On Some Optimal Control Problems for Electric Circuits

Kristof Altmann, Simon Stingelin, and Fredi Tröltzsch

October 12, 2010

Abstract

Some optimal control problems for linear and nonlinear ordinary differential equations related to the optimal switching between different magnetic fields are considered. The main aim is to move an electrical initial current by a controllable voltage in shortest time to a desired terminal current and to hold it afterwards. Necessary optimality conditions are derived by Pontryagin's principle and a Lagrange technique. In the case of a linear system, the principal structure of time-optimal controls is discussed. The associated optimality systems are solved by a one-shot strategy using a multigrid software package. Various numerical examples are discussed.

1 Introduction

We consider a class of optimal control problems governed by linear and nonlinear systems of ordinary differential equations (ODEs) with some relation to eddy current problems in continuum mechanics. The background of our research in real applications is the magnetization of certain 3D objects. More precisely, it is the optimal switching between different magnetization fields that stands behind our investigations.

Problems of this type occur in measurement sensors. If the measurement signal is much smaller than the noise, one uses differential measurement techniques, where one has to switch from one state to an other. In this paper, the signal source would be an electromagnetic field. In practice, the magnetization process is performed by the control of certain induction coils, and the associated model is described by the Maxwell equations. The numerical solution of these partial differential equations is demanding and the treatment of associated optimal control problems is quite challenging. To get a first impression of the optimal control functions to be expected in such processes, we study here simplified models for electrical circuits based on ordinary differential equations. They resemble the behavior that is to be expected also from 3D models with Maxwell's equations.

We shall study different types of electrical circuits and associated optimal control problems. In all problems, we admit box constraints to bound the control functions. Moreover, we also consider an electrical circuit with a nonlinear induction function.

Mathematically, our main goal is twofold. From an application point of view, we are interested in the form of optimal controls for different types of electrical circuits. We discuss several linear and nonlinear models with increasing level of difficulty. On the other hand, we present an optimization method that is, to our best knowledge, not yet common in the whole optimal control community. By a commercial code, we directly solve nonsmooth optimality systems consisting of the state equation, the adjoint equation and a projection formula for the control. This method was already used by several authors. In particular, it was suggested by Neitzel et al. [10] for optimal control problems governed by partial differential equations. This approach works very reliably also for our problems.

We derive the optimality systems in two different ways. First, we invoke Pontryagin's principle to obtain the necessary information. While this approach is certainly leading to the most detailed information on the optimal solutions, is may be difficult to apply in some cases. Therefore, we later establish the optimality conditions also by the well known Lagrange principle of nonlinear optimization in function spaces, see e.g. Ioffe and Tikhomirov [6]. This approach is easier to apply and turns out to be completely sufficient for numerical purposes.

The structure of the paper is as follows: Each model is discussed in a separate section. Here, we motivate the used ordinary differential equation, then we define the related optimal control problem and show its well-posedness. Next, we derive necessary optimality conditions. Finally, we suggest a numerical method and show its efficiency by numerical tests.

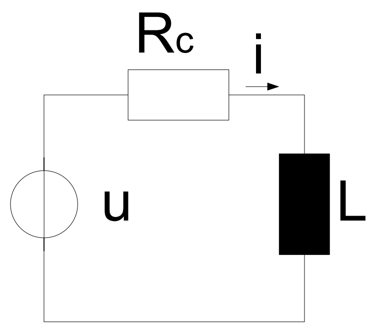


Figure 1: One-loop circuit, model 1

2 Numerical analysis of a one-loop electrical circuit

2.1 The optimal control problem

Let us start with a first simple electrical circuit to illustrate how the optimization theory is developed up to the numerical solution. Magnetic fields are generated and coupled with electric fields. In our simplified setting, we model this coupling by electrical circuits. We begin with a model described by the equation

$$\begin{aligned} L\,\frac{di}{dt}(t) + R_c\, i(t) &= u(t) \\ i(0) &= -i_0 \end{aligned} \tag{1}$$

for an electric circuit, where the *control* u is a controllable voltage, i is the current associated with u, the *state* of the control system, and $-i_0 \in \mathbb{R}$ is its initial value. Moreover, L is the inductance and R_c the resistance. We assume that i_0, L, and R_c are positive real numbers. The situation is shown in Figure 1. Let us mention that, throughout the paper, we use unit quantities to avoid the introduction of units.

Suppose that we want to switch the current i as fast as possible from $-i_0$ to i_0 and to hold the new state i_0 afterwards. This aim is expressed by the cost functional

$$J(i,u) = \frac{\lambda_T}{2}(i(T)-i_0)^2 + \frac{\lambda_Q}{2}\int_0^T (i(t)-i_0)^2\, dt + \frac{\lambda_U}{2}\int_0^T u(t)^2\, dt \tag{2}$$

with positive constants λ_T, λ_Q, and λ_U and a fixed final time $T > 0$. Moreover, the control has to obey the constraints

$$u \in U_{ad} = \{u \in L^2(0,T) \,|\, u_a \le u(t) \le u_b \text{ for a.a. } t \in (0,T)\}, \tag{3}$$

where $u_a < u_b$ are fixed real numbers.

This is a fixed-time optimal control problem, although we have time optimality in mind. The main reason for considering this modification is the following: In standard time optimal control problems, a final target state has to be reached exactly. This is a meaningful problem for the optimal control of ODEs. Here, the state space is finite-dimensional so that the reachability of states is a standard assumption. We discuss this type problems in Section 3.3.

The main goal of our paper is the study of effects that might also occur in the control of 3D magnetic fields, where the physical background has some similarities with the behavior of electric circuits. Since magnetic fields are modeled by the Maxwell equations, an infinite-dimensional state space occurs, where the exact reachability of states is a fairly strong assumption. Therefore, we prefer to approximate the target state closely in a short time rather than to reach it. This is expressed in the setting above.

We search for a couple (i^*, u^*), which minimizes this functional subject to the initial-value problem (1), and the constraints (3). The state function i is considered as a function of $H^1(0,T)$. The functions i^* and u^* are called *optimal state* and *optimal control*, respectively.

The cost functional will be minimal, if the current $i^*(T)$ is close to i_0 (first term in J) and reaches this goal quite fast (second term), while the cost of u is considered by the third term. The weights λ_T, λ_Q, and λ_U can be freely chosen to balance the importance of these 3 terms. We should underline that we need a positive regularization parameter λ_U for the theory and our numerical approach. First of all, this parameter is needed to set up our optimality systems in the form we need for our numerical method. For instance, we solve the optimality systems (15) or (16), where λ_U must not vanish. To have a convex functional, $\lambda_U > 0$ is needed. Moreover, it is known that other numerical optimization methods (say gradient methods, e.g.) behave less stable without regularisation parameter.

Since the state equation is linear, the objective functional is strictly convex, and the set of admissible controls is nonempty and bounded in $L^2(0,T)$, the existence of a unique optimal control is a standard conclusion. We shall indicate optimality by a star, i.e., the optimal control is u^*, while the associated optimal state is denoted by i^*.

2.2 The Pontryagin maximum principle

Let us derive the associated first-order necessary optimality conditions. There are at least two approaches. The first is based on the theory of nonlinear programming in function spaces and uses the well-known concept of Lagrange multipliers associated with the different types of constraints. To apply this technique, a Lagrange function has to be set up and then the necessary conditions are deduced.

On the other hand, the celebrated Pontryagin maximum principle is a standard approach to tackle this problem. It contains the most complete information on the necessary optimality conditions. Therefore, we will apply this principle and transform the associated results in a way that is compatible with the first approach. This is needed to interprete the adjoint state as the Lagrange multiplier associated with the state equation (1). In later sections, we prefer to apply the Lagrange principle.

Let us first recall the Pontryagin principle, formulated for the following class of optimal control problems with objective functional of Bolza type:

$$\min\ J(u) := g(x(T)) + \textstyle\int_0^T f_0(x,u)\,dt.$$

subject to

$$\begin{aligned} \dot{x}_k &= f_k(x,u) && \text{for } k = 1,\ldots,n, \\ x(0) &= x^0, && \\ u(t) &\in \Omega && \text{for a.a. } t \in (0,T), \end{aligned} \tag{4}$$

where $f_k : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $k = 0,\ldots,n$, and $g : \mathbb{R}^n \to \mathbb{R}$ are sufficiently smooth functions, 0 is the fixed initial time, x^0 is a fixed initial vector and $\Omega \subset \mathbb{R}^m$ is a nonempty convex set of admissible control vectors; $T > 0$ is a fixed final time.

Associated with this problem, we introduce the dynamical cost variable x_0 by

$$\dot{x}_0(t) = f_0(x(t), u(t)), \quad x_0(0) = 0,$$

and the *Hamilton function* $H : \mathbb{R}^{n+1} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ by

$$H(\psi, x, u) := \sum_{k=0}^{n} \psi_k\, f_k(x,u). \tag{5}$$

For $n = 2$, this reads

$$H(\psi, x, u) := \psi_0\, f_0(x,u) + \psi_1\, f_1(x,u) + \psi_2\, f_2(x,u). \tag{6}$$

By the Hamilton function, the adjoint equation associated with a pair $(x(\cdot), u(\cdot))$ is defined by

$$\dot{\psi}_i = -\frac{\partial H}{\partial x_i} \quad \text{for } i = 0,\ldots,n, \tag{7}$$

or, written in more explicit terms,

$$\dot{\psi}_i(t) = -\sum_{k=0}^{n} \frac{\partial f_k}{\partial x_i}(x(t), u(t))\psi_k(t) \quad \text{for } i = 0, \ldots, n. \tag{8}$$

This is a linear system of differential equation for the vector function ψ. Notice that (8) implies $\dot{\psi}_0 = 0$ because f does not depend on x_0, hence ψ_0 is identically constant. Each solution $\psi : [0, T] \to \mathbb{R}^{n+1}$, $\psi(\cdot) = (\psi_0(\cdot), \ldots, \psi_n(\cdot))^\top$, of this adjoint equation is called *adjoint state* associated with the pair $(x(\cdot), u(\cdot))$.

Now we have all pre-requisites to formulate the Pontryagin maximum principle. Basic references for this optimality condition are Boltjanski et al. [12] or Ioffe / Tikhomirov [6], at least for a variable final time and a cost functional containing only the integral part of J. Functionals of Bolza type with fixed final time T, as the problem is posed here, require some standard transformations to obtain the associated Pontryagin principle in its most common form. We refer to the detailed discussion in Athans and Falb [2] or to the formulation in Pinch [11] or Feichtinger [3]. A review on different versions of the Pontryagin principle was given by Hartl et al. [5]. For problems with functional of Bolza type, unrestricted final state and fixed final time, which are autonomous in time, the following version holds true:

Theorem 1 (Pontryagin Maximum Principle)**.** *Let $u^*(\cdot)$ be an admissible control for the optimal control problem (4) with corresponding state $x^* = (x_1^*, \ldots, x_n^*)$. Then, in order that u^* and x^* minimize J subject to the given constraints, the following conditions must be satisfied: There exists a non-trivial adjoint state $\psi(\cdot) = (\psi_0, \psi_1, \ldots, \psi_n)^T$: $[0, T] \to \mathbb{R}^{n+1}$ that satisfies a.e. on $(0, T)$ the adjoint equation (8) with $x(t) := x^*(t)$ and $u(t) := u^*(t)$ together with the terminal condition*

$$\psi_i(T) = \psi_0 \frac{\partial g}{\partial x_i}(x^*(T)), \quad i = 1, \ldots, n,$$

and the Hamilton function H attains almost everywhere in $(0, T)$ its maximum with respect to u at $u = u^*(t)$, i.e.

$$\max_{v \in \Omega} H(\psi(t), x(t), v) = H(\psi(t), x^*(t), u^*(t)) \quad \textit{for a.a. } t \in (0, T). \tag{9}$$

Moreover, ψ_0 is a nonpositive constant.

The adjoint equation (8) is a linear system for $\psi(\cdot)$. Therefore, the adjoint state $\psi(\cdot)$ cannot vanish at any time unless it is identically zero. The latter case is excluded by the maximum principle. Therefore, it must hold $\psi_0 \neq 0$, since otherwise the terminal condition of the maximum would imply $\psi(T) = 0$ and hence $\psi(\cdot) = 0$. Re-defining ψ by $\psi := \psi/\psi_0$ and dividing (9) by $|\psi_0|$ we can therefore assume w.l.o.g that $\psi_0 = -1$. This is a standard conclusion for the type of problems above.

2.3 Necessary optimality conditions for the optimal control problem

In this part, we apply the Pontryagin maximum principle to derive the first order necessary optimality conditions for the optimal control $\bar{u}$ of the control problem (1) – (3).

Let us first explain how (1) – (3) fits in the problem (4), to which the Pontryagin principle refers. In (4), we must define $n,\ g,\ f_0,\ \ldots,\ f_n, x,\ x_0$, and Ω. In (1) – (3), we have $n = 1$ and we define $x_1 := i$, $x^0 := -i_0$. Moreover, f_0, f_1, and g are defined below. The set Ω is the interval $[u_a, u_b]$. In this way, (1) – (3) is related to (4), and we can apply the Pontryagin principle to it. As a result, we will obtain the system (10), (11). Although (11) contains the necessary information on an optimal control, it does not have the form we need for our numerical method. Therefore, in a second step, we transform (10), (11) to arrive at (15). This is the form we use for our numerical method. We recall that the control is pointwise bounded by real constants $u_a \leq u_b$, i.e. $u \in U_{ad}$. To cover the unbounded case, also $u_a = -\infty$ and/or $u_b = +\infty$ is formally allowed.

We re-write the differential equation (1) as

$$\frac{di}{dt}(t) = -\frac{R_c}{L}\, i(t) + \frac{1}{L} u(t)$$

and define

$$f_1(i, u) := -\frac{R_c}{L}\, i + \frac{1}{L} u.$$

The Bolza type objective functional in (4) is defined by

$$\begin{aligned} f_0(i,u) &:= \frac{\lambda_Q}{2}(i-i_0)^2 + \frac{\lambda_U}{2}u^2, \\ g(i) &:= \frac{\lambda_T}{2}(i-i_0)^2. \end{aligned}$$

As we explained at the end of the last section, we are justified to assume $\psi_0 = -1$. Writing for convenience $\psi := \psi_1$, we obtain

$$H = -\big(\frac{\lambda_Q}{2}\,\big(i - i_0\big)^2 + \frac{\lambda_U}{2}\,u^2\big) + \psi\,(-\frac{R_c}{L}\,i + \frac{u}{L}).$$

From this definition of H, we obtain the adjoint equation

$$\frac{d\,\psi}{d\,t}(t) = -\frac{\partial\,H}{\partial\,i}(t) = \lambda_Q\,\big(i(t) - i_0\big) + \psi(t)\frac{R_c}{L}.$$

Furthermore, the Pontryagin principle yields the final condition

$$\psi(T) = -\frac{\partial\,g}{\partial\,i}(i(T)) = -\lambda_T\big(i(T) - i_0\big).$$

We have derived the following adjoint system:

$$\begin{aligned} \frac{d\psi}{dt}(t) &= -\psi(t)\,\frac{R_c}{L} + \lambda_Q\big(i(t) - i_0\big) \\ \psi(T) &= -\lambda_T\,(i(T) - i_0). \end{aligned} \tag{10}$$

As before, we denote the optimal quantities by i^* and u^* and omit the star at the associated adjoint state ψ. The constraint (3) is considered by the following variational inequality that is easily derived from the maximum condition (9),

$$\frac{\partial\,H}{\partial\,u}(i^*(t), u^*(t), \psi(t))(u - u^*(t)) = (-\lambda_u u^*(t) + \psi(t)/L)(u - u^*(t)) \le 0 \quad \forall\,u \in U_{ad}, \;\text{ for a.a. } t \in (0,T).$$

We write this in the more common form of a variational inequality,

$$-\frac{\partial\,H}{\partial\,u}(i^*(t), u^*(t), \psi(t))(u - u^*(t)) = (\lambda_u u^*(t) - \psi(t)/L)(u - u(t)^*) \ge 0 \quad \forall\,u \in U_{ad}. \tag{11}$$

In some sense, this form does not really fit to numerical computations. In order to arrive at the standard form of nonlinear programming with associated Lagrange multipliers, we perform a substitution and transform the adjoint state ψ to a Lagrange multiplier p. To this aim, we set

$$p(t) = -\frac{\psi(t)}{L}, \tag{12}$$

which implies

$$\frac{dp}{dt}(t) = -\frac{1}{L}\frac{d\,\psi}{d\,t}(t).$$

We replace ψ by p in equations (10) and obtain the system

$$\begin{aligned} -L\frac{dp}{dt}(t) &= -R_c\,p(t) + \lambda_Q\big(i(t) - i_0\big) \\ -L\,p(T) &= -\lambda_T\,(i(T) - i_0). \end{aligned} \tag{13}$$

Using the same substitution in equation (11), we find

$$(\lambda_U u^*(t) + p(t))(u - u^*(t)) \ge 0 \quad \forall\,u \in [u_a, u_b] \text{ and a.a. } t \in (0,T). \tag{14}$$

Therefore, we can completely describe the optimal control u^* by the following case study:

$$
\begin{array}{llll}
(i) & \lambda_U u^*(t) + p(t) > 0 & \Rightarrow & u^*(t) = u_a \\
(ii) & \lambda_U u^*(t) + p(t) < 0 & \Rightarrow & u^*(t) = u_b \\
(iii) & \lambda_U u^*(t) + p(t) = 0 & \Rightarrow & u^*(t) = -\dfrac{1}{\lambda_U} p(t).
\end{array}
$$

A further discussion of these relations yields that the optimal control satisfies almost everywhere the well known projection formula

$$u^*(t) = \mathbb{P}_{[u_a,u_b]}\{-\frac{1}{\lambda_U}p(t)\} = \max(\min(u_b, -\frac{1}{\lambda_U}p(t)), u_a),$$

where $\mathbb{P}_{[\alpha,\beta]}$ denotes the projection of a real number on the interval $[\alpha,\beta]$. For the derivation of this formula, we refer the reader e.g. to [14]. After inserting this expression for u^* in (1) while considering also (13), the two unknowns i^* and p satisfy the forward-backward optimality system

$$
\begin{aligned}
L\frac{d\,i}{d\,t}(t) + R_c\, i(t) &= \mathbb{P}_{[u_a,u_b]}\{-\frac{1}{\lambda_U}p(t)\} \\
i(0) &= -i_0 \\
-L\,\frac{d\,p}{d\,t}(t) + R_c\, p(t) &= \lambda_Q(i(t) - i_0) \\
L\, p(T) &= \lambda_T(i(T) - i_0).
\end{aligned}
\tag{15}
$$

Although this is a nonlinear and nondifferentiable system, it can be easily solved numerically by available commercial codes. In our computations, we applied the code COMSOL Multiphysics®[1]. In the next section, we demonstrate the efficiency of this idea by numerical examples. There are various other numerical methods for solving systems of ordinary differential equations. We only refer to the standard reference Hairer et al. [4]. Moreover, we mention methods tailored to the numerical treatment of electric circuits, which are presented in Rosłoniec [13]. We prefer COMSOL Multiphysics®, because we were able to implement the solution of our nonsmooth optimality systems in an easy way.

For convenience, we also mention the unbounded case $u_a = -\infty$ and $u_b = \infty$. Here, we have

$$\mathbb{P}_{[u_a,u_b]}\{-\frac{1}{\lambda_U}p(t)\} = -\frac{1}{\lambda_U}p(t),$$

hence the associated optimality system is in the unbounded case

$$
\begin{aligned}
L\frac{d\,i}{d\,t}(t) + R_c\, i(t) &= -\frac{1}{\lambda_U}p(t) \\
i(0) &= -i_0 \\
-L\frac{dp}{dt}(t) + R_c\, p(t) &= \lambda_Q(i(t) - i_0) \\
L\, p\,(T) &= \lambda_T(i(T) - i_0).
\end{aligned}
\tag{16}
$$

The unique solution to this system is the pair (i^*, p), and the optimal control is given by $u^* = -\lambda_U^{-1} p$.

2.4 Numerical examples

In the examples, we select the bounds $u_a := -300$ and $u_b := 300$. Moreover, we consider the time interval $[0, 1]$ and fix the constants $i_0 := 1,\ L := 3.5$ and $R_c := 145$.

First, we investigate the influence of the parameters λ_T, λ_Q and λ_U. As explained above, we solve the nonlinear system (15) numerically to find the optimal solution. Let us start with the influence of λ_Q. This parameter controls the current i along the time. We fix λ_T and λ_U by 1 and vary λ_Q. Figure 2 shows how the current i approaches i_0 for increasing λ_Q. The larger λ_Q is, the faster the current approaches the target. The end of the curves is depending on the ratio between λ_Q and λ_T. We discuss this below.

Next, we consider the influence of the control parameter λ_U. The other parameters are fixed by 1. This parameter expresses the cost of the control u. If λ_U is large, the control is expensive and it is more important to minimize $|u|$, cf. Figure 3. In its right hand side figure, we see that the upper bound u_b is active at the beginning.

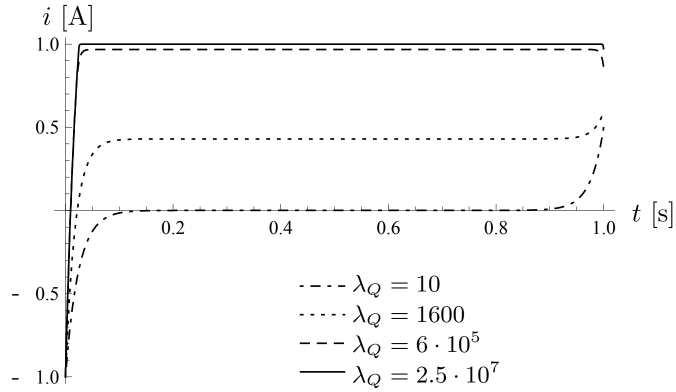
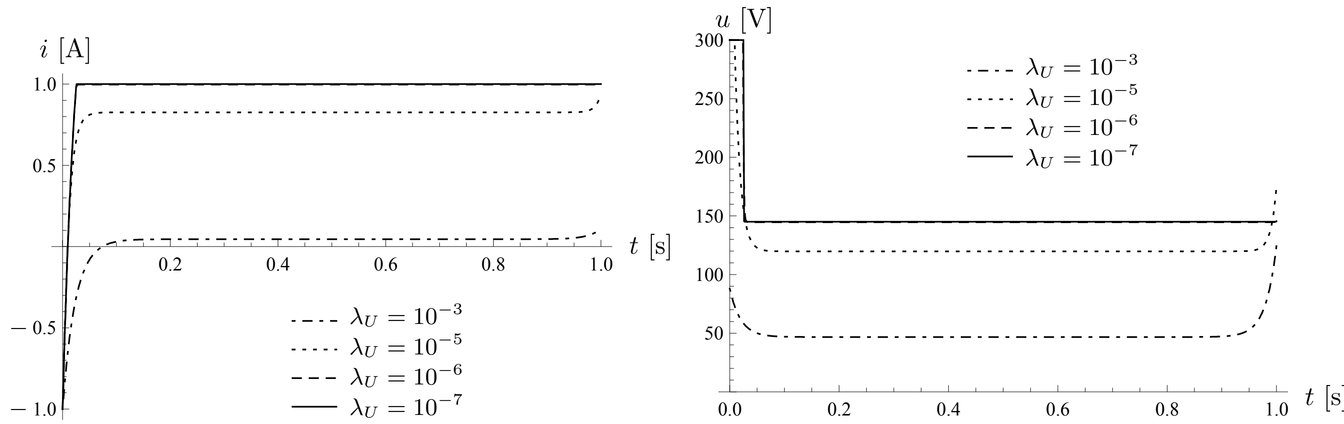


Figure 2: Electric current i by different λ_Q



(a) Electric current i depending on λ_U

(b) Voltage u depending on λ_U

Figure 3: Current i by different λ_U

Finally, we analyse all parameters more detailed and plot only the integral part

$$J_Q := \int_0^T \big(i(t) - i_0\big)^2 dt$$

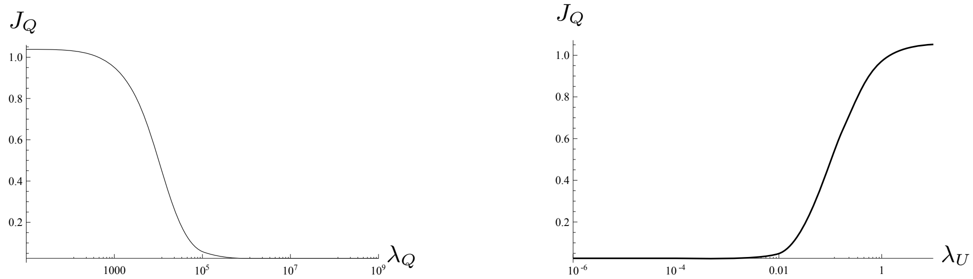
of the cost functional depending on the parameters. We are most interested in this part, since it expresses the speed of approaching i_0. Our aim is to select parameters λ_Q and λ_U, where J_Q is as small as possible. We fix $\lambda_T = 1$ again. We vary only one of the parameters λ_Q and λ_U and set the other one to 1. Figure 4 shows the value of J_Q depending on λ_Q, respectively λ_U. The x-axis are scaled logarithmically. Both curves contain critical points beyond which J_Q grows considerably. In view of this, we recommend the choice $\lambda_Q > 10^5$ and $\lambda_U < 10^{-4}$.

With nearly this ratio between λ_Q and λ_U, we try to find out a suitable λ_T. Fixing $\lambda_Q = 10^6$ and $\lambda_U = 10^{-6}$, we vary λ_T. The left hand side of Figure 5 shows that the influence of λ_T on i is marginal almost everywhere. However, a zoom of the graph close to the final time T reveals that this parameter influences the final behavior. If λ_T is large enough, i^* ends near by i_0.

We started with a simple circuit. In the next sections, we discuss some more examples of higher complexity.

2.5 A two-loop circuit modeled by a single ODE

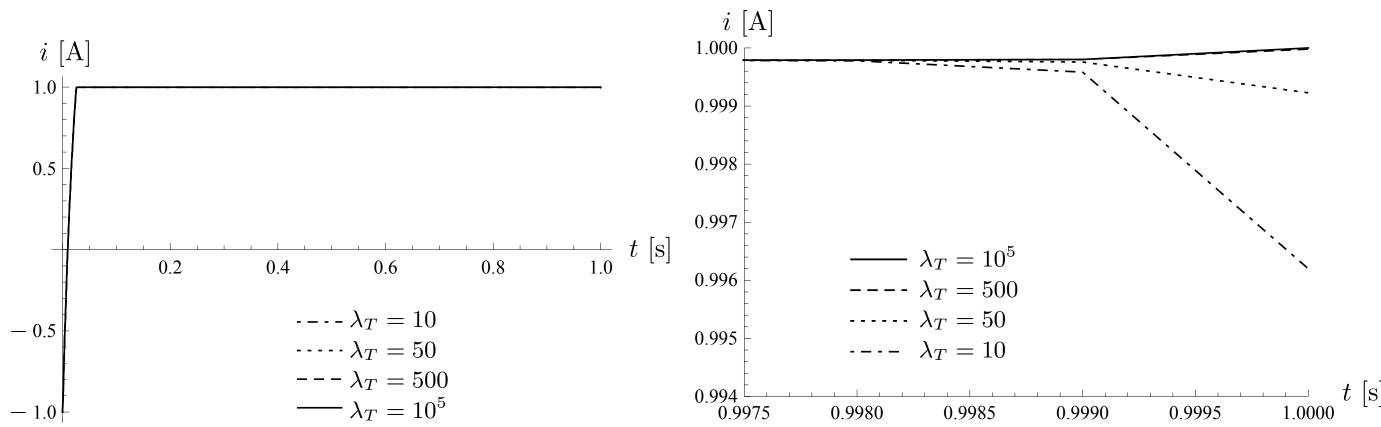
In the remainder of this paper, we will consider several cases of electrical circuits with two loops. We start with a simple one and show that it can be still modeled by a single ODE. Therefore, this circuit still fits to the theory of this section. This subsection is only to show how we transform this case to the single circuit case. Later, we consider linear and nonlinear systems of two ODEs.



(a) The value of J_q influenced by λ_Q

(b) The value of J_q influenced by λ_U

Figure 4: Value J_Q by different parameters



(a) Electric current i depending on λ_T)

(b) Electric current i depending on λ_T near T

Figure 5: Electric current depending on λ_T

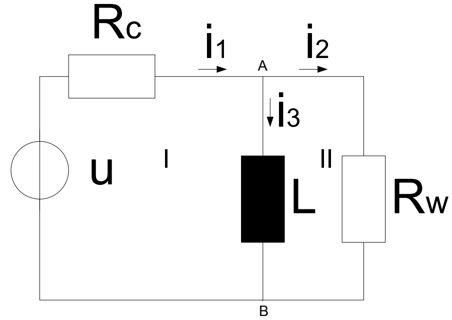


Figure 6: Electrical circuit with two loops

Consider the circuit of Figure 6 that contains an additional resistance R_w. Let us set up the differential equations for this circuit. Applying the Kirchhoff current law to node A, we get the equation

$$i_1(t) = i_2(t) + i_3(t) \tag{17}$$

at any time t. The electrical circuit is composed of the loops I and II. By the Kirchhoff voltage law, we obtain the equations

$$L\frac{di_3}{dt}(t) + R_c\, i_1(t) \quad = \quad u(t) \tag{18}$$

for loop I and

$$-\,L\frac{di_3}{dt}(t) + R_w\, i_2(t) \quad = \quad 0 \tag{19}$$

for loop II. Our main interest is to control the current i_3, which in the sequel is re-named by $i := i_3$. From (17), (18) and (19), we deduce the following equation for the current $i = i_3$:

$$L(1 + \frac{R_c}{R_w})\frac{di}{dt}(t) + R_c\, i(t) = u(t).$$

With the initial condition $i(0) = -i_0$, we arrive at the initial value problem

$$\begin{aligned} a\,\frac{di}{dt}(t) + R_c\, i(t) &= u(t) \\ i(0) &= -i_0, \end{aligned} \tag{20}$$

where a is defined by

$$a := L\,(1 + \frac{R_c}{R_w}).$$

Our goal is to control the electrical currents such that the objective functional (2) is minimized. By our transformation above, this modified problem, which consists of the new state equation (20), original cost functional (2) and control constraints (3), is a particular case of the problem (1) – (3). Therefore, we obtain an optimality system that is equivalent to (15) in the case of a bounded control and (16) for unbounded control. Because of this similarity, we abdicate of numerical examples.

3 Two loops coupled by inductances

3.1 The optimal control problem

In this section, we consider the optimal control of electrical circuits, which cannot be transformed easily to a single ODE. In the preceding sections, we derived the optimality system from the Pontryagin maximum principle and transformed this result into the Lagrangian form. Here, we directly apply the Karush-Kuhn-Tucker theory (KKT theory) of nonlinear programming problems in Banach spaces to deduce the optimality system. We do this in a formal way that does not explicitely mention the function spaces underlying the theory. The precise application of the standard Karush-Kuhn-Tucker theory to problems with differential equations is fairly involved, we refer, e.g., to Ioffe and Tikhomirov [6], Luenberger [8], or Jahn [7]. The assumptions of the KKT theory are met by all of our problems so that the same results will be obtained as the ones we derive below by the formal Lagrange technique, which is adopted from the textbook [14]. The optimality conditions obtained in this way are closer to the applications of nonlinear programming techniques.

We consider the electrical circuit shown in Figure 7 with two loops coupled by inductances L_1 and L_2 via a coupling variable γ. In some sense, this model simulates the coupling between a magnetic field and eddy currents. For physical reasons, we require $\gamma \in [0, 1]$.

This electrical circuit is modeled by the initial-value problem

$$\begin{aligned} L_1\,\frac{di_1}{dt}(t) + C\,\frac{di_2}{dt}(t) + R_c\, i_1(t) &= u(t) \\ C\,\frac{di_1}{dt}(t) + L_2\,\frac{di_2}{dt}(t) + R_w i_2(t) &= 0 \\ i_1(0) &= -i_0 \\ i_2(0) &= 0, \end{aligned} \tag{21}$$

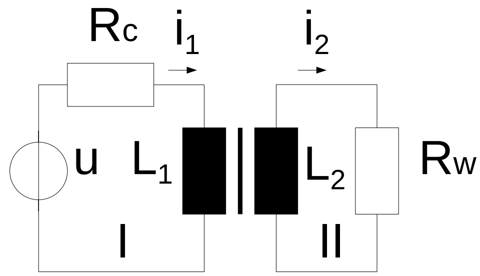


Figure 7: Model 2

which contains the ODEs for loop I and loop II, respectively. The constant C is defined by $C := \gamma\sqrt{L_1 L_2}$. Now, the cost functional to be minimized is defined by

$$J(i_1, i_2, u) := \frac{\lambda_T}{2}(i_1(T) - i_0)^2 + \frac{1}{2}\int_0^T \left(\lambda_Q\,(i_1(t) - i_0)^2 + \lambda_I\, i_2(t)^2 + \lambda_U\, u(t)^2\right)\,dt \tag{22}$$

with nonnegative constants λ_T, λ_Q, λ_I and $\lambda_U > 0$.

The optimal control problem we consider in this subsection is to minimize (22) subject to the state equation (21) and to the box constraint (3).

To derive the associated optimality system, we might multiply the system of ODEs in (21) by the matrix E standing in front of the derivatives di_j/dt, $j = 1, 2$,

$$E = \begin{bmatrix} L_1 & C \\ C & L_2 \end{bmatrix} \tag{23}$$

to obtain a system of ODEs in standard form. This is, however, not needed here. We will use this approach for the discussion of controllability in the next section. In contrast to the Pontryagin maximum principle, the formal Lagrange principle is directly applicable to the system (21). Following the Lagrange principle, we "eliminate" the ODEs above by two Lagrange multiplier functions p_1 and p_2, respectively, while we keep the initial conditions and the box constraint $u \in U_{ad}$ as explicit constraints. Therefore, the associated Lagrange function is defined by

$$\begin{aligned}\mathcal{L}(i_1, i_2, u, p_1, p_2) := J(i_1, i_2, u) &- \int_0^T \left(L_1 \frac{di_1}{dt}(t) + C\,\frac{di_2}{dt}(t) + R_c\, i_1(t) - u(t)\right) p_1(t)\,dt \\ &- \int_0^T \left(L_2 \frac{di_2}{dt}(t) + C\,\frac{di_1}{dt}(t) + R_w\, i_2(t)\right) p_2(t)\,dt.\end{aligned}$$

According to the Lagrange principle, the optimal solution (i_1^*, i_2^*, u^*) has to satisfy the necessary optimality conditions of the problem to minimize $\mathcal{L}$ with respect to i_1, i_2, u without the ODE constraints but subject to the initial conditions and the box constraints. Notice that i_1, i_2, u are here decoupled. For i_1, i_2, this means

$$\frac{\partial\mathcal{L}}{\partial i_j}(i_1^*, i_2^*, u^*, p_1, p_2)\,h = 0 \qquad \forall h \text{ with } \; h(0) = 0,\; j = 1, 2.$$

With respect to u, the variational inequality

$$\frac{\partial\mathcal{L}}{\partial u}(i_1^*, i_2^*, u^*, p_1, p_2)\,(u - u^*) \geq 0 \qquad \forall\, u \in U_{ad}$$

must be satisfied. Let us write for short $\mathcal{L} = \mathcal{L}(i_1^*, i_2^*, u^*, p_1, p_2)$. For $\partial\mathcal{L}/\partial u$, we obtain

$$\frac{\partial\mathcal{L}}{\partial u}(i_1^*, i_2^*, u^*, p_1, p_2)\,u = \int_0^T (\lambda_U\, u(t) + p_1(t))\; u(t)\,dt,$$

hence we have the variational inequality

$$\int_0^T (\lambda_U\, u^*(t) + p_1(t))\; (u(t) - u^*(t))\, dt \ge 0 \quad \forall\, u \in U_{ad}.$$

Again, the pointwise evaluation of the inequality leads finally to the projection formula

$$u^*(t) = \mathbb{P}_{[u_a,u_b]}\{-\frac{1}{\lambda_U} p_1(t)\}. \tag{24}$$

For $\partial\,\mathcal{L}/\partial\, i_1$, we deduce after integrating by parts

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial i_1}(i_1, i_2, u, p_1, p_2)\, h &= \lambda_T (i_1(T) - i_0)\, h(T) + \lambda_Q \int_0^T (i_1(t) - i_0)\, h(t)\, dt \\
&\quad - \int_0^T (L_1 \frac{dh}{dt}(t) + R_c\, h(t))\, p_1(t)\, dt - \int_0^T C\, \frac{dh}{dt}(t)\, p_2(t)\, dt, \\
&= \{\lambda_T (i_1(T) - i_0) - L_1\, p_1(T) - C\, p_2(T)\}\; h(T) \\
&\quad + \int_0^T \left\{ \lambda_Q\, (i_1(t) - i_0) - R_c\, p_1(t) + L_1\, \frac{dp_1}{dt}(t) + C\, \frac{dp_2}{dt}(t) \right\}\; h(t)\, dt \\
&= 0 \quad \forall\, h \text{ with } h(0) = 0.
\end{aligned}$$

Since $h(T)$ and $h(\cdot)$ can vary arbitrarily, both terms in braces must vanish; hence the equations

$$\begin{aligned}
-L_1\, \frac{dp_1}{dt}(t) - C\, \frac{dp_2}{dt}(t) + R_c p_1(t) &= \lambda_Q (i_1(t) -\, i_0(t)) \\
L_1\, p_1(T) + C\, p_2(T) &= \lambda_T (i_1(T) - i_0)
\end{aligned} \tag{25}$$

must hold. In the same way, we derive from $\partial\,\mathcal{L}/\partial\, i_2 = 0$ the equations

$$\begin{aligned}
-C\, \frac{dp_1}{dt}(t) - L_2\, \frac{dp_2}{dt}(t) + R_w p_2(t) &= 0 \\
C\, p_1(T) \,+ L_2\, p_2(T) &= 0.
\end{aligned} \tag{26}$$

The equations (25) and (26) form the adjoint system. Collecting the equations (21), (24), (25), and (26), we obtain the following optimality system to be satisfied by the optimal triple (i_1^*, i_2^*, u^*) and the associated pair of Lagrange multipliers p_1, p_2:

$$\begin{aligned}
L_1\, \frac{di_1}{dt}(t) + C\, \frac{di_2}{dt}(t) + R_c\, i_1(t) &= \mathbb{P}_{[u_a,u_b]}\{-\lambda_U^{-1} p_1(t)\} \\
C\, \frac{di_1}{dt}(t) + L_2\, \frac{di_2}{dt}(t) + R_w\, i_2(t) &= 0 \\
-L_1\, \frac{dp_1}{dt}(t) - C\, \frac{dp_2}{dt}(t) + R_c\, p_1(t) &= \lambda_Q\, (i_1(t) - i_0(t)) \\
-C\, \frac{dp_1}{dt}(t) - C\, L_2\, \frac{dp_2}{dt}(t) + R_w p_2(t) &= \lambda_I\, i_2(t) \\
i_1(0) &= -i_0 \\
i_2(0) &= 0 \\
L_1\, p_1(T) + C\, p_2(T) &= \lambda_T (i_1(T) - i_0) \\
C\, p_1(T) \,+ L_2\, p_2(T) &= 0.
\end{aligned} \tag{27}$$

If the box constraints missing on u, then we can replace the nonlinear and nonsmooth term $\mathbb{P}_{[u_a,u_b]}\{-\lambda_U^{-1} p_1(t)\}$ by the linear and smooth expression $-\lambda_U^{-1} p_1(t)$. To handle the numerical examples below, we directly solved the nonsmooth optimality system (27) by the package COMSOL Multiphysics®.

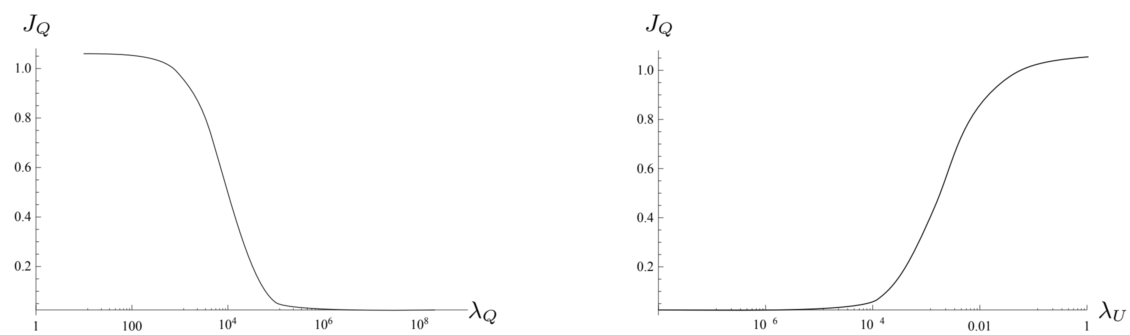
3.2 Numerical examples

In the problem above, we fix the final time $T = 1$ and the constants $i_0 := 1,\ L := 3.5, L_2 := 2,\ R_c := 145,\ R_w := 3500$ and $\gamma = 0.9$. Moreover, we impose the bounds $u_a := -300$ and $u_b := 300$ on the control.

First, we discuss a reasonable choice of the parameters λ_Q, and λ_U. Therefore, we define again

$$J_Q := \int_0^T \big(i_1(t) - i_0\big)^2 dt \tag{28}$$

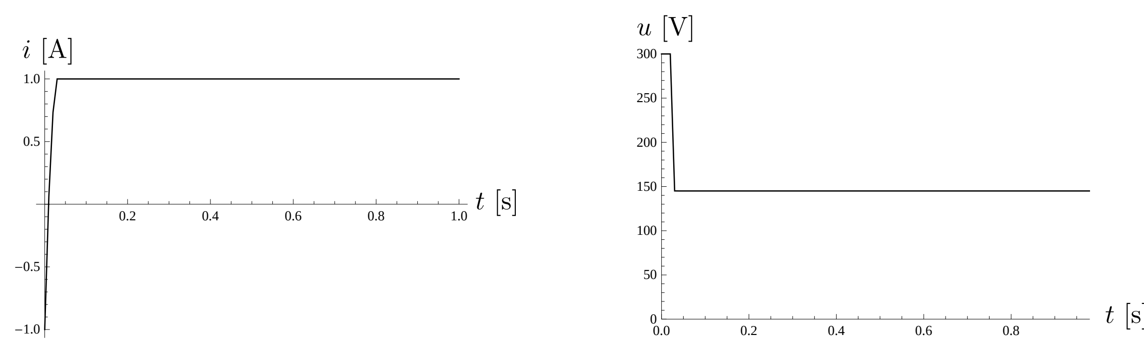
plot J_Q for different parameters. As in Subsection 2.4, we select parameters, where J_Q is as small as possible. The parameters λ_T and λ_I have a very low influence on the size of J_Q and we fix $\lambda_T = 10000$ and $\lambda_I = 100$. We vary only one of the parameters λ_Q and λ_U and set the other one to 1. The left hand side of Figure 8 shows the value of J_Q depending on λ_Q and the right hand side depending on λ_U. We have nearly the same critical points as in model 1 (cf. Figure 4).



(a) Value of J_q depending on λ_Q (λ_U=1) (b) Value of J_q depending on λ_U (λ_Q=1)

Figure 8: Value of J_Q by different parameters

Let us now plot the current i and the control u with the parameters $\lambda_Q = 10^6$, $\lambda_U = 10^{-6}$, $\lambda_T = 10000$ and $\lambda_I = 100$ which have been found above. On the left hand side of Figure 9, we see how the current i approaches $i_0 = 1$. Near the time $t = 0$ the control u stays at the upper bound u_b, before it approaches the lower holding level (right hand side of Figure 9).



(a) Electric current i_1 (b) Voltage u

Figure 9: Solution of Model 2

Additionally, we concentrate on the current i_2. Because i_2 is coupled with i_1, we can neglect the time where i_1 is almost constant. Therefore, Figure 10 shows only a part of the time interval.

3.3 Time-optimal control

In the previous sections, we considered optimal control problems in a fixed interval of time. We formulated integral terms in the cost functional to reach the target state in short time. As we have explained in Section 2.1, we need

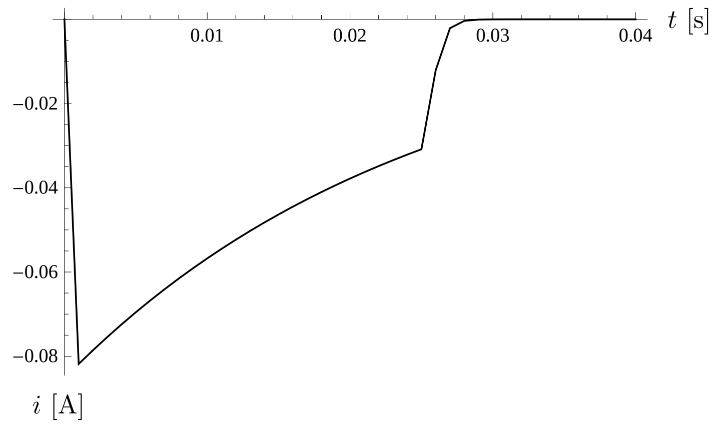


Figure 10: Current i_2 in Model 2

this alternative, because this problem is only a first step for solving time-optimal control problems in the optimal switching between magnetic fields. Then, the 3D Maxwell equations occur and it is really difficult to formulate a time-optimal problem in the right way.

However, in the case of ODEs, it is also useful to reach certain desirable states in minimal time. Such aspects of time optimality are the topic of this section. We consider again the state equation (21) as the model of the electrical circuit. Now, the vector of currents (i_1, i_2) should reach the desired value $(i_0, 0)$ in shortest time. After $(i_0, 0)$ is reached, the currents should hold this state. Again, the initial value is $(-i_0, 0)$.

To this aim, we consider the time-optimal control problem

$$\min T \tag{29}$$

subject to $u \in U_{ad}$ and the state equation

$$\begin{aligned} L_1 \frac{di_1}{dt}(t) + C \frac{di_2}{dt}(t) + R_c\, i_1(t) &= u(t) \\ C \frac{di_1}{dt}(t) + L_2 \frac{di_2}{dt}(t) + R_w i_2(t) &= 0 \\ i_1(0) &= -i_0 \\ i_2(0) &= 0 \end{aligned} \tag{30}$$

and the end conditions

$$\begin{aligned} i_1(T) &= i_0 \\ i_2(T) &= 0. \end{aligned} \tag{31}$$

It turns out that the solution to this problem is also the one that steers the initial state in shortest time to a holdable state $i_1 = i_0$.

In what follows, we shall consider the row vector $i := (i_1, i_2)^\top$. First, we discuss the problem of reachability of the terminal vector $(i_0, 0)^\top$ in finite time. To apply associated results of the controllability theory of linear autonomous systems, we first transform the system (30) to the standard form

$$i'(t) = A\, i(t) + B\, u(t). \tag{32}$$

This equation is obtained after multiplying (30) by

$$E^{-1} = \frac{1}{D} \begin{bmatrix} L_2 & -C \\ -C & L_1 \end{bmatrix}, \tag{33}$$

where E is defined in (23) and $D = (1-\gamma^2)L_1 L_2$ is its determinant. We obtain

$$A = -E^{-1} \begin{bmatrix} R_c & 0 \\ 0 & R_w \end{bmatrix} = \frac{1}{D} \begin{bmatrix} -L_2 R_c & \gamma\sqrt{L_1 L_2} R_w \\ \sqrt{L_1 L_2} R_c & -L_1 R_w \end{bmatrix}, \qquad B = E^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{D} \begin{bmatrix} L_2 \\ -C \end{bmatrix}. \tag{34}$$

Let us first formally determine the limit $\lim_{t\to\infty} i(t)$ for a constant control function $u(t) \equiv u$. Inserting $i(t) = 0$ in (30), we find that $i_1 = R_c^{-1} u$ and $i_2 = 0$ are the limits, provided they exist.

Theorem 2. *For $0 < \gamma < 1$, the eigenvalues of the matrix A are real, negative and distinct. For $u(t) \equiv u$, it holds*

$$\lim_{t\to\infty} i(t) = \left[\begin{array}{c} R_c^{-1} u \\ 0 \end{array} \right].$$

Proof. The characteristic equation for the eigenvalues $\lambda_{1/2}$ is

$$\det(A - \lambda I) = \lambda^2 + \lambda\,(L_1 R_w + L_2 R_c) + (1-\gamma^2) L_1 L_2 R_c R_w = 0,$$

hence

$$\lambda_{1/2} = \frac{1}{2}\left(-(L_1 R_w + L_2 R_c) \pm \sqrt{L_1^2 R_w^2 + (2 - 4(1-\gamma^2)) L_1 L_2 R_c R_w + L_2^2 R_c^2} \right).$$

Since $0 < \gamma < 1$, a simple estimate shows

$$\begin{aligned} 0 \le (L_1 R_w - L_2 R_c)^2 &= L_1^2 R_w^2 - 2 L_1 L_2 R_c R_w + L_2^2 R_c^2 \\ &< L_1^2 R_w^2 + (2 - 4(1-\gamma^2)) L_1 L_2 R_c R_w + L_2^2 R_c^2. \end{aligned}$$

Therefore, the square root is real and nonvanishing, both eigenvalues are negative and $\lambda_1 \neq \lambda_2$.

To show the second claim, we recall the variation of constants formula

$$\begin{aligned} i(t) &= e^{At} \left[\begin{array}{c} -i_0 \\ 0 \end{array} \right] + \int_0^t e^{A(t-s)} B \left[\begin{array}{c} u \\ 0 \end{array} \right] ds \\ &= e^{At} \left[\begin{array}{c} -i_0 \\ 0 \end{array} \right] + A^{-1}(I - e^{At}) B \left[\begin{array}{c} u \\ 0 \end{array} \right]; \end{aligned}$$

notice that u is a constant. Since the eigenvalues of A are negative, the matrix exponential function tends to zero as $t \to \infty$, hence

$$\lim_{t\to\infty} i(t) = A^{-1} B \left[\begin{array}{c} u \\ 0 \end{array} \right].$$

Moreover, we analogously have $i'(t) \to 0$. Therefore the claimed limes follows directly by passing to the limit in (30). $\square$

Next, we confirm that the well-known Kalman condition

$$\operatorname{rank}[B, A\,B] = 2 \tag{35}$$

is satisfied. In fact, the vectors B and $A\,B$ are linearly independent iff $E\,B$ and $E\,A\,B$ have the same property, because E is invertible. We have

$$EAB = -E\,E^{-1} \left[\begin{array}{cc} R_c & 0 \\ 0 & R_w \end{array} \right] E^{-1} \left[\begin{array}{c} 1 \\ 0 \end{array} \right] = \frac{1}{D} \left[\begin{array}{c} R_c L_2 \\ -C\,L_2 \end{array} \right] \quad \text{and} \quad E\,B = E E^{-1} \left[\begin{array}{c} 1 \\ 0 \end{array} \right] = \left[\begin{array}{c} 1 \\ 0 \end{array} \right],$$

and both vectors are obviously linearly independent. The Kalman condition is the main assumption for proving the next result.

Theorem 3. *Suppose that $u_b > R_c i_0$. Then there is a control function $u \in U_{ad}$ that steers the initial value $(-i_0, 0)^\top$ in finite time to the terminal state $(i_0, 0)^\top$.*

Proof. In a first time interval $[0, t_1]$, we fix the control by the hold-on voltage, i.e.

$$u(t) = R_c i_0, \quad t \in [0, t_1].$$

We know by Theorem 2 that $i(t) \to (i_0, 0)^\top$ as $t_1 \to \infty$. Therefore, given an $\varepsilon > 0$

$$|i(t_1) - (i_0, 0)^\top| < \varepsilon,$$

if t_1 is sufficiently large. We then have $i(t_1) = (i_0, 0)^\top + \eta$ with $|\eta| < \varepsilon$. Starting from $i(t_1)$ as new initial value, we reach $(i_0, 0)^\top$ after time $t_2 > 0$ as follows: We fix $\tau > 0$ and set

$$u(t_1 + \tau) = R_c i_0 + v(\tau), \quad \tau \in [0, t_2].$$

By the variation of constants formula, we obtain

$$\begin{aligned} i(t_1 + \tau) &= e^{A\tau}\begin{pmatrix} i_0 \\ 0 \end{pmatrix} + \int_0^\tau e^{A(\tau-s)} B\, R_c i_0\, ds + \int_0^\tau e^{A(\tau-s)} B\, v(s)\, ds. && (36) \\ &= \begin{pmatrix} i_0 \\ 0 \end{pmatrix} + \int_0^\tau e^{A(\tau-s)} B\, v(s)\, ds + \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}, && (37) \end{aligned}$$

because the first two items in the right-hand side of (36) express the stationary current $(i_0, 0)^\top$ associated with the hold-on voltage $R_c i_0$. At time $\tau = t_2$, we want to have $i(t_1 + \tau) = (i_0, 0)^\top$, hence by (37) we have to find v such that

$$-\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \int_0^{t_2} e^{A(t_2-s)} B\, v(s)\, ds,$$

i.e. equivalently

$$-e^{-A\,t_2}\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \int_0^{t_2} e^{-A\,s} B\, v(s)\, ds. \tag{38}$$

Thanks to the Kalman condition, the operator on the right-hand side maps any open ball of $L^\infty(0, t_2)$ of radius $\delta > 0$ around 0 in an open ball of $\mathbb{R}^2$ with radius $\varepsilon > 0$, cf. Macki and Strauss [9], proof of Theorem 3, chapter II, (3). By our assumption $u_b > R_c i_0$, we have that $R_c i_0 + v(t)$ is an admissible control, if $|v(t)| \le \delta$ and $\delta > 0$ is sufficiently small. We fix now $t_2 > 0$ and select $\delta > 0$ so small that $R_c i_0 + v(t)$ is admissible. Moreover, we take t_1 so large that $\eta = i(t_1) - (i_0, 0)^\top$ belongs to the ball of $\mathbb{R}^2$ around zero mentioned above.

Then the control

$$u(t) = \begin{cases} R_c i_0 & \text{on } [0, t_1] \\ R_c i_0 + v(t - t_1) & \text{on } (t_1, t_1 + t_2] \end{cases}$$

has the desired properties. $\square$.

Now we have all prerequisites for a discussion of the time-optimal control.

Theorem 4. *If $\gamma \in (0, 1)$ and $u_b > R_c i_0$ is assumed, then the time-optimal control problem* (29) – (31) *has a unique optimal control. This control is bang-bang and has at most one switching point.*

Proof. It follows from the last theorem that the final state $(i_0, 0)^\top$ is reachable in finite time by admissible controls. Therefore, Theorem 1 of [9], chpt. II yields the existence of at least one optimal control. (This theorem deals with the problem to reach the target $(0, 0)^\top$, but the proof works for our situation with a marginal change).

We know that the Kalman condition is satisfied by our problem, which is here equivalent to the normality of the system, because B consists of a single column vector, cf. [9], chpt. III, Theorem 7. Therefore, thanks to Theorem 4 of [9], chpt. III, the optimal control is unique and bang-bang with at most finitely many switching points.

Now we apply the maximum principle for linear time-optimal control problems, [9], Theorem 3, chpt. III. It says that there exists a nonzero constant vector $h \in \mathbb{R}^2$ such that the optimal control u satisfies

$$u(t) = \begin{cases} u_b & \text{if } h^\top exp(-At)B > 0 \\ u_a & \text{if } h^\top exp(-At)B < 0 \end{cases} \tag{39}$$

for almost all $t \in [0, T]$, where T is the optimal time.

Because the eigenvalues of A are real and distinct and $h \neq 0$, the function $h^\top exp(-At)B$ is a sum of two exponential functions, which can have at most one zero. From this, the claim on the number of switching points follows immediately. $\square$

By the maximum principle (39), the optimal control has at most one switching point t_1. Thanks to the underlying background of physics, it can be found numerically as follows:

In a first interval $[0, t_1]$, we fix the control u at the upper bound u_b and require that $i_1(t_1) > i_0$. In $(t_1, T]$, we set the control to the lower bound u_a until it reaches the value i_0 at the time T, i.e.

$$i_1(T) = i_0.$$

Of course, T depends on t_1, say $T = \tau(t_1)$. We determine t_1 by the condition $i_2(T) = 0$, i.e.

$$i_2(\tau(t_1)) = 0. \tag{40}$$

Time optimal control to a holding level. In the problem (29) – (31), the target is the terminal state $(i_0, 0)$. After having reached this state at time T, we are able to fix the control to the holding voltage

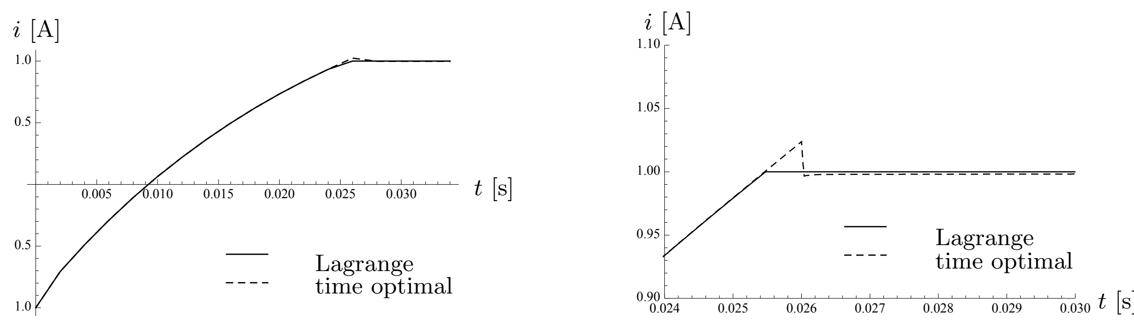
$$u_0 := R_c i_0.$$

Then, for $t > T$, the current i is identically constant, $i(t) = (i_0, 0)\ \forall t > T$. In this way, the solution of our time-optimal control problem solves also the problem to reach the holdable state $i_1 = i_0$ in shortest time.

To a certain extent, the aim of reaching the holdable state in shortest time can also be covered by the cost functional J in (22). The first term forces i_1 to reach i_0 while the second will become small, if this is done quite fast. However, the use of this functional will only lead to an approximation of time optimality.

Now we will compare the above method with the optimal control for the problem (22) of Subsection 3.1. It turns out that the optimal time to reach the holdable state, can also be obtained from (22) by a suitable choice of the parameters. We fixed them as follows:

The control bounds are again $u_a := -300$ and $u_b := 300$. We have seen, that a time close to $t = 0$ is important and select $T = 0.035$. Furthermore we fix $i_0 := 1$, $L := 3.5$, $L_2 := 2$, $R_c := 145$, $R_w := 3500$, $\gamma = 0.9$, $\lambda_T = 10000$, $\lambda_Q = 10^6$, $\lambda_U = 10^{-6}$, and $\lambda_I = 10000$. The solution with the associated cost functional J is plotted in all pictures as a solid line. The solution obtained by optimizing the points t_1, t_2 is drawn by a dashed line. Let us comment first on the current i_1. Figure 11 shows that the currents obtained by the two different methods are very similar. Differences appear in a short interval after the first time, where i_0 is reached. To analyse this, we consider the right-hand picture in Figure 11, where this part is zoomed. While the line current approaches directly i_0, the dashed line solution first exceeds the value i_0 to later come back to the value of i_0.



(a) Electrical current i_1

(b) Electrical current i_1 by $t = 0.025$

Figure 11: Solutions for the current i_1 of Model 2

In contrast to the optimal currents, there are big differences between the two optimal controls. This is shown in Figure 12. The control drawn as a dashed line stays on the upper bound until the first switching point is reached and moves to the lower bound up to the second switching point. After t_2, the control is fixed at the holding level. In contrast to this, the control drawn as line stays at the upper bound and decreases directly to the holding level.

We observe that the two methods compute extremely different optimal controls, although the optimal values and currents are fairly similar.

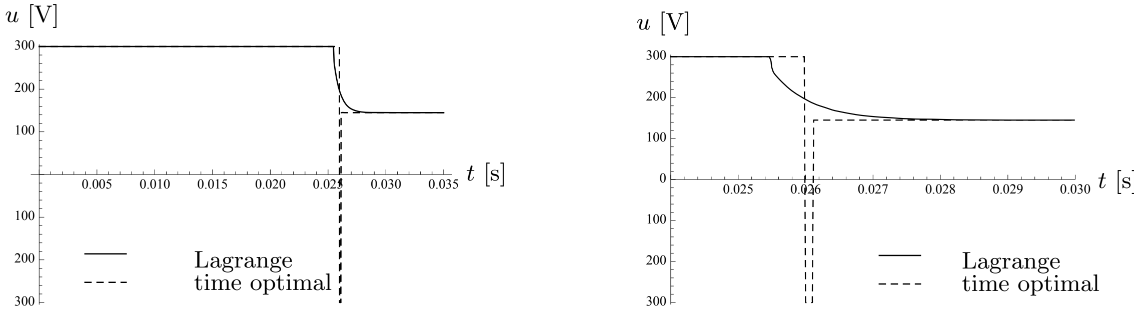
To complete the numerical comparison, we plot the current i_2 in Figure 13.

4 Inductance depending on the electrical current

4.1 The nonlinear control problem

Finally, we investigate cases, where the induction L depends on the electric current; then a nonlinear state equation arises. Having Model 1 in mind, see Figure 1, we define the induction L by

$$L(i) := a + 3\,b\,i^2.$$



(a) Control u for $t \in [0, 0.035)$

(b) Control u for $t \in [0.025, 0.03)$

Figure 12: Comparison of the controls u

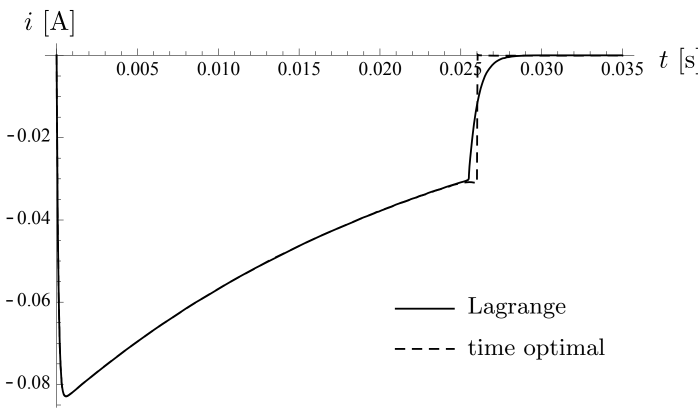


Figure 13: Comparison of the current i_2

Therefore, we derive from (1) the nonlinear equation

$$\left(a + 3\, b\, i^2(t)\right)\, \frac{di}{dt}(t) + R_c\, i(t) = u(t). \tag{41}$$

Because a and b are positive, we can transform this equation to

$$\frac{di}{dt} = -\frac{R_c}{a + 3\, b\, i^2} + \frac{1}{a + 3\, b\, i^2}\, u. \tag{42}$$

The optimal control problem under consideration is now to minimize the objective functional (2) subject to the state equation (41) and the control constraints (3).

4.2 Necessary optimality conditions

Again, the optimality conditions can be derived in two ways. The first is to invoke the Pontryagin maximum principle on the basis of the state equation in the form (42). The second, which we prefer and discuss, is the use of the Lagrange principle. Here, we directly use the equation (41) to define the Lagrange function

$$\begin{aligned}\mathcal{L}(i, u, p) =& \frac{\lambda_T}{2} (i(T) - i_0)^2 + \frac{\lambda_Q}{2} \int_0^T (i(t) - i_0)^2\, dt + \frac{\lambda_U}{2} \int_0^T u(t)^2\, dt \\ & - \int_0^T \left(R_c\, i(t) + \left(a + 3\, b\, i^2(t)\right)\, \frac{di}{dt}(t) - u(t) \right)\, p(t)\, dt.\end{aligned}$$

As in the former sections, the necessary conditions are derived by the derivatives $\partial\mathcal{L}/\partial u$ and $\partial\mathcal{L}/\partial i$. We find

$$\frac{\partial \mathcal{L}(i, u, p)}{\partial u} h = \int_0^T \left(\lambda_U\, u(t) + p(t)\right)\, h(t)\, dt,$$

and, assuming $h(0) = 0$,

$$
\begin{aligned}
\frac{\partial \mathcal{L}(i,u,p)}{\partial i} h =& \lambda_T (i(T) - i_0) - \int_0^T \lambda_Q (i(t) - i_0)\, h(t)\, dt - \int_0^T R_c\, p(t)\, h(t)\, dt \\
& - \int_0^T 6\, b\, i(t)\, \frac{di}{dt}(t)\, h(t)\, p(t)\, dt - \int_0^T (a + 3\, b\, i^2(t))\, \frac{dh}{dt}\, p(t)\, dt \\
=& \lambda_T\, (i(T) - i_0) - \int_0^T \lambda_Q (i(t) - i_0)\, h(t)\, dt - \int_0^T R_c\, p(t)\, h(t)\, dt \\
& - \int_0^T 6\, b\, i(t)\, \frac{di}{dt}(t)\, h(t)\, p(t)\, dt - (a + 3\, b\, i^2(T))\, p(T)\, h(T) \\
& + \int_0^T \frac{d\left(((a + 3\, b i^2))\ p\right)}{dt}(t)\, h(t)\, dt \\
\\
=& \left(\lambda_T (i(T) - i_0) - (a + 3\, b\, i^2)\, p(T)\right)\ h(T) \\
& + \int_0^T -\lambda_Q (i(t) - i_0)\, h(t) - R_c\, p\, h(t) + \left(a + 3\, b\, i^2(t)\right)\ \frac{dp}{dt}(t)\, h(t)\, dt.
\end{aligned}
$$

Since the last expression must vanish for all directions h with initial zero value, we obtain the adjoint equation

$$
\begin{aligned}
-\left(a + 3\, b\, i^2(t)\right)\ \frac{dp}{dt}(t) + R_c\, p &= \lambda_Q\, (i(t) - i_0) \\
\left(a + 3\, b\, i^2(T)\right)\ p(T) &= \lambda_T (i(T) - i_0).
\end{aligned}
$$

Finally, we obtain as in the former sections the optimality system

$$
\begin{aligned}
R_c\, i(t) + \left(a + 3\, b\, i^2(t)\right)\ \frac{di}{dt}(t) &= \mathbb{P}_{[u_a, u_b]}\{-\lambda_U^{-1} p\} \\
-R_c\, p + \left(a + 3\, b\, i^2(t)\right)\ \frac{dp}{dt}(t) &= -\lambda_Q\, (i(t) - i_0) \\
i(0) &= -i_0 \\
\left(a + 3\, b\, i^2(T)\right)\ p(T) &= \lambda_T (i(T) - i_0).
\end{aligned}
\tag{43}
$$

We report on the direct numerical solution of this system in the next subsection.

4.3 Numerical examples

We take the same bounds and constants as in the other models, i. e. $u_a := -300$ and $u_b := 300$. As before, we vary useful parameters λ_Q, λ_U and λ_T. For the known reasons, we define again

$$
J_Q := \int_0^1 (i(t) - i_0)^2 dt.
$$

We have seen that λ_T is less important and set $\lambda_T = 1000$. Again, we fix $\lambda_U = 1$, vary λ_Q and interchange these roles of the parameters afterwards. So we selected as before reasonable parameters by

$$
\lambda_Q = 10^6 \text{ and } \lambda_U = 10^{-6}.
$$

At first view, there are no significant differences to the other models. However, if we concentrate on the time close to $t = 0$, we see a nonlinear increase of the electrical current i. This is shown in Figure 14 by different λ_U. The behaviour of the control u is similar to the linear case. First, the control stays on the upper bound and next it decays to the holding level, see Figure 17.

Now we consider the induction L. Because L depends on i, we need to analyse L only near the initial time. In Figure 15 we see how the induction increases up to the value 1 and grows up to the holding level. This is caused by

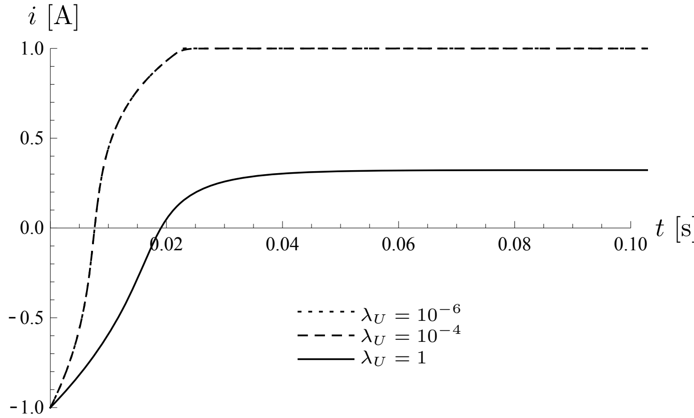


Figure 14: Electrical current i by different λ_U

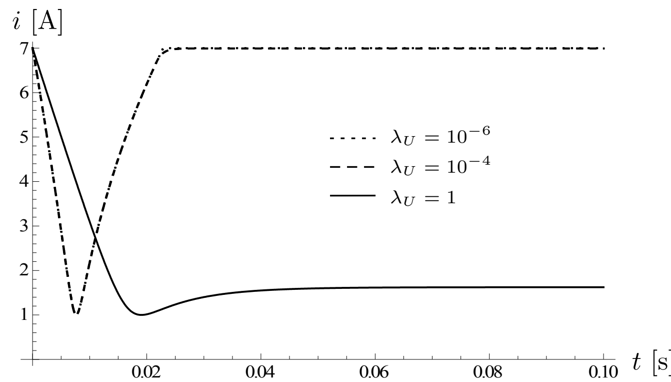


Figure 15: Induction L by different λ_U

the fact that i is starting at $-i_0 = -1$. At the value $i(t) = 0$, the induction must be equal to $L(0) = a + b \cdot 0^2 = 1$ and admit a minimum there. After that point, the induction grows again up to a holding level.

Finally, we compare the case with constant induction of Model 1 (cf. 2.4) and with the last model with time dependent induction. We use the same parameters as before. In Figure 16 we can see the differences in the state i. With constant induction, the electrical current i grows smoothly. In the case with time dependent induction L, the current grows slower at the beginning. After some short time, the current increases faster than by constant induction and reaches earlier the value of i_0. In Figure 17, the curves of the controls u are plotted. Here we can see

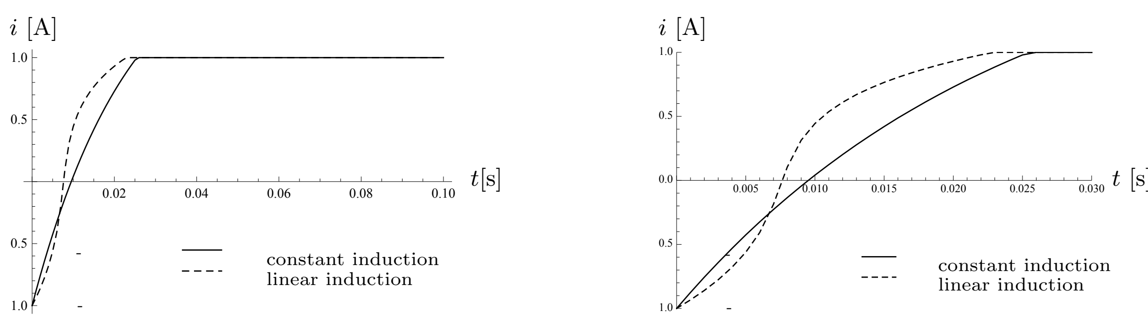


Figure 16: Comparison of the electrical current by different inductions

that in the case with constant induction the control stays longer at the upper bound. However, both curves have the same characteristic behaviour.

The computational time for solving the nonlinear optimal control problem is about 850 seconds (Single Core 3 GHz,4 GB RAM) in contrast to the linear problem of Section 2.4 with about 180 seconds.

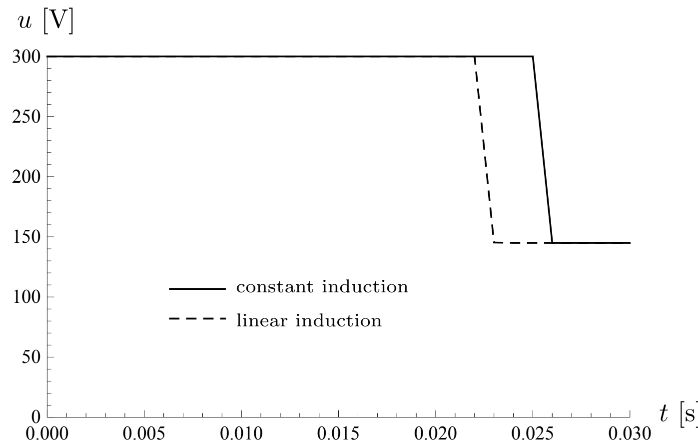


Figure 17: Comparison of the voltage by different inductions

Remark 5. *By our numerical method, we just solve the first-order optimality system. This system of necessary conditions is not sufficient for local optimality, if a nonlinear state equation models the electric current. However comparing the results for the nonlinear equation with those for the linear one, we can be sure to have computed the minimum. Another hint on optimality is that the current reaches the desired state in short time.*

Conclusions. We have considered a class of optimal control problems related to the control of electric circuits. It is shown that, by an approiate choice of parameters, the time optimal transfer of an initial state to another one can be reduced to a fixed time control problem with quadratic objective functional. Invoking the kown theory of first-order necessary optimality conditions, a non-smooth optimality system is set up, which can be directly solved by available software for differential equations. This direct and slightly non-standard numerical approach turned out to be very efficient and robust for this class of optimal control problems.

References

[1] *Software COMSOL Multiphysics* ®. internet reference: www.comsol.com.

[2] M. Athans and P. L. Falb. *Optimal control. An introduction to the theory and its applications.* McGraw-Hill Book Co., New York, 1966.

[3] G. Feichtinger and R. F. Hartl. *Optimale Kontrolle ökonomischer Prozesse.* Walter de Gruyter & Co., Berlin, 1986. Anwendungen des Maximumprinzips in den Wirtschaftswissenschaften.

[4] E. Hairer, S. P. Nørsett, and G. Wanner. *Solving ordinary differential equations. I, Nonstiff problems*, volume 8 of *Springer Series in Computational Mathematics.* Springer-Verlag, Berlin, second edition, 1993.

[5] R. F. Hartl, S. P. Sethi, and R. G. Vickson. A survey of the maximum principles for optimal control problems with state constraints. *SIAM Rev.*, 37(2):181–218, 1995.

[6] A. D. Ioffe and V. M. Tihomirov. *Theory of extremal problems*, volume 6 of *Studies in Mathematics and its Applications.* North-Holland Publishing Co., Amsterdam, 1979.

[7] J. Jahn. *Introduction to the Theory of Nonlinear Optimization.* Springer-Verlag, Berlin, 1994.

[8] D. G. Luenberger. *Optimization by Vector Space Methods.* Wiley, New York, 1969.

[9] J. W. Macki and A. Strauss. *Introduction to optimal control theory.* Springer-Verlag, New York, 1982. Undergraduate Texts in Mathematics.

[10] I. Neitzel, U. Prüfert, and T. Slawig. Strategies for time-dependent PDE control with inequality constraints using an integrated modeling and simulation environment. *Numer. Algorithms*, 50(3):241–269, 2009.

[11] E. R. Pinch. *Optimal Control and the Calculus of Variations.* Oxford Science Publications, 1993.

[12] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes.* A Pergamon Press Book. The Macmillan Co., New York, 1964.

[13] Stanisław Rosłoniec. *Fundamental Numerical Methods for Electrical Engineering.* Springer Publishing Company, Incorporated, 2008.

[14] F. Tröltzsch. *Optimal Control of Partial Differential Equations. Theory, Methods and Applications.* Graduate Studies in Mathematics, Volume 112. American Mathematical Society, Providence, 2010.