Delta \eta = 0 \quad \text{in } \Omega, \quad \eta = u \quad \text{on } \Gamma, \quad (3.8)$$

and pointwise state constraints on the particle position

$$r(t) \in \tilde{\Omega}. \quad (3.9)$$

Herein, $J_1, J_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$ are given functions which reflect the goal of the optimization to steer the beam on the overall time interval and at end time, respectively. Moreover, the Tikhonov parameter α is a positive real number. Finally, $\tilde{\Omega} \subset \Omega$ is a closed subdomain fulfilling

$$\operatorname{dist}(\tilde{\Omega}, \Gamma) > R,$$

where R is the number defining the support of the smeared out delta distribution, cf. (3.1).

Remark 3.1. *Note that now the integrands on the right-hand side of (3.7a) are well-defined in any case, even if $r(t) \notin \Omega$ for some $t \in [0, T]$. However, in this case, the model becomes physically meaningless. In this way the state constraint in (3.9) ensures that the model does not lose its physical validity. Moreover, in applications, it is important to keep the particles inside the accelerator tube, which is also reflected by the condition (3.9).*

3.1. Standing assumptions and notation. We start by introducing several function spaces which will be useful in the sequel.

Definition 3.2 ($H(\operatorname{curl}; \Omega)$ -spaces). *By X we denote the space $X = L^2(\Omega; \mathbb{R}^3)$. For convenience of notation the scalar products and corresponding norms in X and $X \times X$ are both denoted by $(\cdot, \cdot)_X$ and $\|\cdot\|_X$, respectively. Moreover, we set*

$$H(\operatorname{curl}; \Omega) := \{\omega \in X : \operatorname{curl} \omega \in X\},$$

where $\operatorname{curl} : X \rightarrow \mathcal{D}'$ denotes the distributional curl-operator. With the obvious scalar product $H(\operatorname{curl}; \Omega)$ becomes a Hilbert space. It is well known that there exists a linear and continuous operator $\tau_n : H(\operatorname{curl}; \Omega) \rightarrow H^{-1/2}(\Gamma; \mathbb{R}^3)$ such that $\tau_n \omega = \omega \times n$ for all $\omega \in H(\operatorname{curl}; \Omega) \cap C(\bar{\Omega}; \mathbb{R}^3)$, see e.g. [20, Chapter 2]. In the sequel we will denote $\tau_n \omega$ by $\omega \times n$ for all $\omega \in H(\operatorname{curl}; \Omega)$ for simplicity and call this operator

tangential trace. For a detailed discussion of the tangential trace we refer to [2]. Furthermore, we define the set

$$H_{\text{curl}}^\Gamma := \{V = (V_1, V_2) \in H(\text{curl}; \Omega) \times H(\text{curl}; \Omega) : V_1 \times n = 0\}.$$

As a closed subspace of a Hilbert space, H_{curl}^Γ is a Hilbert space itself.

Definition 3.3 ($H(\text{div}; \Omega)$ -spaces). We define the set

$$H(\text{div}; \Omega) := \{\omega \in X : \text{div } \omega \in L^2(\Omega)\},$$

where $\text{div} : X \rightarrow \mathcal{D}'$ denotes the distributional divergence. Equipped with the obvious scalar product, $H(\text{div}; \Omega)$ becomes a Hilbert space. Functions in $H(\text{div}; \Omega)$ admit a normal trace, i.e., there is a linear and continuous operator $\gamma_n : H(\text{div}; \Omega) \rightarrow H^{-1/2}(\Gamma)$ such that $\gamma_n \omega = \omega \cdot n$ for all $\omega \in H(\text{div}; \Omega) \cap C(\bar{\Omega}; \mathbb{R}^3)$, see e.g. [43, Theorem 1.2]. As above, we denote the normal trace by $\omega \cdot n$ for all ω in $H(\text{div}; \Omega)$. Furthermore, we define the set

$$\mathcal{H} := \{v \in H_0^1(\Omega) : \nabla v \in H(\text{div}; \Omega), \partial_n v \in L^2(\Gamma)\},$$

where we set $\partial_n v := n \cdot \nabla v$. Endowed with the norm

$$\|v\|_{\mathcal{H}} = (\|v\|_{H^1(\Omega)}^2 + \|\Delta v\|_{L^2(\Omega)}^2 + \|\partial_n v\|_{L^2(\Gamma)}^2)^{\frac{1}{2}}$$

and the corresponding scalar product, it is a Hilbert space, too. Here and in the following, $\Delta := \text{div } \nabla : \mathcal{H} \rightarrow L^2(\Omega)$ denotes the Laplacian.

Now we are in the position to state the assumptions on the domain Ω .

Assumption 3.4 (Regularity of the domain).

- (1) The domain $\Omega \subset \mathbb{R}^3$ is open, bounded, and simply connected.
- (2) The subdomain $\tilde{\Omega}$ can be represented by

$$\tilde{\Omega} = \{x \in \mathbb{R}^3 : g_i(x) \leq 0, i = 1, \dots, m\}$$

where $m \in \mathbb{N}$ and $g_i \in C^1(\mathbb{R}^3, \mathbb{R})$ with absolutely continuous derivatives g'_i .

- (3) Furthermore, Ω is such that for all $g \in L^2(\Omega)$ there exists a unique solution $w \in \mathcal{H}$ of

$$\int_{\Omega} \nabla w \cdot \nabla v \, dx = \int_{\Omega} g v \, dx \quad \forall v \in H_0^1(\Omega) \quad (3.10)$$

and the following a priori estimate

$$\|w\|_{\mathcal{H}} \leq C \|g\|_{L^2(\Omega)}$$

is fulfilled with a constant $C > 0$ independent of g and w .

Remark 3.5. By the Lax-Milgram Lemma (3.10) admits a unique solution in $w \in H_0^1(\Omega)$ and, due to $g \in L^2(\Omega)$, it immediately follows that $\nabla w \in H(\text{div}; \Omega)$. The additional condition $\partial_n w \in L^2(\Gamma)$ is satisfied under rather mild assumptions on the boundary of Ω , cf. [12, Chapter 6].

Assumption 3.6 (Problem data). We assume the following assumptions on the data in (P):

- $r_0 \in \tilde{\Omega}$.
- The first two contributions to the objective fulfill $J_1, J_2 \in C^1(\mathbb{R}^3)$. Furthermore, we assume that J_1 and J_2 are bounded from below by constants $\underline{c}_1 > -\infty$ and $\underline{c}_2 > -\infty$.
- The Tikhonov regularization parameter satisfies $\alpha \in \mathbb{R}, \alpha > 0$.
- The smeared out delta distribution φ fulfills the assumptions in (3.1).
- ϵ, μ, q are positive constants.

- $E_0, B_0 \in X$.
- $g_1, \dots, g_m \in C^1(\mathbb{R}^3)$.

Given a linear normed space \mathcal{X} we denote by $C_{\{0\}}([0, T]; \mathcal{X})$ the space of functions from $C([0, T]; \mathcal{X})$ which vanish at $t = 0$. The space $C_{\{0\}}^1([0, T]; \mathcal{X})$ is defined analogously. By

$$Y := \{(r, p) \in H^1([0, T]; \mathbb{R}^3)^2 : r(0) = p(0) = 0\}, \quad Z := L^2([0, T]; \mathbb{R}^3)^2$$

we denote the state space, which comes into play in Section 6. To keep the notation concise, we also denote the space $\{r \in H^1([0, T]; \mathbb{R}^3) : r(0) = 0\}$ by Y . In addition, the Jacobian of the electric current j as given in (3.3) is denoted by

$$j'(r, p) := (\partial_r j(r, p), \partial_p j(r, p)) = \begin{pmatrix} \partial_r j_1(r, p) & \partial_p j_1(r, p) \\ \partial_r j_2(r, p) & \partial_p j_2(r, p) \end{pmatrix} \quad (3.11)$$

If \mathcal{X} and \mathcal{Y} are linear normed spaces, we write $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ for the space of linear and bounded operators from \mathcal{X} to \mathcal{Y} . Furthermore, $|v|_2$ is the Euclidean norm of a vector $v \in \mathbb{R}^3$. Abusing the notation slightly, we denote the Euclidean norm on $\mathbb{R}^3 \times \mathbb{R}^3$ by the same symbol, i.e., $|(v, w)|_2 := \sqrt{|v|_2^2 + |w|_2^2}$ for $v, w \in \mathbb{R}^3$. If $A \in \mathbb{R}^{3 \times 3}$, then $|A|_F$ denotes the Frobenius norm of A . Finally, throughout the paper, C is a generic constant.

3.2. Mathematically rigorous formulation of the optimal control problem. In the following we define a rigorous notion of solutions to the system of state equations in (3.6)–(3.8). We start with Maxwell's equation and define the linear and unbounded operator

$$\mathcal{A} : X \times X \rightarrow X \times X, \quad \mathcal{A} := \begin{pmatrix} 0 & -\text{curl} \\ \text{curl} & 0 \end{pmatrix}$$

with its domain of definition $D(\mathcal{A}) = H_{\text{curl}}^\Gamma$. In view of Remark 2.1, Maxwell's equation can then be reformulated by the following *Cauchy-Problem*:

$$\left. \begin{aligned} \frac{\partial}{\partial t} \begin{pmatrix} E(t) \\ B(t) \end{pmatrix} + \mathcal{A} \begin{pmatrix} E(t) \\ B(t) \end{pmatrix} &= j \quad \text{a.e. in } [0, T] \\ \begin{pmatrix} E(0) \\ B(0) \end{pmatrix} &= \begin{pmatrix} E_0 \\ B_0 \end{pmatrix} \end{aligned} \right\} \quad (3.12)$$

As shown in [13, Chapter XVII.B., Section 4] and [14, Chapter IX, Section 3], $-i\mathcal{A}$ is self-adjoint, i.e., $-i\mathcal{A} = i\mathcal{A}^* = -(i\mathcal{A})^*$, and consequently the theorem of Stone states that \mathcal{A} is the infinitesimal generator of a C_0 -semigroup, see [38]. We denote this semigroup and its two components by

$$\mathcal{G}(t) : X \times X \rightarrow X \times X, \quad \mathcal{G}(t) := \begin{pmatrix} \mathcal{E}(t) \\ \mathcal{B}(t) \end{pmatrix}. \quad (3.13)$$

As \mathcal{G} is strongly continuous, the following notion of solutions to (3.12) is meaningful:

Definition 3.7 (Mild solution of Maxwell's equations). *Let $(E_0, B_0) \in X \times X$ and $j \in L^1([0, T]; X)^2$ be given. Then we call $(E, B) \in C([0, T]; X)^2$, given by*

$$\begin{pmatrix} E(t) \\ B(t) \end{pmatrix} = \mathcal{G}(t) \begin{pmatrix} E_0 \\ B_0 \end{pmatrix} + \int_0^t \mathcal{G}(t-\tau) j(r, p)(\tau) \, d\tau \quad 0 \leq t \leq T, \quad (3.14)$$

mild solution of the Cauchy problem (3.12) on $[0, T]$.

Note that the strong continuity of \mathcal{G} implies that the right-hand side in (3.14) indeed defines an element of $C([0, T]; X)^2$. Moreover, by strong continuity, there are constants $M \geq 1$ and $\omega \geq 0$ such that

$$\|\mathcal{G}(t)\|_{\mathcal{L}(X \times X, X \times X)} \leq M e^{\omega t} \quad \forall t \in [0, T] \quad (3.15)$$

giving in turn the following a priori estimate

$$\|(E, B)\|_{C([0,T];X \times X)} \leq 2Me^{\omega T} (\|(E_0, B_0)\|_{X \times X} + \|j\|_{L^1([0,T];X \times X)}). \quad (3.16)$$

Next we turn to the Poisson equation (3.8). As the Dirichlet data are given by the control function $u \in L^2(\Gamma)$, we employ the following notion of solutions:

Definition 3.8 (Very weak solution of Poisson equation). *For given $u \in L^2(\Gamma)$ we call $\eta \in L^2(\Omega)$ very weak solution of (3.8), if it solves the very weak formulation*

$$-\int_{\Omega} \eta \Delta v \, dx + \int_{\Gamma} u \partial_n v \, d\varsigma = 0 \quad \forall v \in \mathcal{H}. \quad (3.17)$$

Lemma 3.9. *For every $u \in L^2(\Gamma)$ there exists a unique solution $\eta \in L^2(\Omega)$ of (3.17) satisfying an a priori estimate*

$$\|\eta\|_{L^2(\Omega)} \leq C \|u\|_{L^2(\Gamma)}$$

with a constant $C > 0$ independent of u and η .

Proof. Assumption 3.4 and the open mapping theorem yield that $-\Delta^{-1} \in \mathcal{L}(L^2(\Omega), \mathcal{H})$ and consequently $(-\Delta^*)^{-1} \in \mathcal{L}(\mathcal{H}^*, L^2(\Omega))$. Moreover, by definition of \mathcal{H} , the mapping

$$R : L^2(\Gamma) \rightarrow \mathcal{H}^*, \quad \langle Ru, v \rangle_{\mathcal{H}^*, \mathcal{H}} := - \int_{\Gamma} u \partial_n v \, d\varsigma, \quad u \in L^2(\Gamma), \quad v \in \mathcal{H}$$

is linear and continuous. Therefore,

$$\eta = (-\Delta^*)^{-1} Ru \quad (3.18)$$

is the unique solution of (3.17). This immediately implies the a priori estimate with $C = \|(-\Delta)^{-1}\|_{\mathcal{L}(L^2(\Omega), \mathcal{H})} \|R\|_{\mathcal{L}(L^2(\Gamma), \mathcal{H}^*)}$. \square

Remark 3.10. *The low regularity of the very weak solution implies that the external magnetic field $b = \nabla \eta$ is in general only a distribution and no proper function. Note however that, thanks to integration by parts in (3.5), only η and u appear on the right hand side of (3.7a).*

Remark 3.11. *We point out that the magnetic field $b = \nabla \eta$ can be extended outside of Ω in a divergence-free manner. The boundary data u , i.e., the control function, can physically be interpreted as a surface current density on Γ . Naturally, one can, in general, not realize such current density in $L^2(\Gamma)$ in practice so that the numerical results presented in Section 7.4 are rather of theoretical interest.*

Based on the above findings, in particular (3.14) and (3.18), we can eliminate E , B , and η from the state system to obtain a system of equations in r , p , and u only. This gives rise to the following definition:

Definition 3.12 (Solution of state system). *Let the mappings j , F_L , and e be defined as follows:*

1. *Current density:*

$$j : C([0, T]; \mathbb{R}^3)^2 \rightarrow C([0, T]; X)^2, \quad j(r, p)(x, t) := \begin{pmatrix} -q \varphi(x - r(t))v(p(t)) \\ 0 \end{pmatrix},$$

2. *Lorentz force:*

$$\begin{aligned} F_L &: C([0, T]; \mathbb{R}^3)^2 \rightarrow C([0, T]; X) \\ F_L(r, p)(x, t) &:= E(x, t) + \beta(p(t)) \times B(x, t) \\ &= \mathcal{E}(t) \begin{pmatrix} E_0 \\ B_0 \end{pmatrix} + \int_0^t \mathcal{E}(t - \tau) j(r, p)(\tau) d\tau \\ &\quad + \beta(p(t)) \times \left(\mathcal{B}(t) \begin{pmatrix} E_0 \\ B_0 \end{pmatrix} + \int_0^t \mathcal{B}(t - \tau) j(r, p)(\tau) d\tau \right), \end{aligned}$$

with the components \mathcal{E} and \mathcal{B} of the semigroup \mathcal{G} , see (3.13)

3. *State system operator:*

$$e : C_{\{0\}}^1([0, T]; \mathbb{R}^3)^2 \times L^2(\Gamma) \rightarrow C([0, T]; \mathbb{R}^3)^2, \quad e(w, z, u) := \begin{pmatrix} e_1(w, z, u) \\ e_2(w, z, u) \end{pmatrix},$$

$$\begin{aligned} e_1(w, z, u)(t) &:= \dot{z}(t) - q \int_{\Omega} \varphi(x - w(t) - r_0) F_L(w + r_0, z + p_0)(t) dx \\ &\quad + q \int_{\Omega} ((-\Delta^*)^{-1} R u) \left[\nabla \varphi(x - w(t) - r_0) \times \beta(z(t) + p_0) \right] dx \\ &\quad - q \int_{\Gamma} u \varphi(x - w(t) - r_0) \beta(z(t) + p_0) \times n d\varsigma \\ e_2(w, z, u)(t) &:= \dot{w}(t) - v(z(t) + p_0). \end{aligned}$$

Then we call a triple $(w, z, u) \in C_{\{0\}}^1([0, T]; \mathbb{R}^3)^2 \times L^2(\Gamma)$ solution of the state system, if it satisfies $e(w, z, u) = 0$.

We point out that, due to the smoothness assumptions on φ in (3.1) and the regularity of the mild solution, see Definition (3.7), the mappings j , F_L , and e indeed possess the asserted mapping properties. Note that both PDEs, i.e., Maxwell's equations as well as the Poisson equation, are incorporated into this notion of solution by means of the solution operators of the respective PDE in form of (3.14) and (3.18). Therefore we call the equation $e(w, z, u) = 0$ *reduced (state) system*, as it only involves the variables w , z , and u .

With this notion of solution to the state system at hand, we are now in the position to state a mathematically rigorous version of the optimal control problem under consideration:

$$\left. \begin{aligned} \min \quad & \mathcal{J}(w + r_0, u) \\ \text{s.t.} \quad & w, z \in C_{\{0\}}^1([0, T]; \mathbb{R}^3), \quad u \in L^2(\Gamma) \\ & e(w, z, u)(t) = 0 \quad \forall t \in [0, T] \\ & g_i(w(t) + r_0) \leq 0, \quad i = 1, \dots, m, \quad \forall t \in [0, T]. \end{aligned} \right\} \quad (\text{P})$$

For the sake of clarity we recall all variables and their meaning in Table 3.1. Here and in all what follows, we denote the couple (w, z) by y . For completeness we also list the adjoint variables arising in the upcoming sections in this table.

4 Analysis of the state equation

We begin the discussion of (P) with an existence and uniqueness result for the reduced state system. To be more precise, we prove that, for every $u \in L^2(\Gamma)$, there exists a unique $y \in C_{\{0\}}^1([0, T]; \mathbb{R}^3)^2$ such that $e(y, u) = 0$. The proof is classical and based on Banach's contraction principle. It follows the lines of [28] and [42, Section 2.4], where existence and uniqueness is shown for the Abraham

Variable	Name of quantity
<u>State variables</u>	
E	electric field
B	magnetic field
r	position of particle
p	relativistic momentum of particle
w	normalized particle position
z	normalized momentum
$y := (w, z)$	
η	solution of Poisson equation
<u>Control variable</u>	
u	boundary data of Poisson equation
<u>Adjoint variables</u>	
Φ	adjoint electric field
Ψ	adjoint magnetic field
ϱ	adjoint particle position
π	adjoint relativistic momentum
$\omega := (\varrho, \pi)$	
χ	adjoint Poisson solution
μ	Lagrange multiplier
<u>Further variables</u>	
j	electric current
F_L	Lorentz force
ρ	charge density
γ	Lorentz factor
b	external magnetic field
e	external electric field
φ	smeared out delta distribution

TABLE 3.1. Overview of arising variables

model for the case $\Omega = \mathbb{R}^3$ and without the Poisson equation for the external magnetic field. Let $u \in L^2(\Gamma)$ be fix but arbitrary. The constraint $e(y, u) = 0$ in (P) is equivalent to

$$\dot{y}(t) = f(y, u)(t) \quad \forall t \in [0, T], \quad y(0) = 0, \quad (4.1)$$

where $f = (f_1, f_2) : C([0, T]; \mathbb{R}^3)^2 \rightarrow C([0, T]; \mathbb{R}^3)^2$ is given by

$$\begin{aligned}
f_1(w, z, u)(t) := & q \int_{\Omega} \varphi(x - w(t) - r_0) F_L(w + r_0, z + p_0)(t) dx \\
& - q \int_{\Omega} ((-\Delta^*)^{-1} R u) \left[\nabla \varphi(x - w(t) - r_0) \times \beta(z(t) + p_0) \right] dx \\
& + q \int_{\Gamma} u \varphi(x - w(t) - r_0) \beta(z(t) + p_0) \times n d\varsigma \\
f_2(w, z, u)(t) := & v(z(t) + p_0).
\end{aligned}$$

For the rest of this section we suppressed the dependency of f on u , as u is fixed throughout this section. In order to apply the Banach's fixed point theorem, we prove the following

Lemma 4.1. *The right hand side in the reduced system (4.1) is globally Lipschitz continuous with respect to y in the following sense*

$$|f(y_1)(t) - f(y_2)(t)|_2 \leq L \|y_1 - y_2\|_{C([0,t];\mathbb{R}^3)^2} \quad \forall t \in [0, T] \quad (4.2)$$

with Lipschitz constant $L \geq 0$.

Proof. First observe that, by definition of v in (2.4), we have

$$|v(p)|_2 \leq c, \quad |v'(p)|_F \leq \frac{\sqrt{3}}{m_0^q} \quad \forall p \in \mathbb{R}^3. \quad (4.3)$$

Moreover, (3.1) implies

$$\begin{aligned} \|\varphi(\cdot - r_1) - \varphi(\cdot - r_2)\|_{L^2(\Omega)} &\leq \sqrt{\frac{4}{3} \pi R^3 L_\varphi} |r_1 - r_2|_2 \quad \forall r_1, r_2 \in \mathbb{R}^3 \\ \|\varphi(\cdot - r_2)\|_{L^2(\Omega)} &\leq \sqrt{\|\varphi\|_{L^\infty(\mathbb{R}^3)} \|\varphi\|_{L^1(\mathbb{R}^3)}} = \sqrt{C_\varphi} \quad \forall r_2 \in \mathbb{R}^3, \end{aligned} \quad (4.4)$$

where $L_\varphi > 0$ denotes the Lipschitz constant of φ and $C_\varphi := \max_{x \in \mathbb{R}^3} |\varphi(x)|$. Note that φ is globally Lipschitz since it is continuously differentiable and has bounded support.

The assertion for f_2 follows from

$$\begin{aligned} &|f_2(y_1)(t) - f_2(y_2)(t)|_2 \\ &= |v(z_1(t) + p_0) - v(z_2(t) + p_0)|_2 \\ &\leq |v'(z_2(t) + p_0 + s(z_1(t) - z_2(t)))|_F |y_2(t) - y_1(t)|_2 \leq \frac{\sqrt{3}}{m_0^q} \|y_1 - y_2\|_{C([0,t];\mathbb{R}^3)^2}. \end{aligned}$$

To verify the global Lipschitz continuity of f_1 , we exemplary consider

$$\hat{f}(y)(t) := q \int_{\Omega} \varphi(x - w(t) - r_0) \left[\mathcal{E}(t) \begin{pmatrix} E_0 \\ B_0 \end{pmatrix} + \int_0^t \mathcal{E}(t - \tau) j(w + r_0, z + p_0)(\tau) d\tau \right] dx,$$

which is one of the terms that arise, if one inserts the definition of F_L into f_1 . Now let $t \in [0, T]$ and $y_1 = (w_1, z_1), y_2 = (w_2, z_2) \in C([0, t]; \mathbb{R}^3)^2$ be arbitrary. Using the abbreviations $r_i = w_i + r_0$ and $p_i = z_i + p_0$, $i = 1, 2$, we obtain by means of (3.15) that

$$\begin{aligned} &|\hat{f}(y_1)(t) - \hat{f}(y_2)(t)|_2 \\ &\leq q \left(\left\| \mathcal{E}(t) \begin{pmatrix} E_0 \\ B_0 \end{pmatrix} \right\|_X \right. \\ &\quad \left. + \int_0^t \|\mathcal{E}(t - \tau) j(r_1, p_1)(\tau)\|_X d\tau \right) \|\varphi(\cdot - r_1(t)) - \varphi(\cdot - r_2(t))\|_{L^2(\Omega)} \\ &\quad + q \int_0^t \|\mathcal{E}(t - \tau) j(r_1, p_1)(\tau) - \mathcal{E}(t - \tau) j(r_2, p_2)(\tau)\|_X d\tau \|\varphi(\cdot - r_2(t))\|_{L^2(\Omega)} \\ &\leq q M e^{\omega T} \left(\|j(r_1, p_1)\|_{L^1([0,t]; X \times X)} + \|(E_0, B_0)\|_{X \times X} \right) \sqrt{\pi} R L_\varphi |r_1(t) - r_2(t)|_2 \\ &\quad + q M e^{\omega T} \sqrt{C_\varphi} \|j(r_1, p_1) - j(r_2, p_2)\|_{L^1([0,t]; X \times X)}. \end{aligned}$$

Concerning the expressions involving j , we find by employing (4.3) and (4.4) that

$$\begin{aligned}
& \|j(r_1, p_1) - j(r_2, p_2)\|_{L^1([0,t]; X \times X)} \\
&= q \int_0^t \|\varphi(\cdot - r_1(\tau))v(p_1(\tau)) - \varphi(\cdot - r_2(\tau))v(p_2(\tau))\|_X d\tau \\
&\leq q \int_0^t \left(\|\varphi(\cdot - r_1(\tau)) - \varphi(\cdot - r_2(\tau))\|_{L^2(\Omega)} |v(p_1(\tau))|_2 \right. \\
&\quad \left. + \|\varphi(\cdot - r_2(\tau))\|_{L^2(\Omega)} |v(p_1(\tau)) - v(p_2(\tau))|_2 \right) d\tau \\
&\leq q T \left(\sqrt{\pi} R L_\varphi c \|r_1 - r_2\|_{C([0,t]; \mathbb{R}^3)} + \sqrt{C_\varphi} \frac{\sqrt{3}}{m_0^q} \|p_1 - p_2\|_{C([0,t]; \mathbb{R}^3)} \right)
\end{aligned}$$

and

$$\begin{aligned}
\|j(r_1, p_1)\|_{L^1([0,t]; X \times X)} &= q \int_0^t \|\varphi(\cdot - r_1(\tau))\|_{L^2(\Omega)} |v(p_1(\tau))|_2 d\tau \\
&\leq q T \sqrt{C_\varphi} c.
\end{aligned} \tag{4.5}$$

By inserting these estimates we end up with

$$\begin{aligned}
|\hat{f}(y_1)(t) - \hat{f}(y_2)(t)| &\leq K (\|r_1 - r_2\|_{C([0,t]; \mathbb{R}^3)} + \|p_1 - p_2\|_{C([0,t]; \mathbb{R}^3)}) \\
&\leq \sqrt{2} K \|y_1 - y_2\|_{C([0,t]; \mathbb{R}^3)^2}
\end{aligned}$$

with a constant $K > 0$ independent of t , y_1 , and y_2 . The Lipschitz continuity of the remaining parts in f_1 can be proven by similar estimates. \square

Remark 4.2. We point out that the Lipschitz constant in (4.2) depends on u so that one should rather write

$$|f(y_1, u)(t) - f(y_2, u)(t)|_2 \leq L(u) \|y_1 - y_2\|_{C([0,t]; \mathbb{R}^3)^2} \quad \forall t \in [0, T].$$

Of course, the proof of existence of a solution to (4.1) for fixed u is not affected by this dependency.

Based on the Lipschitz-estimate in Lemma 4.1, existence and uniqueness can now be shown by Banach's contraction principle. The arguments are classical and follow the lines of [42, Section 2.4]. For convenience of the reader we sketch the proof in Appendix A.

Theorem 4.3. For all $u \in L^2(\Gamma)$ there exists a unique solution $y \in C_{\{0\}}^1([0, T]; \mathbb{R}^3)^2$ of the reduced system (4.1) and the following a priori estimate is fulfilled

$$\|y\|_{C^1([0,T]; \mathbb{R}^3)^2} \leq C_1 \|u\|_{L^2(\Gamma)} + C_2$$

with a constants $C_1, C_2 > 0$ independent of u and y .

5 Existence of an optimal control

With the existence result for the reduced state system in Theorem 4.3 at hand, it is now straightforward to establish the existence of a globally optimal control.

Theorem 5.1. Assume that there is a control $u \in L^2(\Gamma)$ such that the associated state $y = (w, z) \in C_{\{0\}}^1([0, T]; \mathbb{R}^3)^2$ satisfies the state constraint $g_i(w(t) + r_0) \leq 0$ for all $i = 1, \dots, m$ and all $t \in [0, T]$. Then there exists at least one globally optimal control for (P).

Proof. By assumption the feasible set of (P) is non-empty. Thus there exists a minimizing sequence $\{y_n, u_n\} = \{w_n, z_n, u_n\} \subset C_{\{0\}}^1([0, T]; \mathbb{R}^3)^2 \times L^2(\Gamma)$, i.e., $e(y_n, u_n) = 0$, $w_n(t) + r_0 \in \tilde{\Omega}$ for all $t \in [0, T]$, and

$$\mathcal{J}(w_n + r_0, u_n) \xrightarrow{n \rightarrow \infty} \inf (P) =: j \in \mathbb{R} \cup \{-\infty\}.$$

From Assumption 3.6 we deduce

$$\frac{\alpha}{2} \|u_n\|_{L^2(\Gamma)}^2 \leq \mathcal{J}(w_n + r_0, u_n) - \underline{c}_1 T - \underline{c}_2$$

so that $\{u_n\}$ is bounded in $L^2(\Gamma)$. As $e(y_n, u_n) = 0$, Theorem 4.3 yields the boundedness of $\{y_n\}$ in $H^1([0, T]; \mathbb{R}^3)^2$. Consequently, there exist weakly converging subsequences, and w.l.o.g. we assume weak convergence of the whole sequences, i.e.

$$u_n \rightharpoonup u^* \text{ in } L^2(\Gamma) \quad \text{and} \quad y_n \rightharpoonup y^* = (w^*, z^*) \text{ in } H^1([0, T]; \mathbb{R}^3)^2.$$

The compactness of the embedding $H^1([0, T]; \mathbb{R}^3)^2 \hookrightarrow C([0, T]; \mathbb{R}^3)^2$ then yields strong convergence of $\{y_n\}$ in the maximum-norm so that Lemma 4.1 and Remark 4.2 give

$$\|f(y_n, u^*) - f(y^*, u^*)\|_{C([0, T]; \mathbb{R}^3)^2} \leq L(u^*) \|y_n - y^*\|_{C([0, T]; \mathbb{R}^3)^2} \xrightarrow{n \rightarrow \infty} 0.$$

Moreover, the strong convergence of the state in $C([0, T]; \mathbb{R}^3)^2$ further implies

$$\|\beta(p_n) - \beta(p^*)\|_{C([0, T]; \mathbb{R}^3)} \rightarrow 0, \quad \|\varphi(\cdot - r_n) - \varphi(\cdot - r^*)\|_{C([0, T]; H^1(\Omega))} \rightarrow 0.$$

As the control only appears linearly in the state system, these convergences allow to pass to the limit in the reduced state equation in weak form, i.e., for every $v = (v_1, v_2) \in L^2(0, T; \mathbb{R}^3)^2$ there holds

$$\begin{aligned} & \int_0^T \dot{y}^*(t) \cdot v(t) dt \\ &= \lim_{n \rightarrow \infty} \int_0^T \dot{y}_n(t) \cdot v(t) dt \\ &= \lim_{n \rightarrow \infty} \int_0^T f(y_n, u_n)(t) \cdot v(t) dt \\ &= \lim_{n \rightarrow \infty} \left(\int_0^T f(y_n, u^*)(t) \cdot v(t) dt \right. \\ & \quad \left. - q \int_{\Omega} ((-\Delta^*)^{-1} R(u_n - u^*)) \int_0^T [\nabla \varphi(x - r_n(t)) \times \beta(p_n(t))] \cdot v_1(t) dt dx \right. \\ & \quad \left. + q \int_{\Gamma} (u_n - u^*) \int_0^T [\varphi(x - r_n(t)) \beta(p_n(t)) \times n] \cdot v_1(t) dt d\varsigma \right) \\ &= \int_0^T f(y^*, u^*)(t) \cdot v(t) dt. \end{aligned}$$

Therefore, we obtain

$$\dot{y}^*(t) = f(y^*, u^*)(t) \quad \text{f.a.a. } t \in [0, T].$$

Because of $y^* \in C([0, T]; \mathbb{R}^3)^2$ the right hand side is continuous such that $y^* \in C^1([0, T]; \mathbb{R}^3)^2$. From $y_n \rightarrow y^*$ in $C([0, T]; \mathbb{R}^3)^2$ we further infer that $y^*(0) = 0$, and consequently y^* coincides with the unique solution of (4.1) associated with u^* . The convergence of the state in $C([0, T]; \mathbb{R}^3)^2$ and the continuity of g_i , $i = 1, \dots, m$, moreover yield

$$g_i(w^*(t) + r_0) \leq 0 \quad \forall i = 1, \dots, m \quad \Leftrightarrow \quad w^*(t) + r_0 \in \tilde{\Omega}$$

for all $t \in [0, T]$ such that the state constraint is also fulfilled in the limit. Therefore, the couple (y^*, u^*) fulfills all constraints in (P).

Finally, the strong convergence of $\{y_n\}$ in $C([0, T]; \mathbb{R}^3)^2$, the weak convergence of $\{u_n\}$ in $L^2(\Gamma)$, and the weak lower semicontinuity of $\|\cdot\|_{L^2(\Gamma)}^2$ allow to pass to the limit in the objective:

$$\begin{aligned} j &= \lim_{n \rightarrow \infty} \mathcal{J}(w_n + r_0, u_n) \\ &\geq \lim_{n \rightarrow \infty} \left(\int_0^T J_1(w_n(t) + r_0) dt + J_2(w_n(T) + r_0) \right) + \liminf_{n \rightarrow \infty} \frac{\alpha}{2} \int_{\Gamma} u_n^2 d\zeta \\ &\geq \mathcal{J}(w^* + r_0, u^*), \end{aligned}$$

which implies the optimality of (y^*, u^*) . \square

6 First-order necessary optimality conditions

For the rest of the paper, we slightly change the functional analytical framework of the optimal control problem under consideration. To be more precise, we weaken the regularity of the state space in order to obtain a more regular adjoint state and treat the state as a function in

$$Y = \{y \in H^1([0, T]; \mathbb{R}^3)^2 : y(0) = 0\}.$$

Thus the mapping associated with the reduced state system becomes $e : Y \times L^2(\Gamma) \rightarrow Z = L^2([0, T]; \mathbb{R}^3)^2$, with a slight abuse of notation still denoted by e . It is easily seen that this modification does not affect the above analysis, in particular the proof of existence of an optimal control, since the state is treated as a function in $H^1([0, T]; \mathbb{R}^3)^2$ there anyway. Note that $H^1([0, T]; \mathbb{R}^3)^2 \hookrightarrow C([0, T]; \mathbb{R}^3)^2$ so that the mappings j and F_L from Definition 3.12 are still well-defined.

Remark 6.1. *If a couple $(y, u) \in Y \times L^2(\Gamma)$ satisfies the constraint $e(y, u) = 0$, i.e.,*

$$\dot{y}(t) = f(y, u)(t) \quad \text{f.a.a. } t \in [0, T], \quad y(0) = 0,$$

then $f(y, u) \in C([0, T]; \mathbb{R}^3)^2$ implies $y \in C^1([0, T]; \mathbb{R}^3)^2$ so that y coincides with the unique solution of (4.1) from Theorem 4.3. In other words, the treatment of (P) in the weaker state space Y does not affect the regularity of the optimal state.

6.1. The linearized state equation. We start the derivation of a qualified optimality system by the analysis of the linearized reduced state system.

Lemma 6.2. *The reduced form e is continuously Fréchet-differentiable from $Y \times L^2(\Gamma)$ to Z . Its partial derivatives at $(y, u) = (w, z, u) \in Y \times L^2(\Gamma)$ in direction $(\phi, h) = (\phi_r, \phi_p, h) \in Y \times L^2(\Gamma)$ are given by*

$$\begin{aligned} \left(\frac{\partial e_1}{\partial y}(y, u)\phi \right)(t) &= \dot{\phi}_p(t) - \left(\frac{\partial f_1}{\partial y}(y, u)\phi \right)(t), \\ \left(\frac{\partial e_2}{\partial y}(y, u)\phi \right)(t) &= \dot{\phi}_r(t) - \left(\frac{\partial f_2}{\partial y}(y, u)\phi \right)(t), \\ \left(\frac{\partial e_1}{\partial u}(y, u)h \right)(t) &= - \left(\frac{\partial f_1}{\partial u}(y, u)h \right)(t), \quad \left(\frac{\partial e_2}{\partial u}(y, u)h \right)(t) = 0 \end{aligned}$$

with

$$\begin{aligned} \left(\frac{\partial f_1}{\partial u}(y, u)h\right)(t) &= q \int_{\Gamma} h\varphi(x - r(t))\beta(p(t)) \times n \, d\zeta \\ &\quad - q \int_{\Omega} (-\Delta^*)^{-1} Rh[\nabla\varphi(x - r(t)) \times \beta(p(t))] \, dx, \\ \left(\frac{\partial f_2}{\partial y}(y, u)\phi\right)(t) &= v'(p(t))\phi_p(t), \end{aligned}$$

and

$$\begin{aligned} \left(\frac{\partial f_1}{\partial y}(y, u)\phi\right)(t) &= -q \int_{\Omega} [\nabla\varphi(x - r(t)) \cdot \phi_r(t)] F_L(r, p)(t) \, dx \\ &\quad + q \int_{\Omega} \varphi(x - r(t)) (\partial_y F_L(r, p)\phi)(t) \, dx \\ &\quad + q \int_{\Gamma} u \left[\varphi(x - r(t)) \beta'(p(t))\phi_p(t) \right. \\ &\quad \quad \left. - [\nabla\varphi(x - r(t)) \cdot \phi_r(t)] \beta(p(t)) \right] \times n \, d\zeta \\ &\quad + q \int_{\Omega} ((-\Delta^*)^{-1} Ru) \left[\nabla^2\varphi(x - r(t))\phi_r(t) \times \beta(p(t)) \right. \\ &\quad \quad \left. - \nabla\varphi(x - r(t)) \times \beta'(p(t))\phi_p(t) \right] \, dx, \end{aligned}$$

with $r = w + r_0$, $p = z + p_0$, the derivative of the Lorentz force term F_L

$$\begin{aligned} (\partial_y F_L(r, p)\phi)(t) &= (\partial_r F_L(r, p)\phi_r + \partial_p F_L(r, p)\phi_p)(t) \\ &= \int_0^t \mathcal{E}(t - \tau) (j'(r, p)(\tau)\phi(\tau)) \, d\tau \\ &\quad + \beta(p(t)) \times \int_0^t \mathcal{B}(t - \tau) (j'(r, p)(\tau)\phi(\tau)) \, d\tau \\ &\quad + \beta'(p(t))\phi_p(t) \times \left(\mathcal{B}(t) \begin{pmatrix} E_0 \\ B_0 \end{pmatrix} + \int_0^t \mathcal{B}(t - \tau) j(r, p)(\tau) \, d\tau \right) \end{aligned}$$

and j' as given in (3.11).

Proof. As a linear and bounded operator the time derivative is clearly continuously Fréchet-differentiable for $H^1([0, T]; \mathbb{R}^3)$ to $L^2([0, T]; \mathbb{R}^3)$. All nonlinear Nemyzki-operators involved in f are differentiated in spaces of continuous functions. Because of its slightly non-standard structure, we exemplarily study the Fréchet-differentiability of $r \mapsto \nabla\varphi(\cdot - r)$ from $C([0, T]; \mathbb{R}^3)$ to $C([0, T]; L^2(\Omega))$:

$$\begin{aligned} &\|\nabla\varphi(\cdot - (r + \phi_r)) - \nabla\varphi(\cdot - r) - \nabla^2\varphi(\cdot - r) \cdot \phi_r\|_{C([0, T]; L^2(\Omega))}^2 \\ &= \max_{t \in [0, T]} \left(\int_{\Omega} |\nabla\varphi(x - r(t) - \phi_r(t)) - \nabla\varphi(x - r(t)) - \nabla^2\varphi(x - r(t))\phi_r(t)|^2 \, dx \right. \\ &= \max_{t \in [0, T]} \int_{\Omega} \left| \int_0^1 \nabla^2\varphi(x - r(t) - \theta\phi_r(t))\phi_r(t) \, d\theta - \nabla^2\varphi(x - r(t))\phi_r(t) \right|^2 \, dx \\ &\leq \max_{t \in [0, T]} \int_{\Omega} \left| \int_0^1 L_{\varphi, 2} \theta |\phi_r(t)|^2 \, d\theta \right|^2 \, dx = \frac{1}{4} L_{\varphi, 2}^2 |\Omega| \|\phi_r(t)\|_{C([0, T]; \mathbb{R}^3)}^4, \end{aligned}$$

where $L_{\varphi, 2}$ denotes the Lipschitz constant of $\nabla^2\varphi$. This gives the partial differentiability of f w.r.t. y . As u only appears linearly, f is moreover partially differentiable w.r.t. u . Furthermore, one readily verifies that these partial derivatives are continuous in (y, u) . Therefore, [8, Theorem 3.7.1] gives the continuous Fréchet-differentiability of e . \square

Lemma 6.3. *Let $(y, u) \in Y \times L^2(\Gamma)$ be given. Then for every $h \in Z$ there exists a unique solution $\phi = (\phi_r, \phi_p) \in Y$ of the linearized equation*

$$\frac{\partial e}{\partial y}(y, u)\phi = h. \quad (6.1)$$

Proof. In view of Lemma 6.2, (6.1) is equivalent to

$$\begin{pmatrix} \dot{\phi}_p(t) \\ \phi_r(t) \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial y}(y, u)\phi \end{pmatrix}(t) + h(t) \quad \text{f.a.a. } t \in [0, T], \quad \phi(0) = 0$$

with $\partial_y f(y, u)\phi = (\partial_y f_1(y, u)\phi, \partial_y f_2(y, u)\phi)$. As in the proof of Theorem 4.3, existence and uniqueness of the equivalent integral equation, given by

$$\begin{pmatrix} \phi_p(t) \\ \phi_r(t) \end{pmatrix} = \int_0^t \left[\begin{pmatrix} \frac{\partial f}{\partial y}(y, u)\phi \end{pmatrix}(\tau) + h(\tau) \right] d\tau,$$

can again be proven by Banach's contraction principle, provided that there is a constant $C > 0$ such that

$$\left| \begin{pmatrix} \frac{\partial f}{\partial y}(y, u)\phi \end{pmatrix}(t) \right|_2 \leq C \|\phi\|_{C([0, t]; \mathbb{R}^3)^2} \quad \forall t \in [0, T],$$

cf. (4.2). (Note in this context that $\phi \mapsto \partial_y f(y, u)\phi$ is a linear mapping so that Lipschitz continuity is equivalent to boundedness.) The latter inequality however can be verified by estimates similar to the proof of Lemma 4.1. \square

6.2. KKT conditions. Having established the differentiability of the reduced state system, we are now in the position to derive first-order optimality system in qualified form, i.e., Karush-Kuhn-Tucker (KKT) conditions involving Lagrange multipliers associated with the constraints in (P). To this end, let $(y^*, u^*) = (w^*, z^*, u^*) \in Y \times L^2(\Gamma)$ be a arbitrary local optimum of (P). As before, we set $r^* = w^* + r_0$ and $p^* = z^* + p_0$ in all what follows. It is known that the existence of Lagrange multipliers requires certain constraint qualifications, see e.g. [51]. In our case, one of these, namely the surjectivity of $\partial_y e(y^*, u^*)$, was established in Lemma 6.3. However, we need an additional condition to obtain a Lagrange multiplier for the pointwise state constraint in (P), too.

Assumption 6.4 (Linearized Slater condition). *We assume that there is a function $\hat{h} \in L^2(\Gamma)$ so that*

$$g_i(r^*(t)) + g'_i(r^*(t))\hat{\phi}_r(t) < 0 \quad \forall t \in [0, T], \quad i = 1, \dots, m, \quad (6.2)$$

where $\hat{\phi} = (\hat{\phi}_r, \hat{\phi}_p) \in Y$ is the solution to (6.1) for $h = \hat{h}$.

Note that the Nemyzki operators associated with g_1, \dots, g_m are Fréchet-differentiable from $C([0, T]; \mathbb{R}^3)$ to $C([0, T])$ by Assumption 3.6. The same holds for the functions J_1 and J_2 within the objective.

Given that Assumption 6.4 is fulfilled, one can establish the existence of Lagrange multipliers, see for instance [25, Section 1.7.3.4]. To be more precise, under Assumption 6.4 there exists $(\pi, \varrho, \lambda) \in Z \times C([0, T]; \mathbb{R}^m)^*$ such that the following

KKT conditions are satisfied:

$$e(y^*, u^*)(t) = 0 \quad \forall t \in [0, T] \quad (6.3a)$$

$$\frac{\partial e}{\partial y}(y^*, u^*)^* \begin{pmatrix} \pi \\ \varrho \end{pmatrix} - \frac{\partial \mathcal{J}}{\partial y}(r^*, u^*) - \begin{pmatrix} g'(r^*)^* \lambda \\ 0 \end{pmatrix} = 0 \quad \text{in } Y^* \quad (6.3b)$$

$$\frac{\partial \mathcal{J}}{\partial u}(r^*, u^*) + \frac{\partial e}{\partial u}(y^*, u^*)^* \begin{pmatrix} \pi \\ \varrho \end{pmatrix} = 0 \quad \text{in } L^2(\Gamma) \quad (6.3c)$$

$$\begin{aligned} g_i(r^*(t)) &\leq 0 \quad \forall t \in [0, T], \\ \lambda_i &\geq 0, \quad \langle \lambda_i, g_i(r^*) \rangle_{C([0, T])^*, C([0, T])} = 0, \quad i = 1, \dots, m. \end{aligned} \quad (6.3d)$$

Herein the inequality $\lambda_i \geq 0$ is to be understood in a distributional sense, i.e., $\langle \lambda_i, v \rangle \geq 0$ for all $v \in C([0, T])$ with $v(t) \geq 0$ for all $t \in [0, T]$. Moreover, we set $g := (g_1, \dots, g_m)$ and denote by g' the associated Jacobian.

For the rest of this section, we aim to transfer (6.3b) to an adjoint system and to evaluate the gradient equation in (6.3c). We start with (6.3b), which in variational form reads as follows

$$\begin{aligned} &\int_0^T \left[\pi(t) \cdot \left(\frac{\partial e_1}{\partial y}(y^*, u^*) \phi \right)(t) + \varrho(t) \cdot \left(\frac{\partial e_2}{\partial y}(y^*, u^*) \phi \right)(t) \right] dt \\ &- \left\langle \frac{\partial \mathcal{J}}{\partial r}(r^*, u^*), \phi_r \right\rangle_{Y^*, Y} - \langle \lambda, g'(r^*) \phi_r \rangle_{C([0, T]; \mathbb{R}^m)^*, C([0, T]; \mathbb{R}^m)} = 0 \end{aligned} \quad (6.4)$$

$\forall \phi \in Y.$

By employing Lemma 6.2 we find for the first term in (6.4)

$$\begin{aligned} &\int_0^T \pi(t) \cdot \left(\frac{\partial e_1}{\partial y}(y^*, u^*) \phi \right)(t) dt \\ &= \int_0^T \dot{\phi}_p(t) \cdot \pi(t) dt - q I_L(\pi, y^*, \phi) \\ &+ q \int_0^T \phi_r(t) \cdot \left(\int_{\Omega} [F_L(r^*, p^*)(t) \cdot \pi(t)] \nabla \varphi(x - r^*(t)) dx \right. \\ &\quad \left. + \int_{\Omega} ((-\Delta^*)^{-1} R u^*) [\beta(p^*(t)) \times \nabla^2 \varphi(x - r^*(t)) \pi(t)] dx \right. \\ &\quad \left. - \int_{\Gamma} u^* [\nabla \varphi(x - r^*(t)) \cdot \pi(t)] (n \times \beta(p^*(t))) d\varsigma \right) dt \\ &- q \int_0^T \phi_p(t) \cdot \left(\int_{\Omega} ((-\Delta^*)^{-1} R u^*) [\beta'(p^*(t)) \pi(t) \times \nabla \varphi(x - r^*(t))] dx \right. \\ &\quad \left. - \int_{\Gamma} u^* \varphi(x - r^*(t)) (n \times \beta'(p^*(t)) \pi(t)) d\varsigma \right) dt, \end{aligned}$$

where $I_L(\pi, y^*, \phi)$ is defined by

$$I_L(\pi, y^*, \phi) := \int_0^T \pi(t) \cdot \int_{\Omega} \varphi(x - r^*(t)) (\partial_y F_L(r^*, p^*) \phi)(t) dx dt.$$

In view of Lemma 6.2, applying Fubini's theorem to this expression leads to

$$\begin{aligned}
& I_L(\pi, y^*, \phi) \\
&= \int_0^T \int_0^t \int_{\Omega} \left[\mathcal{E}(t-\tau) (j'(r^*, p^*)(\tau) \phi(\tau)) \right. \\
&\quad \left. + \beta(p^*(t)) \times \mathcal{B}(t-\tau) (j'(r^*, p^*)(\tau) \phi(\tau)) \right] \cdot \varphi(x - r^*(t)) \pi(t) dx d\tau dt \\
&\quad + \int_0^T \int_{\Omega} \varphi(x - r^*(t)) \beta'(p(t)) \phi_p(t) \times B^*(t) \cdot \pi(t) dx dt \\
&= \int_0^T \phi(t) \cdot \int_{\Omega} j'(r^*, p^*)(t)^\top \int_t^T \mathcal{G}(\tau-t)^* \kappa(r^*, p^*, \pi)(\tau) d\tau dx dt \\
&\quad + \int_0^T \phi_p(t) \cdot \int_{\Omega} B^*(t) \times \varphi(x - r^*(t)) \beta'(p^*(t)) \pi(t) dx dt,
\end{aligned}$$

where we abbreviated

$$B^*(t) := \mathcal{B}(t) \begin{pmatrix} E_0 \\ B_0 \end{pmatrix} + \int_0^t \mathcal{B}(t-\tau) j(r, p)(\tau) d\tau$$

and set

$$\begin{aligned}
& \kappa : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3 \times \mathbb{R}^3 \\
& \kappa(r, p, \pi) := \begin{pmatrix} \varphi(x-r) \pi \\ \pi \times \varphi(x-r) \beta(p) \end{pmatrix}.
\end{aligned}$$

Moreover let us define

$$\begin{pmatrix} \Phi(t) \\ \Psi(t) \end{pmatrix} := \int_t^T \mathcal{G}(\tau-t)^* \kappa(r^*, p^*, \pi)(\tau) d\tau \quad (6.5)$$

Since $-i\mathcal{A}$ is self-adjoint, the theorem of Stone implies that $\mathcal{G}(t)^*$ is the semigroup generated by the adjoint operator

$$\mathcal{A}^* : X \times X \rightarrow X \times X, \quad \mathcal{A}^* = \begin{pmatrix} 0 & \text{curl} \\ -\text{curl} & 0 \end{pmatrix}$$

with domain $D(\mathcal{A}^*) = D(\mathcal{A}) = H_{\text{curl}}^\Gamma$. Thus $(\Phi, \Psi) \in C([0, T]; X)^2$ is the mild solution of the following backward-in-time problem:

$$\begin{aligned}
& -\frac{\partial}{\partial t} \begin{pmatrix} \Phi(t) \\ \Psi(t) \end{pmatrix} + \mathcal{A}^* \begin{pmatrix} \Phi(t) \\ \Psi(t) \end{pmatrix} = \begin{pmatrix} \varphi(\cdot - r^*(t)) \pi(t) \\ \pi(t) \times \varphi(\cdot - r^*(t)) \beta(p^*(t)) \end{pmatrix} \\
& \Phi(T) = \Psi(T) = 0.
\end{aligned} \quad (6.6)$$

By setting $\omega := (\varrho, \pi)$ and summarizing the above transformations, we obtain for the first two addends in (6.4)

$$\begin{aligned}
& \int_0^T \left[\pi(t) \cdot \left(\frac{\partial e_1}{\partial y}(y^*, u^*) \phi \right)(t) + \varrho(t) \cdot \left(\frac{\partial e_2}{\partial y}(y^*, u^*) \phi \right)(t) \right] dt \\
&= \int_0^T \dot{\phi}(t) \cdot \omega(t) dt + \int_0^T \phi(t) \cdot A(y^*, u^*, \omega)(t) dt
\end{aligned}$$

with

$$\begin{aligned}
A(y^*, u^*, \omega)(t) &= \begin{pmatrix} A_r(y^*, u^*, \omega)(t) \\ A_p(y^*, u^*, \omega)(t) \end{pmatrix} \\
&:= q \begin{pmatrix} \int_{\Omega} [F_L(r^*, p^*)(t) \cdot \pi(t)] \nabla \varphi(x - r^*(t)) dx \\ - \int_{\Omega} B^*(t) \times \varphi(x - r^*(t)) \beta'(p^*(t)) \pi(t) dx \end{pmatrix} \\
&\quad + q \begin{pmatrix} \int_{\Omega} \eta^* [\beta(p^*(t)) \times \nabla^2 \varphi(x - r^*(t)) \pi(t)] dx \\ - \int_{\Omega} \eta^* [\beta'(p^*(t)) \pi(t) \times \nabla \varphi(x - r^*(t))] dx \end{pmatrix} \\
&\quad + q \begin{pmatrix} - \int_{\Gamma} u^* [\nabla \varphi(x - r^*(t)) \cdot \pi(t)] (n \times \beta(p^*(t))) d\varsigma \\ \int_{\Gamma} u^* \varphi(x - r^*(t)) (n \times \beta'(p^*(t)) \pi(t)) d\varsigma \end{pmatrix} \\
&\quad + q^2 \begin{pmatrix} - \int_{\Omega} v(p^*(t)) [\nabla \varphi(x - r^*(t)) \cdot \Phi(t)] dx \\ \int_{\Omega} \varphi(x - r^*(t)) v'(p^*(t)) \Phi(t) dx \end{pmatrix} - \begin{pmatrix} 0 \\ v'(p^*(t)) \varrho(t) \end{pmatrix},
\end{aligned} \tag{6.7}$$

where $\eta^* = (-\Delta^*)^{-1} R u^*$. Thus the adjoint equation (6.4) becomes

$$\begin{aligned}
\int_0^T \dot{\phi}(t) \cdot \omega(t) dt + \int_0^T \phi(t) \cdot A(y^*, u^*, \omega)(t) dt - \left\langle \frac{\partial \mathcal{J}}{\partial r}(r^*, u^*), \phi_r \right\rangle_{Y^*, Y} \\
- \langle \lambda, g'(r^*) \phi_r \rangle_{C([0, T]; \mathbb{R}^m)^*, C([0, T]; \mathbb{R}^m)} = 0 \quad \forall \phi \in Y.
\end{aligned} \tag{6.8}$$

By the Riesz representation theorem $\lambda \in C([0, T]; \mathbb{R}^m)^*$ can be identified with a function of bounded variations. This leads to the following result, whose detailed proof is given in Appendix B.

Lemma 6.5. *The adjoint particle position ϱ and the adjoint momentum π satisfy $\varrho \in \text{BV}([0, T]; \mathbb{R}^3)$ and $\pi \in W^{1, \infty}([0, T]; \mathbb{R}^3)$. Together with a function $\mu \in \text{NBV}([0, T]; \mathbb{R}^m)$ they fulfill the following ODEs backward in time:*

$$-\dot{\pi}(t) = -A_p(y^*, u^*, \varrho, \pi)(t) \quad \text{a.e. in }]0, T[\tag{6.9}$$

$$\pi(T) = 0 \tag{6.10}$$

$$-\dot{\varrho}(t) = -A_r(y^*, u^*, \varrho, \pi)(t) + \nabla J_1(r^*(t)) - g'(r^*(t))^\top \dot{\mu}(t) \quad \text{a.e. in }]0, T[\tag{6.11}$$

$$\varrho(T) = \nabla J_2(r^*(T)). \tag{6.12}$$

In addition, μ is monotone increasing and satisfies

$$\int_0^T g(r^*(t)) \cdot d\mu(t) = 0.$$

Moreover, ϱ only admits finitely many points of discontinuity t_1, \dots, t_ℓ in $]0, T[$, at each of which

$$\varrho(t_i) - \lim_{\varepsilon \searrow 0} \varrho(t_i - \varepsilon) = g'(r^*(t_i))^\top \left(\lim_{\varepsilon \searrow 0} \mu(t_i - \varepsilon) - \mu(t_i) \right), \quad i = 1, \dots, \ell, \tag{6.13}$$

holds true.

Next we turn to the gradient equation (6.3c). Focusing on the second addend in (6.3c), we obtain by means of Lemma 6.2 that

$$\begin{aligned}
& \int_{\Gamma} \left(\frac{\partial e}{\partial u}(y^*, u^*)^* \omega \right) \phi_u d\zeta \\
&= \int_0^T \left(\frac{\pi(t)}{\varrho(t)} \right) \cdot \left(\frac{\partial e}{\partial u}(y^*, u^*) \phi_u \right)(t) dt \\
&= q \int_0^T \pi(t) \cdot \int_{\Omega} ((-\Delta^*)^{-1} R \phi_u) [\nabla \varphi(x - r^*(t)) \times \beta(p^*(t))] dx dt \\
&\quad - q \int_0^T \pi(t) \cdot \int_{\Gamma} \phi_u \varphi(x - r^*(t)) \beta(t, p^*(t)) \times n d\zeta dt \\
&= \int_{\Gamma} \phi_u q \int_0^T \left(R^*(-\Delta)^{-1} \left[(\nabla \varphi(x - r^*(t)) \times \beta(p^*(t))) \cdot \pi(t) \right] \right. \\
&\quad \left. - [\varphi(x - r^*(t)) \beta(p^*(t)) \times n] \cdot \pi(t) \right) dt d\zeta.
\end{aligned}$$

Let us define the adjoint Poisson solution by

$$\chi(t) := -\Delta^{-1} \left[(\nabla \varphi(\cdot - r^*(t)) \times \beta(p^*(t))) \cdot \pi(t) \right] \in \mathcal{H}.$$

Note that the regularity w.r.t. time carries over from π to χ so that

$$\chi \in W^{1,\infty}([0, T]; \mathcal{H}).$$

Then, in view of $\partial_u \mathcal{J}(r^*, u^*) = \alpha u^* \in L^2(\Gamma)$ and $R^* = -\partial_n : \mathcal{H} \rightarrow L^2(\Gamma)$, the gradient equation (6.3c) becomes

$$\begin{aligned}
& \int_{\Gamma} \left(q \int_0^T \left[-\partial_n \chi(t) - [\varphi(x - r^*(t)) \beta(p^*(t)) \times n] \cdot \pi(t) \right] dt + \alpha u^* \right) \phi_u d\zeta = 0 \\
& \qquad \qquad \qquad \forall \phi_u \in L^2(\Gamma)
\end{aligned}$$

and the fundamental lemma of calculus of variations yields

$$u^*(x) = \frac{q}{\alpha} \int_0^T \left[\partial_n \chi(x, t) + [\varphi(x - r^*(t)) \beta(p^*(t)) \times n] \cdot \pi(t) \right] dt \quad \text{a.e. on } \Gamma.$$

Summarizing the results we have, thus, derived the following first-order necessary optimality conditions for (P):

Theorem 6.6 (KKT conditions). *Let $u^* \in L^2(\Gamma)$ be a locally optimal boundary control with associated states $(E^*, B^*, \eta^*, r^*, p^*) \in C([0, T]; X)^2 \times L^2(\Omega) \times C^1([0, T]; \mathbb{R}^3)^2$. Assume further that the linearized Slater condition in Assumption 6.4 is fulfilled. Then there exist adjoint states*

$$(\Phi, \Psi, \chi, \varrho, \pi) \in C([0, T]; X)^2 \times W^{1,\infty}([0, T]; \mathcal{H}) \times \text{BV}([0, T]; \mathbb{R}^3) \times W^{1,\infty}([0, T]; \mathbb{R}^3)$$

and a Lagrange multiplier $\mu \in \text{NBV}([0, T]; \mathbb{R}^m)$ so that following optimality system is fulfilled:

State equations:

Maxwell equations:

$$\begin{aligned}
& \frac{\partial}{\partial t} \begin{pmatrix} E^*(t) \\ B^*(t) \end{pmatrix} + \mathcal{A} \begin{pmatrix} E^*(t) \\ B^*(t) \end{pmatrix} = \begin{pmatrix} -q \varphi(\cdot - r^*(t)) v(p^*(t)) \\ 0 \end{pmatrix} \quad \text{a.e. in } [0, T] \\
& E^*(0) = E_0, \quad B^*(0) = B_0
\end{aligned}$$

Newton-Lorenz equation:

$$\begin{aligned} \dot{p}^*(t) &= q \left(\int_{\Omega} \varphi(x - r^*(t)) (E^*(t) + \beta(p^*(t)) \times B^*(t)) dx \right. \\ &\quad \left. + \int_{\Gamma} u^* \varphi(x - r^*(t)) \beta(p^*(t)) \times n d\zeta \right. \\ &\quad \left. - \int_{\Omega} \eta^* [\nabla \varphi(x - r^*(t)) \times \beta(p^*(t))] dx \right) \quad \forall t \in [0, T] \\ p^*(0) &= 0 \\ \dot{r}^*(t) &= v(p^*(t)) \quad \forall t \in [0, T] \\ r^*(0) &= 0 \end{aligned}$$

Poisson's equation in very weak form:

$$\int_{\Omega} \eta^* \Delta v \, dx = \int_{\Gamma} u^* \partial_n v \, d\varsigma \quad \forall v \in \mathcal{H}$$

Adjoint equations:

Adjoint Maxwell equations:

$$\begin{aligned} -\frac{\partial}{\partial t} \begin{pmatrix} \Phi(t) \\ \Psi(t) \end{pmatrix} + \mathcal{A}^* \begin{pmatrix} \Phi(t) \\ \Psi(t) \end{pmatrix} &= \begin{pmatrix} \varphi(\cdot - r^*(t)) \pi(t) \\ \pi(t) \times \varphi(\cdot - r^*(t)) \beta(p^*(t)) \end{pmatrix} \\ \Phi(T) = \Psi(T) &= 0 \end{aligned}$$

Adjoint ODE system:

$$\begin{aligned}
-\dot{\pi}(t) = & q \int_{\Omega} B^*(t) \times \varphi(x - r^*(t)) \beta'(p^*(t)) \pi(t) dx \\
& + q \int_{\Omega} \eta^* [\beta'(p^*(t)) \pi(t) \times \nabla \varphi(x - r^*(t))] dx \\
& - q \int_{\Gamma} u^* \varphi(x - r^*(t)) (n \times \beta'(p^*(t)) \pi(t)) d\zeta \\
& - q^2 \int_{\Omega} \varphi(x - r^*(t)) v'(p^*(t)) \Phi(t) dx + v'(p^*(t)) \varrho(t)
\end{aligned} \tag{6.14}$$

a.e. in $]0, T[$

$$\pi(T) = 0$$

$$\begin{aligned}
-\dot{\varrho}(t) = & -q \int_{\Omega} \eta^* [\beta(p^*(t)) \times \nabla^2 \varphi(x - r^*(t)) \pi(t)] dx \\
& + q \int_{\Gamma} u^* [\nabla \varphi(x - r^*(t)) \cdot \pi(t)] (n \times \beta(p^*(t))) d\varsigma \\
& - q \int_{\Omega} [(E^*(t) + \beta(p^*(t)) \times B^*(t)) \cdot \pi(t)] \nabla \varphi(x - r^*(t)) dx \\
& + q^2 \int_{\Omega} v(p^*(t)) [\nabla \varphi(x - r^*(t)) \cdot \Phi(t)] dx \\
& + \nabla J_1(r^*(t)) - g'(r^*(t))^\top \dot{\mu}(t) \quad \text{a.e. in }]0, T[\\
\varrho(T) = & \nabla J_2(r^*(T))
\end{aligned} \tag{6.15}$$

Jump conditions:

$$\varrho(t_i) - \lim_{\varepsilon \searrow 0} \varrho(t_i - \varepsilon) = g'(r^*(t_i))^\top (\lim_{\varepsilon \searrow 0} \mu(t_i - \varepsilon) - \mu(t_i)), \quad i = 1, \dots, \ell, \quad (6.16)$$

Adjoint Poisson equation:

$$\begin{aligned} -\Delta\chi(x, t) &= (\nabla\varphi(x - r^*(t)) \times \beta(p^*(t))) \cdot \pi(t) & f.a.a. (x, t) \in]0, T[\times \Omega \\ \chi(x, t) &= 0 & f.a.a. (x, t) \in]0, T[\times \Gamma \end{aligned}$$

Gradient equation:

$$u^*(x) = \frac{q}{\alpha} \int_0^T \left[\partial_n \chi(x, t) + [\varphi(x - r^*(t))\beta(p^*(t)) \times n] \cdot \pi(t) \right] dt \quad a.e. \text{ on } \Gamma$$

Complementary relations:

$$\mu_j \text{ monotone increasing, } \int_0^T g_j(r^*(t)) d\mu_j(t) = 0, \quad g_j(r^*(t)) \leq 0 \quad \forall t \in [0, T] \\ \text{for all } j = 1, \dots, m$$

Remark 6.7. As a function of bounded variation, μ can be decomposed as

$$\mu = \mu_a + \mu_d + \mu_s,$$

where $\mu_a \in AC([0, T]; \mathbb{R}^m)$ is absolutely continuous and $\mu_d \in L^\infty([0, T]; \mathbb{R}^m)$ is a step function covering the discontinuities of μ . Moreover, $\mu_s \in C([0, T]; \mathbb{R}^m)$ is the singular part, which is non-constant and whose derivative vanishes almost everywhere. Consequently, μ in (6.15) can be replaced by μ_a , while (6.16) holds also with μ_d instead of μ .

Remark 6.8. By integration by parts one can formally derive a strong formulation of the adjoint Maxwell equations in Theorem 6.6:

$$-\frac{\partial}{\partial t} \Phi(x, t) + \text{curl } \Psi(x, t) = \varphi(x - r^*(t))\pi(t) \quad \text{in } \Omega \times [0, T] \quad (6.17)$$

$$-\frac{\partial}{\partial t} \Psi(x, t) - \text{curl } \Phi(x, t) = \pi(t) \times \varphi(x - r^*(t))\beta(p^*(t)) \quad \text{in } \Omega \times [0, T] \quad (6.18)$$

$$\text{div} \left(\frac{\partial}{\partial t} \Phi(x, t) \right) = -\text{div} (\varphi(x - r^*(t))\pi(t)) \quad \text{in } \Omega \times [0, T] \quad (6.19)$$

$$\text{div} \left(\frac{\partial}{\partial t} \Psi(x, t) \right) = -\text{div} (\pi(t) \times \varphi(x - r^*(t))\beta(p^*(t))) \quad \text{in } \Omega \times [0, T] \quad (6.20)$$

$$\Phi(\varsigma, t) \times n = 0, \quad \frac{\partial}{\partial t} \Psi(\varsigma, t) \cdot n = -\pi(t) \times \varphi(\varsigma - r^*(t))\beta(p^*(t)) \cdot n \quad \text{in } \Gamma \times [0, T] \quad (6.21)$$

$$\Phi(x, T) = 0, \quad \Psi(x, T) = 0 \quad \text{in } \Omega. \quad (6.22)$$

Note that the right hand side in (6.17)–(6.18) does, in general, not satisfy a conservation of charge, which gives rise to non-standard equations in (6.19) and (6.20) and the unusual boundary condition in (6.21).

7 Numerical investigations

In the following we illustrate by means of a representative example that the optimal control problem (P) can be treated numerically. We follow the analytical approach and use the reduced state system of Definition 3.12 for our numerical investigations. After a brief description of the numerical method we will present some exemplary results.

7.1. Discretization of the state system. We start the description of the numerical method with the discretization of the state system. Inspired from the analytical treatment of Maxwell's equations by means of semigroup theory, we approximate the solution of Maxwell's equations with the help of their fundamental solution, i.e., the semigroup arising if $\Omega = \mathbb{R}^3$. We thus neglect the influence of any boundary conditions. In case of a single point charge, i.e., charge and current as in (2.2) and (2.3), this fundamental solution allows an explicit representation of the arising electromagnetic fields, the so called Liénard-Wiechert fields, cf. e.g. [27, 42]:

$$E(x, t) = \frac{q}{4\pi\epsilon} \frac{(1 - |\beta(p(t_{ret}))|_2^2)}{|R_v(t_{ret}, p(t_{ret}))|_2^3} R_v(t_{ret}, p(t_{ret})) \\ + \frac{q}{4\pi\epsilon c^2 |R_v(t_{ret}, p(t_{ret}))|_2^3} R(t_{ret}) \times (R_v(t_{ret}, p(t_{ret})) \times \dot{v}(p(t_{ret}))) \quad (7.1)$$

$$B(x, t) = c\epsilon\mu \frac{R(t_{ret})}{|R(t_{ret})|_2} \times E(x, t) \quad (7.2)$$

with

$$R(t) := x - r(t), \quad t_{ret} := t_{ret}(x, t) = t - \frac{R(t_{ret})}{c}, \\ R_v(t, p) := (R(t) - \beta(p)R(t)).$$

For the numerical realization these expressions are further simplified. Firstly, we neglect the difference between t and t_{ret} . Moreover, we leave out the terms arising from an acceleration of the charge, i.e., the second addend on the right hand side of (7.1). In contrast to the first addend which is of order $1/R^2$, this term grows with $1/R$ and thus models the far field, whose influence on the movement of the particles can be neglected, see [28].

The Poisson equation in (3.17) is discretized by means of finite elements. We use a uniform hexahedral mesh and piecewise trilinear and continuous ansatz functions for both, solution and test function, which represents a variational crime due to the low regularity of the very weak solution. A priori error analysis for this procedure can be found in [4]. The linear system of equations arising by this discretization is solved by the CG method preconditioned via an incomplete LU decomposition of the stiffness matrix.

Finally, the relativistic Newton-Lorentz equations (3.7) are solved numerically by the so called Boris scheme, a second-order time stepping scheme especially tailored to this type of equations of motions, described in [5, 7]. It is frequently used in plasma physics and especially for particle accelerators (as part of particle-in-cell methods), since it is an explicit and energy conserving scheme. The physical quantities and constants involved in (3.7) differ by several orders of magnitude, cf. Table 7.1 below. In order to avoid numerical cancellation effects, we introduce a nondimensionalization factor in the Newton-Lorentz equations. In addition, cancellation also occurs in the numerical evaluation of the integrals involving φ in (3.7a). This is due to the small support of φ , whose diameter amounts 10^{-6} and causes larger slopes of φ due to the normalization in (3.1). To circumvent these problems, we use a linear transformation to enlarge the support. The transformed integrals are approximated by the Simpson rule weighted with $\varphi(\cdot - r)$ and $\nabla\varphi(\cdot - r) \times \beta(p)$, respectively.

7.2. Optimization algorithm. To keep the model physically meaningful it is of major importance to fulfill the pointwise state constraint in (3.9), see Remark 3.1. This is guaranteed by a purely primal interior point approach in form of a log-barrier method, see e.g. [36, Chapter 19]. In [40, 41] this method has been proven

to work in function space for one dimensional problems, i.e., problems involving ODEs as in our case. The reduction of the homotopy parameter associated with the primal interior point method follows an update strategy by [36, Section 19.3].

For the optimization algorithm we reduce the optimal control problem to an optimization problem in the control variable u only, which is justified by Theorem 4.3. The major advantage of this procedure is a significant reduction of the number of optimization variables, since the control u is only one dimensional. It does not depend on time, and has its support on Γ instead of the whole domain Ω . The dimension of the optimization problem reduced to the control variable, thus, amounts to the number of nodes on the boundary only. This allows to employ optimization methods, which require large memory demand like the BFGS method, see e.g. [36, Section 6.1]. Thanks to the reduction of the dimension the BFGS method can be run for a moderate number of degrees of freedom on a computer with 4GB RAM without any limited memory modification. In order to globalize the method, we perform a curvature test to switch from the BFGS direction to the negative gradient of the objective, if necessary, and apply a line-search according to the Armijo rule.

As a consequence of this reduction approach the mapping $u \mapsto \mathcal{J}(r(u), u)$ as well as its derivative have to be evaluated in every iteration of the optimization algorithm. Here $r(u)$ denotes the r -component of the solution of state system associated with u . The derivative of $u \mapsto \mathcal{J}(r(u), u)$ is computed numerically by means of the automatic differentiation tool ADiMat [6]. As the number of control variables is much higher than the number of output variables, which is just a real number, we use the reverse mode. Moreover, we exclude the linear parts of the solution mapping of the state system from automatic differentiation to differentiate them by hand. This especially concerns the iterative solver of Poisson's equation. To summarize we thus follow a first-discretize-then optimize approach. It is not clear whether the discrete adjoint equation arising in this way can be interpreted as a suitable discretization of the adjoint system in Theorem 6.6. In particular, the adjoint Boris scheme gives rise to future research with regard to its stability and consistency.

7.3. Test setting. For the numerical realization we chose an electron as particle. The mass at rest and the charge are chosen appropriately, see Table 7.1. The computational domain Ω is a cube of size length $2 \cdot 10^{-3}$ m. For the subdomain $\tilde{\Omega}$ arising in the state constraint (3.9) we chose an inner cube of size length $2 \cdot 10^{-4}$ m. As the electron is almost moving with the speed of light, the end time was set to $T = 2 \cdot 10^{-10}$ s.

Quantity	Symbol	Value (in SI units)
speed of light (in vacuum)	c	$2.9979 \cdot 10^8 \text{ m s}^{-1}$
permittivity of free space	ϵ	$8.8541 \cdot 10^{-12} \text{ F m}^{-1}$
permeability of free space	μ	$4\pi \cdot 10^{-7} \text{ H m}^{-1}$
electron rest mass	m_0^q	$9.1093 \cdot 10^{-31} \text{ kg}$
electric charge	q	$1.6021 \cdot 10^{-19} \text{ C}$

TABLE 7.1. Physical constants.

For the numerical computations we focus on optimizing the particle position at end time, i.e., we choose

$$J_1(r) \equiv 0, \quad J_2(r) = \frac{1}{2} |r - r_d|_2^2$$

for the contributions to the objective in (\tilde{P}) and (P) , respectively. Furthermore, the Tikhonov parameter α in the objective is set to $\alpha = 10^{-9}$ to compensate for the comparatively large values of the control. Consequently we are mainly interested in steering the particle beam at a given end time to a fixed position r_d . As a stopping criterion for the overall algorithm we check if the relative error between the desired particle position r_d and the computed one is below a given tolerance.

For the computations presented in the following section, we used an equidistant mesh with 17,576 nodes. This amounts to 7,504 nodes on the boundary, i.e., the number of unknown control variables, which corresponds to the dimension of the optimization problem. For the numerical integration of the ODE we used an equidistant time step of 10^{-12} s.

7.4. Numerical results. The particle trajectories for selected iterations of the optimization algorithms are shown in Figures 7.2 to 7.6. While the particle is colored in black, we marked the desired end position in the upper left corner in grey. It is to be noted that the control u only influences the magnetic field, which in turn cannot slow down or accelerate this particle beam since its contribution to the Lorentz force only acts perpendicular to the direction of motion, cf. (2.6a). This causes spiral shaped trajectories such as the ones depicted in figures 7.1 to 7.6. The desired end position has been reached after 47 iterations of the optimization algorithm with an accuracy of $3.5 \cdot 10^{-8}$ m.

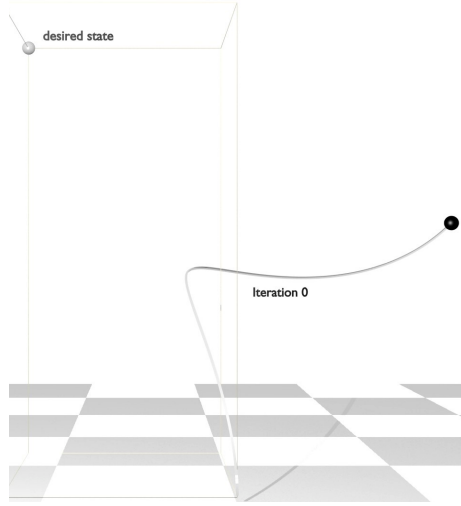


FIGURE 7.1. Particle trajectory in iteration 0.

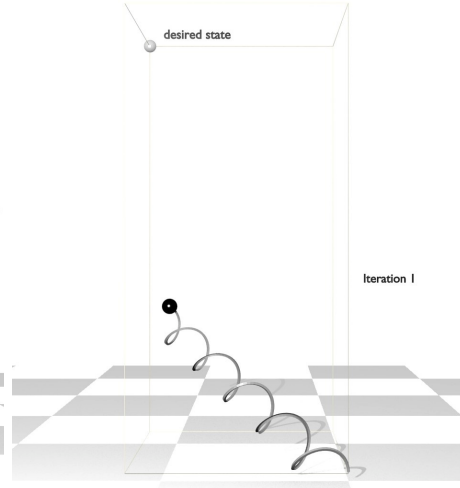


FIGURE 7.2. Particle trajectory in iteration 1.

The optimal external magnetic field on the boundary of Ω generated by the optimal control u^* is shown in Figures 7.7 and 7.8.

Table 7.2 shows the convergence history of the globalized BFGS interior point method. Beside the objective value and Euclidean norm of the gradient, Table 7.2 shows the used descent direction for selected iterations, where “BFGS” refers to the BFGS direction and “Grad” is the negative gradient.

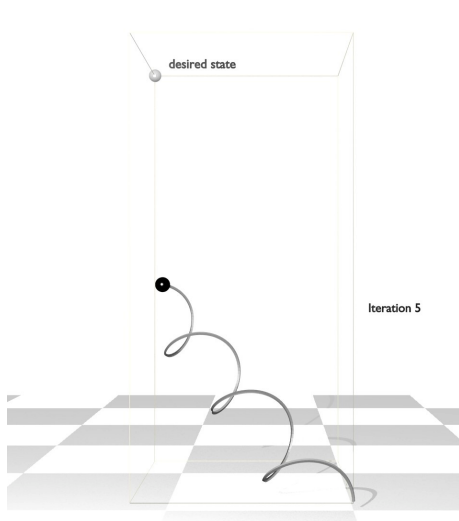


FIGURE 7.3. Particle trajectory in iteration 5.

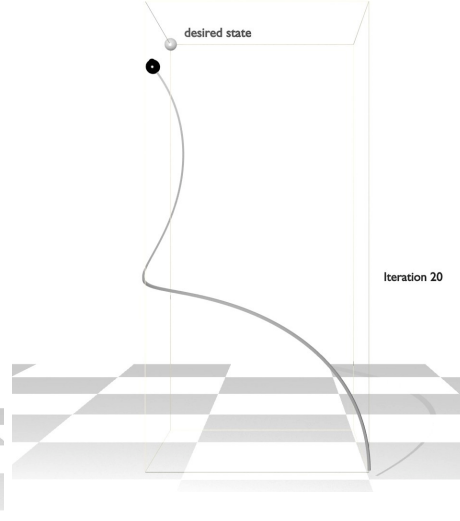


FIGURE 7.4. Particle trajectory in iteration 20.

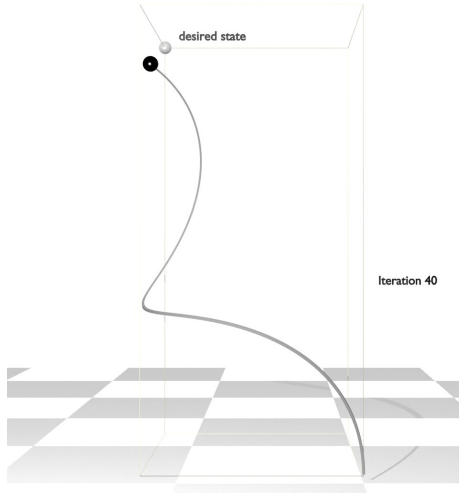


FIGURE 7.5. Particle trajectory in iteration 40.

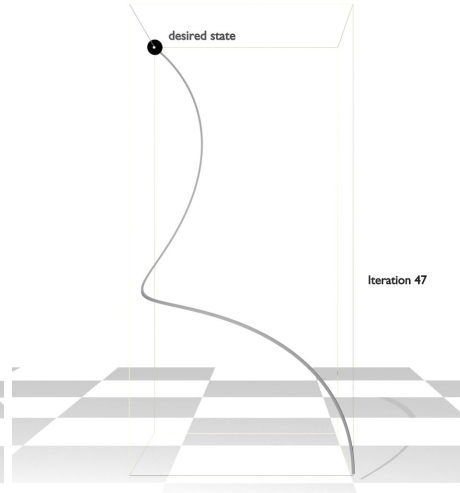


FIGURE 7.6. Particle trajectory in iteration 47.

A Proof of Theorem 4.3

Clearly, y solves (4.1) if and only if it is a fixed point of

$$G : C([0, T]; \mathbb{R}^3)^2 \rightarrow C([0, T]; \mathbb{R}^3)^2, \quad G(y)(t) := \int_0^t f(y)(\tau) d\tau.$$

We show that G is contractive, if we equip the set of continuous functions with the following equivalent norm

$$\|y\|_G := \max_{t \in [0, T]} e^{-Lt} |y(t)|_2.$$

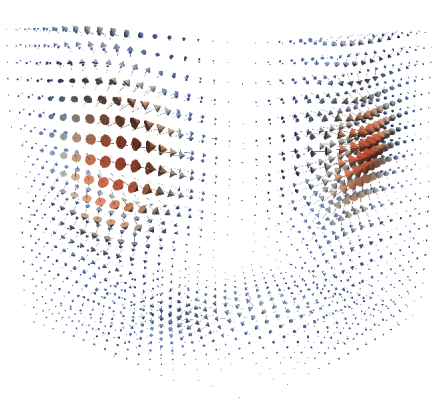


FIGURE 7.7. Front view of external magnetic field.

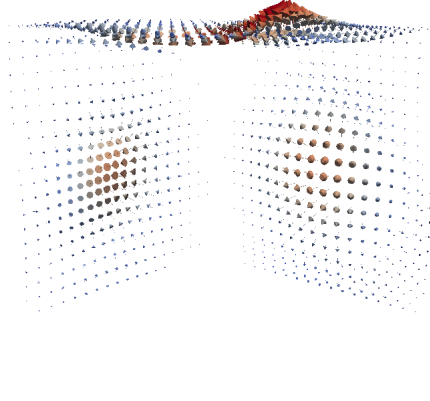


FIGURE 7.8. Back view of external magnetic field.

Iteration	f-value	gradient	step	Iteration	f-value	gradient	step
0	0.3835	-	-	35	0.0060	1.7E-5	BFGS
1	0.1509	0.0024	Grad	40	1.6E-5	1.7E-5	Grad
5	0.0748	6.0E-4	BFGS	42	9.9E-8	6.0E-6	BFGS
10	0.0091	2.6E-4	BFGS	44	4.0E-8	2.7E-7	BFGS
20	0.0086	2.3E-4	Grad	46	2.5E-9	1.4E-8	BFGS
25	0.0064	6.0E-5	BFGS	47	6.1E-10	5.4E-9	BFGS
30	0.0061	3.9E-5	BFGS				

TABLE 7.2. Convergence history of the optimization algorithm.

To this end, observe that for every $v \in C([0, t]; \mathbb{R}^3)^2$ and every $\tau \in [0, T]$ there holds

$$\|v\|_{C([0, \tau]; \mathbb{R}^3)^2} \leq \max_{s \in [0, \tau]} e^{Ls} \max_{s \in [0, \tau]} e^{-Ls} |v(s)|_2 \leq e^{L\tau} \|v\|_G.$$

Then we obtain by means of Lemma 4.1

$$\begin{aligned} \|G(y) - G(v)\|_G &\leq \max_{t \in [0, T]} e^{-Lt} \int_0^t |f(y)(\tau) - f(v)(\tau)| d\tau \\ &\leq L \max_{t \in [0, T]} e^{-Lt} \int_0^t \|y - v\|_{C([0, \tau]; \mathbb{R}^3)^2} d\tau \\ &\leq L \max_{t \in [0, T]} e^{-Lt} \int_0^t e^{L\tau} d\tau \|y - v\|_G \leq (1 - e^{-LT}) \|y - v\|_G, \end{aligned}$$

i.e., the desired contractivity of G . Thus Banach's fixed point theorem gives the existence of a unique solution to (4.1) as claimed.

To prove the a priori estimate we again abbreviate $(r, p) := y + (r_0, p_0)$. Then (4.3) implies for an arbitrary $t \in [0, T]$ that

$$|f_2(y)(t)| \leq c.$$

Beside (4.4), the conditions on φ in (3.1) clearly give that for every $r \in \mathbb{R}^3$

$$\|\nabla \varphi(x - r)\|_X \leq \sqrt{|\Omega|} \max_{x \in \mathbb{R}^3} |\nabla \varphi(x)| < \infty, \quad \|\varphi(x - r)\|_{L^2(\Gamma)} \leq C_\varphi \sqrt{|\Gamma|} < \infty.$$

Thus, (3.15), (4.3), (4.4), and the definition of β in (2.5) give

$$\begin{aligned} |f_1(y)(t)| &\leq q \left(\|\varphi(\cdot - r(t))\|_{L^2(\Omega)} \|F_L(r, p)(t)\|_X \right. \\ &\quad + \|(-\Delta^*)^{-1} R\|_{\mathcal{L}(L^2(\Gamma), L^2(\Omega))} \|u\|_{L^2(\Gamma)} \|\nabla \varphi(\cdot - r(t))\|_{L^2(\Omega)} |\beta(p(t))|_2 \\ &\quad \left. + \|u\|_{L^2(\Gamma)} \|\varphi(\cdot - r(t))\|_{L^2(\Gamma)} |\beta(p(t))|_2 \right) \\ &\leq C_1 \|u\|_{L^2(\Gamma)} + C \|F_L(r, p)(t)\|_X. \end{aligned}$$

In view of $|\beta(p)|_2 \leq 1$ for all $p \in \mathbb{R}^3$, cf. again (2.5) and (4.3), F_L can be estimated by

$$\|F_L(r, p)(t)\|_X \leq 2Me^{\omega T} \left(\|(E_0, B_0)\|_{X \times X} + \|j(r, p)\|_{L^1([0, T]; X)} \right)$$

with

$$\|j(r, p)\|_{L^1([0, T]; X)} \leq qT \sqrt{C_\varphi} c,$$

see (4.5). Therefore, we arrive at

$$|\dot{y}(t)| = |f(y)(t)| \leq C_1 \|u\|_{L^2(\Gamma)} + C_2 \quad \forall t \in [0, T]$$

with constants $C_1, C_2 > 0$ independent of t , u , and y . As $y(0) = 0$, this gives the desired estimate.

B Proof of Lemma 6.5

By the Riesz representation theorem there exists a unique function $\mu \in \text{NBV}([0, T]; \mathbb{R}^m)$ such that

$$\begin{aligned} \langle \lambda, g'(r^*) \phi_r \rangle_{C([0, T]; \mathbb{R}^m)^*, C([0, T]; \mathbb{R}^m)} &= \int_0^T (g'(r^*(t)) \phi_r(t)) \cdot d\mu(t) \\ &\quad \forall \phi_r \in C([0, T]; \mathbb{R}^3). \end{aligned} \quad (\text{B.1})$$

Moreover, $\lambda \geq 0$ implies that μ is monotonically increasing as claimed. Taking the definition of \mathcal{J} into account we find

$$\left\langle \frac{\partial \mathcal{J}}{\partial r}(r^*, u^*), \phi_r \right\rangle_{Y^*, Y} = \int_0^T J'_1(r^*(t)) \phi_r(t) dt + J'_2(r^*(T)) \phi_r(T).$$

By inserting this together with (B.1) in (6.8) we arrive at

$$\begin{aligned} \int_0^T \dot{\phi}(t) \cdot \omega(t) dt + \int_0^T \phi(t) \cdot \left[A(y^*, u^*, \omega)(t) - \begin{pmatrix} \nabla J_1(r^*(t)) \\ 0 \end{pmatrix} \right] dt \\ - \int_0^T (g'(r^*(t)) \phi_r(t)) \cdot d\mu(t) = 0 \quad \forall \phi \in C_0^\infty([0, T]; \mathbb{R}^3)^2. \end{aligned}$$

In view of (6.7), the continuity of B^* , E^* , and y^* w.r.t. time and $\omega \in L^2([0, T]; \mathbb{R}^3)^2$ implies $A(y^*, u^*, \omega) \in L^2([0, T]; \mathbb{R}^3)^2$. Thus, according to the Du Bois Raymond theorem for Stieltjes integrals, see e.g. [19, Lemma 3.1.9], the equivalence class ω admits a representation as BV-function, denoted by the same symbol for simplicity, which fulfills for all $t \in [0, T]$

$$\begin{aligned} \varrho(t) &= \varrho(T) - \int_t^T [A_r(y^*, u^*, \omega)(\tau) - \nabla J_1(r^*(\tau))] d\tau \\ &\quad + \int_t^T g'(r^*(\tau))^\top d\mu(\tau) \end{aligned} \quad (\text{B.2})$$

$$\pi(t) = \pi(T) - \int_t^T A_p(y^*, u^*, \omega)(\tau) d\tau. \quad (\text{B.3})$$

The later equation immediately implies (6.9) and, since $\pi, \varrho \in \text{BV}([0, T]; \mathbb{R}^3) \hookrightarrow L^\infty([0, T]; \mathbb{R}^3)$, this ODE gives the desired regularity of π .

As a function of bounded variation μ has at most countably many discontinuities and is differentiable almost everywhere in $]0, T[$. Moreover, since μ is in addition monotonically increasing, there holds

$$\frac{d}{dt} \int_t^T g'(r^*(\tau))^\top d\mu(\tau) = -g'(r^*(t))^\top \dot{\mu}(t) \quad \text{f.a.a. } t \in]0, T[$$

see e.g. [19, Lemma 2.1.26]. Thus (B.2) gives (6.11).

Integrating the last integral in (B.2) by parts leads to

$$\begin{aligned} & \varrho(t) + g'(r^*(t)) \cdot \mu(t) \\ &= \varrho(T) + g'(r^*(T)) \cdot \mu(T) \\ & \quad - \int_t^T \left[A_r(y^*, u^*, \omega)(\tau) - \nabla J_1(r^*(\tau)) + \sum_{j=1}^m \mu_j(\tau) g_j''(r^*(\tau)) \dot{r}^*(\tau) \right] d\tau. \end{aligned}$$

As the right hand side is continuous, the discontinuities of ϱ are therefore located at the same points as the ones of μ . Moreover, as μ is of bounded variation, one has

$$\int_t^T g'(r^*(\tau))^\top d\mu(\tau) - \lim_{\varepsilon \searrow 0} \int_{t-\varepsilon}^T g'(r^*(\tau))^\top d\mu(\tau) = g'(r^*(t))^\top \left(\lim_{\varepsilon \searrow 0} \mu(t-\varepsilon) - \mu(t) \right)$$

for every $t \in]0, T]$, cf. e.g. [19, p. 66]. Since $A_r(y^*, u^*, \omega)(\cdot) - \nabla J_1(r^*(\cdot)) \in L^2([0, T]; \mathbb{R}^3)$, (B.2) therefore implies (6.13).

Integrating the first integral in (6.8) by parts yields

$$\begin{aligned} & - \int_0^T \phi(t) \cdot d\omega(t) + \int_0^T \phi(t) \cdot \left[A(y^*, u^*, \omega)(t) - \begin{pmatrix} \nabla J_1(r^*(t)) \\ 0 \end{pmatrix} \right] dt \\ & - \int_0^T (g'(r^*(t)) \phi_r(t)) \cdot d\mu(t) = J_2'(r^*(T)) \phi_r(T) - \phi(T) \cdot \omega(T) \quad \forall \phi \in Y. \end{aligned} \tag{B.4}$$

The continuity of $g'(r^*(\cdot))$ gives that

$$\nu(t) = \int_t^T g'(r^*(\tau))^\top d\mu(\tau)$$

is of bounded variation. Since ϕ_r also continuous, we arrive at

$$\int_0^T (g'(r^*(t)) \phi_r(t)) \cdot d\mu(t) = - \int_0^T \phi_r(t)^\top d\nu(t),$$

cf. e.g. [19, p. 67]. Hence, thanks to (B.2) and (B.3), (B.4) gives $\omega(T) \cdot \phi(T) = J_2'(r^*(T)) \phi_r(T)$ for all $\phi \in Y$, which in turn yields the desired final time conditions in (6.10) and (6.12).

References

- [1] W Ackermann and et al. Operation of a free-electron laser from the extreme ultraviolet to the water window. *Nature Photonics*, 1(6):336–342, June 2007.
- [2] A. Alonso and A. Valli. Some remarks on the characterization of the space of tangential traces of $\mathbf{H}(\text{rot}; \omega)$ and the construction of an extension operator. *manuscripta mathematica*, 89(1):159–178, 1996.
- [3] G. Bauer, D.-A. Deckert, and D. Dürr. Maxwell-Lorentz dynamics of rigid charges. *Communications in Partial Differential Equations*, 38(9):1519–1538, 2013.

- [4] M. Berggren. Approximations of very weak solutions to boundary-value problems. *Siam J. Numer. Anal.*, 42(2):860–877, 2004.
- [5] C.K. Birdsall and A.B. Langdon. *Plasma Physics via Computer Simulation*. Series in Plasma Physics. Taylor & Francis, 2004.
- [6] C.H. Bischof, H.M. Bücker, B. Lang, A. Rasch, and A. Vehreschild. Combining source transformation and operator overloading techniques to compute derivatives for MATLAB programs. In *Proceedings of the Second IEEE International Workshop on Source Code Analysis and Manipulation (SCAM 2002)*, pages 65–72, Los Alamitos, CA, USA, 2002. IEEE Computer Society.
- [7] J.P. Boris. Relativistic plasma simulation - optimization of a hybrid code. In *Proc. Fourth Conf. Numerical Simulation of Plasmas*, pages 3–67, 1970.
- [8] H.P. Cartan. *Differential Calculus*. Hermann, 1983.
- [9] E. Casas. Control of an elliptic problem with pointwise state constraints. *SIAM Journal on Control and Optimization*, 24(6):1309–1318, 1986.
- [10] E. Casas. Boundary control of semilinear elliptic equations with pointwise state constraints. *SIAM Journal on Control and Optimization*, 31(4):993–1006, 1993.
- [11] E. Casas and J. Raymond. Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations. *SIAM Journal on Control and Optimization*, 45(5):1586–1611, 2006.
- [12] M. Dauge. *Elliptic boundary value problems on corner domains : smoothness and asymptotics of solutions*. Lecture notes in mathematics. Springer, Berlin, 1988.
- [13] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for science and technology*, volume 5. Spinger, Berlin, 1992.
- [14] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for science and technology*, volume 3. Spinger, Berlin, 1992.
- [15] K. Deckelnick, A. Günther, and M. Hinze. Finite element approximation of Dirichlet boundary control for elliptic pdes on two- and three-dimensional curved domains. *SIAM Journal on Control and Optimization*, 48(4):2798–2819, 2009.
- [16] P. Druet, O. Klein, J. Sprekels, F. Tröltzsch, and I. Yousept. Optimal control of three-dimensional state-constrained induction heating problems with nonlocal radiation effects. *SIAM Journal on Control and Optimization*, 49(4):1707–1736, 2011.
- [17] M. Falconi. Global solution of the electromagnetic field-particle system of equations. *Journal of Mathematical Physics*, 55(10):1–12, 2014.
- [18] CGR Geddes, C Toth, J van Tilborg, E Esarey, C B Schroeder, D Bruhwiler, C Nieter, J Cary, and W P Leemans. High-quality electron beams from a laser wakefield accelerator using plasma-channel guiding. *Nature*, 431(7008):538–541, 2004.
- [19] M. Gerds. *Optimal control of ODEs and DAEs*. De Gruyter, Berlin, 2012.
- [20] V. Girault and P.-A. Raviart. *Finite Element Methods for Navier-Stokes Equations*. Springer, Berlin, 1986.
- [21] E. Gjonaj, T. Lau, S. Schnepf, F. Wolfheimer, and T. Weiland. Accurate modelling of charged particle beams in linear accelerators. *New Journal of Physics*, 8:1–21, August 2006.
- [22] R. Griesse and K. Kunisch. Optimal control for a stationary MHD system in velocity-current formulation. *SIAM Journal on Control and Optimization*, 45(5):1822–1845, 2006.

- [23] R. F. Hartl, S. P. Sethi, and R. G. Vickson. A survey of the maximum principles for optimal control problems with state constraints. *SIAM Review*, 37(2):181–218, 1995.
- [24] M. Hintermüller and K. Kunisch. Feasible and non-interior path-following in constrained minimization with low multiplier regularity. *SIAM J. Control and Optim.*, 45:1198–1221, 2006.
- [25] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, New York, 2009.
- [26] V. Imaikin, A. Komech, and N. Mauser. Soliton-type asymptotics for the coupled Maxwell-Lorentz equations. *Annales Henri Poincaré*, 5:1117–1135, 2004.
- [27] J. D. Jackson. *Classical electrodynamics*. Wiley, New York, 3rd edition, 1999.
- [28] A. Komech and H. Spohn. Long-time asymptotics for the coupled Maxwell-Lorentz equations. *Communications in Partial Differential Equations*, 25:559–584, 2000.
- [29] K. Kunisch and B. Vexler. Constrained Dirichlet boundary control in L^2 for a class of evolution equations. *SIAM Journal on Control and Optimization*, 46(5):1726–1753, 2007.
- [30] Z Li, V Akcelik, A Candel, S Chen, L Ge, A Kabel, L Q Lee, C Ng, E Prudencio, G Schussman, R Uplenchwar, L Xiao, and K Ko. Towards simulation of electromagnetics and beam physics at the petascale. In *2007 IEEE Particle Accelerator Conference (PAC)*, pages 889–893. IEEE, 2007.
- [31] S. May, R. Rannacher, and B. Vexler. Error analysis for a finite element approximation of elliptic Dirichlet boundary control problems. *SIAM Journal on Control and Optimization*, 51(3):2585–2611, 2013.
- [32] C. Meyer, A. Rösch, and F. Tröltzsch. Optimal control of pdes with regularized pointwise state constraints. *Comp. Optim. Appl.*, 33:209–228, 2006.
- [33] S. Nicaise, S. Stingelin, and F. Tröltzsch. On two optimal control problems for magnetic fields. *Computational Methods in Applied Mathematics*, 14:555–573, 2014.
- [34] S. Nicaise, S. Stingelin, and F. Tröltzsch. Optimal control of magnetic fields in flow measurement. *Discrete and Continuous Dynamical Systems-S*, 8:579–605, 2015.
- [35] S. Nicaise and F. Tröltzsch. A coupled Maxwell integrodifferential model for magnetization processes. *Mathematische Nachrichten*, 287(4):432–452, 2014.
- [36] J. Nocedal and S. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, 2006.
- [37] G. Of, T. X. Phan, and O. Steinbach. Boundary element methods for Dirichlet boundary control problems. *Mathematical Methods in the Applied Sciences*, 33(18):2187–2205, 2010.
- [38] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44. Springer, New York, 1st edition, 1983.
- [39] J. Rossbach and P. Schmüser. Basic course on accelerator optics. In S. Turner, editor, *CAS - CERN Accelerator School: 5th General accelerator physics course*, volume I, 1992.
- [40] A. Schiela. Barrier methods for optimal control problems with state constraints. *Siam J. Optim.*, 20(2):1002–1031, 2009.
- [41] A. Schiela. An interior point method in function space for the efficient solution of state constrained optimal control problems. *Mathematical Programming*, 138:83–114, 2013.

- [42] H. Spohn. *Dynamics of charged particles and their radiation field*. Cambridge University Press, Cambridge, 2004.
- [43] R. Temam. *Navier-Stokes equations: theory and numerical analysis*. North-Holland, Amsterdam, 1977.
- [44] F. Tröltzsch and I. Yousept. PDE-constrained optimization of time-dependent 3D electromagnetic induction heating by alternating voltages. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46:709–729, 7 2012.
- [45] K. Wille. *The Physics of Particle Accelerators*. Oxford University Press, 2000.
- [46] I. Yousept. Optimal control of a nonlinear coupled electromagnetic induction heating system with pointwise state constraints. *Ann. Acad. Rom. Sci. Ser. Math. Appl.*, 2(1):45–77, 2010.
- [47] I. Yousept. Finite element analysis of an optimal control problem in the coefficients of time-harmonic eddy current equations. *J. Optim. Theory Appl.*, 154:879–903, 2012.
- [48] I. Yousept. Optimal control of Maxwell’s equations with regularized state constraints. *Computational Optimization and Applications*, 52:559–581, 2012.
- [49] I. Yousept. Optimal bilinear control of eddy current equations with grad-div regularization and divergence penalization. *Journal of Numerical Mathematics*, accepted for publication, 2013.
- [50] I. Yousept. Optimal control of quasilinear $\mathbf{H}(\text{curl})$ -elliptic partial differential equations in magnetostatic field problems. *SIAM Journal on Control and Optimization*, 51(5):3624–3651, 2013.
- [51] J. Zowe and S. Kurcyusz. Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics and Optimization*, 5(1):49–62, 1979.

(C. Meyer) TU DORTMUND, FACULTY OF MATHEMATICS, VOGELPOTHSWEG 87, 44227 DORTMUND, GERMANY.

E-mail address: cmeyer@math.tu-dortmund.de

(S. M. Schnepp) INSTITUTE OF GEOPHYSICS, DEPARTMENT OF EARTH SCIENCES, ETH ZURICH, SONNEGSTRASSE 5, CH-8092 ZURICH, SWITZERLAND.

E-mail address: schnepps@ethz.ch

(O. Thoma) TU DORTMUND, FACULTY OF MATHEMATICS, VOGELPOTHSWEG 87, 44227 DORTMUND, GERMANY.

E-mail address: othoma@math.tu-dortmund.de