Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 15, No. 1-2/2023

ON THE BANG-BANG PRINCIPLE FOR PARABOLIC OPTIMAL CONTROL PROBLEMS*

F. Tröltzsch†

DOI https://doi.org/10.56082/annalsarscimath.2023.1-2.286

Dedicated to Dr. Dan Tiba on the occasion of his 70^{th} anniversary

Abstract

Optimal control problems for the linear heat equation with final observation and pointwise constraints on the control are considered, where the control depends only on the time. It is shown that to each finite number of given switching points, there is a final target such that the optimal objective value is positive, the optimal control is bang bang, and has the desired switching structure. The theory is completed by numerical examples.

MSC: 49K20, 49K30, 49M05

keywords: optimal control, heat equation, bang-bang principle, finitely many switching points

1 Introduction

In this paper, we discuss the construction of parabolic optimal control problems with time-dependent control, such that the optimal control has a desired switching structure. Our main result is that, for each given set of

*Accepted for publication on April 19-th, 2023

†troeltzsch@math.tu-berlin.de, Institut für Mathematik der Technischen Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany

real numbers $0 < s_1 < \ldots < s_k < T$, there is a target state such that the optimal control of the boundary control problem below has the switching points $s_1, \ldots, s_k$. Even though the method is constructive, its numerical application is limited to a small number k of switching points. We present numerical examples for $k = 2$, 3.

In the first and main part of the paper, we consider the spatially one-dimensional optimal boundary control problem

$$\min \int_0^1 |y(x,T) - y_\Omega(x)|^2 dx \tag{1}$$

subject to

$$\begin{aligned} \partial_t y(x,t) - \partial_{xx} y(x,t) &= 0 && \text{in } (0,1) \times (0,T) \\ \partial_x y(0,t) &= 0 && \text{in } (0,T) \\ \partial_x y(1,t) + \lambda\, y(1,t) &= u(t) && \text{in } (0,T) \\ y(x,0) &= 0 && \text{in } (0,1), \end{aligned} \tag{2}$$

and

$$|u(t)| \le 1 \quad \text{a.e. in } (0,T). \tag{3}$$

Here, $T > 0$ and $\lambda > 0$ are given constants, while $y_\Omega \in L^2(0,1)$ is the given target state. Let us set for convenience $\Omega := (0,1)$.

Remark 1 *For simplicity, the heat equation is given with homogeneous right-hand side and initial data. All results remain true for the inhomogeneous equation $\partial_t y - \partial_{xx} y = e$ with initial condition $y(\cdot, 0) = y_0$, where $e \in L^2(\Omega \times (0,T))$ and $y_0 \in L^2(\Omega)$. The only difference is that then the control-to-state operator $S : u \mapsto y$ is affine instead of linear. The necessary optimality conditions remain true without any change.*

Our main question in the first part is the following: How can we construct the target state y_Ω, such that the optimal control has a desired bang-bang structure while the optimal value of the objective functional is positive?

In what follows, we fix $\lambda = 0$ so that the state equation (2) has pure Neumann boundary conditions. The case $\lambda > 0$ will be needed for some references only. Thanks to the simple Neumann boundary condition, the weak solution $y \in W(0,T) = \{y \in L^2(0,T; H^1(\Omega)) : \partial_t y \in L^2(0,T; H^1(\Omega)')\}$ is given by the following Fourier expansion:

$$\begin{aligned} y(x,t) &= 1 + 2\sum_{n=1}^{\infty} (-1)^n \cos(n\pi x) \int_0^t e^{-n^2\pi^2(t-s)} u(s)\,ds \\ &= \int_0^t G(x,1,t-s)u(s)\,ds, \end{aligned}$$

where $G : \Omega \times \Omega \times [0, \infty) \to \mathbb{R}$ is the Green's function

$$G(x, \xi, t) = 1 + 2\sum_{n=1}^{\infty} \cos(n\pi x) cos(n\pi\xi) e^{-n^2\pi^2 t}.$$

For this well-known fact, we refer e.g. to [30], Section 3.2.2, where also the Green's function for $\lambda > 0$ is contained.

2 Necessary optimality conditions

It is easy to show by a standard weak compactness argument that the optimal control problem (1)-(3) has at least one optimal control. Let $\bar{u}$ be an optimal control with associated state $\bar{y}$ and adjoint state $\bar{\varphi}$ (throughout the paper, optimality is indicated by a bar).

The adjoint equation for $\bar{\varphi}$ is

$$\begin{aligned} -\partial_t \varphi(x,t) - \partial_{xx}\varphi(x,t) &= 0 && \text{in } (0,1) \times (0,T) \\ \partial_x \varphi(0,t) &= 0 && \text{in } (0,T) \\ \partial_x \varphi(1,t) &= 0 && \text{in } (0,T) \\ \varphi(x,T) &= \bar{y}(x,T) - y_\Omega(x) && \text{in } (0,1); \end{aligned} \tag{4}$$

notice that we concentrate on the case $\lambda = 0$. The optimal control $\bar{u}$ must obey the variational inequality

$$\int_0^T \bar{\varphi}(1,t)(u(t) - \bar{u}(t))\, dt \geq 0 \quad \forall\, |u(\cdot)| \leq 1. \tag{5}$$

A simple pointwise discussion yields the equivalent pointwise condition

$$\bar{u}(t) = -\text{sign}\, \bar{\varphi}(1,t) \text{ a.e. in } (0,T). \tag{6}$$

Each of the conditions (5) and (6) is necessary and sufficient for the optimality of $\bar{u}$. The sufficiency follows from the convexity of the problem (1)-(3).

In view of (6), for given $\bar{\varphi}$, the optimal control is uniquely determined in the points t, where $\varphi(1,t) \neq 0$, while (6) does not provide explicit information on $\bar{u}$ in the set $\{t \in [0,T] : \bar{\varphi}(1,t) = 0\}$. It is interesting and well known that only two cases can happen:

In the first, $\bar{\varphi}$ is the zero function. In this case, the target y_Ω is reached by the optimal state, i.e. we have $\|\bar{y}(\cdot,T) - y_\Omega\|_{L^2(0,1)} = 0$.

In the second, it holds $\|\bar{y}(\cdot,T) - y_\Omega\|_{L^2(0,1)} > 0$, the target cannot be reached under the constraints (3). Then the function $t \mapsto \bar{\varphi}(1,t)$ can have at most countably many zeros that may accumulate only at $t = T$. In other words, in any subinterval $[0, T - \tau]$ with $0 < \tau < T$, the function $t \mapsto \bar{\varphi}(1,t)$ has only finitely many zeros. The proof is based on the fact that $t \mapsto \bar{\varphi}(1,t)$ can be extended to a holomorphic function on the complex set $\{z \in \mathbb{C} : Re(z) < T\}$; we refer to [12].

Definition 1 (Switching point) *If the function $t \mapsto \bar{\varphi}(1,t)$ changes its sign in $s \in (0,T)$, then s is called a switching point of $\bar{u}$.*

We recall the celebrated bang-bang principle:

Theorem 1 (Bang-Bang Principle) *Let $\bar{u}$ be an optimal control of the parabolic boundary control problem* (1)-(3) *with associated state $\bar{y}$ and associated adjoint state $\bar{\varphi}$. If $\|\bar{y}(\cdot,T) - y_\Omega\|_{L^2(0,1)} > 0$, then $|\bar{u}(t)| = 1$ holds a.e. on $(0,T)$. The optimal control has at most countably many switching points that can accumulate only at $t = T$. All switching points are zeros of the function $t \mapsto \bar{\varphi}(1,t)$, $t \in [0,T]$.*

For the proof, the reader is referred to [12], one of the earliest references for this result. The theorem and the proof can also be found in [30, Section 3.2.4]. The bang-bang principle is also known for more general distributed and boundary control problems for partial differential equations in higher dimension, see also Section 6. A short survey is given in Section 7. In the control theory for ordinary differential equations, this principle has a much longer history, we refer only to the textbook [18].

For given y_Ω, the following simple question arises: Does the associated optimal control have finitely or infinitely many switching points? There is an example in [24] for problem (1)-(3) with $\lambda = 1$: The target function is

$$y_\Omega(x) = \frac{1}{2}(1 - x^2).$$

Numerical computations deliver an optimal control with exactly one switching point. Is this the only one? In [6], this is confirmed by quite tedious estimates. Students of the author tested the function $y_\Omega(x) = \frac{1}{2}(1 - x)$. Here, more switching points were computed. Is their number finite?

To the best knowledge of the author, there is no general result that answers the question whether, for a given y_Ω, the number of switching points of the associated optimal control is finite or not. The author does also not have such a result.

Instead, in this paper we reverse the question: How to construct the target y_Ω such that the optimal control $\bar{u}$ has finitely many and prescribed switching points $s_1, \ldots, s_k$. It will be shown that to each positive integer k and each set of real numbers $0 < s_1 < \ldots < s_k < T$, a target state y_Ω exists such that the associated optimal control has switching points exactly in $s_1, \ldots, s_k$. This method can be numerically implemented for small k.

Currently, there is an active research on optimal control problems for PDEs without Tikhonov regularization term, where optimal controls are expected to be bang-bang. Under suitable assumptions on the zero set of the adjoint states, the stability of optimal controls with respect to certain perturbations of the problem can be shown and finite element error estimates become possible, cf. [3, 5]. Therefore, test examples are useful, where these assumptions are satisfied. This paper is a contribution to this issue for parabolic equations.

To the best knowledge of the author, an example where the optimal control has infinitely many switching points is not yet known. This seems to be a difficult and open problem.

3 Main idea of the construction

We want to construct a target y_Ω, such that the optimal control $\bar{u}$ has the switching points

$$0 < s_1 < \ldots < s_k < T,$$

and

$$\bar{u}(t) = (-1)^{i+1}, \quad t \in [s_{i-1}, s_i), \quad i = 1, \ldots, k+1. \tag{7}$$

Here we have set $s_0 = 0$, $s_{k+1} = T$. Notice that then $\bar{u}$ starts with 1,

$$\bar{u}(t) = 1 \text{ in } [0, s_1). \tag{8}$$

A trivial way of constructing a problem with a given $\bar{u}$ as optimal solution is obtained by setting $y_\Omega = y_{\bar{u}}$, where $y_{\bar{u}}$ denotes the state associated with $\bar{u}$. Then, the adjoint state is identically zero and, according to our definition, switching points do not exist. This is, what we do not have in mind.

From now on, the switching points s_i and their number k are given and $\bar{u}$ is the control defined above, based on $s_1, \ldots, s_k$. The condition (6) is sufficient for optimality of $\bar{u}$. Therefore, having found an y_Ω and a control $\bar{u}$ with adjoint state $\bar{\varphi}$ that obey (6), then $\bar{u}$ is optimal.

We proceed as follows: First, we define

$$y_\Omega^0 = y_{\bar{u}}(\cdot, T). \tag{9}$$

Then we have $y_{\bar{u}}(\cdot,T) - y_\Omega^0 = 0$. Next, we set

$$y_\Omega = y_\Omega^0 + z, \tag{10}$$

where z is to be constructed. Associated with z, we define ψ_z as the unique solution of the auxiliary adjoint equation

$$\begin{array}{rcll} -\partial_t\psi(x,t) - \partial_{xx}\psi(x,t) & = & 0 & \text{in } (0,1)\times(0,T) \\ \partial_x\psi(0,t) & = & 0 & \text{in } (0,T) \\ \partial_x\psi(1,t) & = & 0 & \text{in } (0,T) \\ \psi(x,T) & = & -z(x) & \text{in } (0,1). \end{array} \tag{11}$$

In view of (9) and (10), ψ_z obeys the final condition

$$\psi_z(x,T) = -z(x) = y_{\bar{u}}(x,T) - y_\Omega^0(x) - z(x) = y_{\bar{u}}(x,T) - y_\Omega(x). \tag{12}$$

Therefore, we will have $\psi_z = \bar{\varphi}$. To match the sign condition (6) for the given $\bar{u}$, the function z must obey the following properties:

- In $(0,T)$, $\psi_z(1,t) = 0$ if and only if $t \in \{s_1,\dots,s_k\}$,
- sign $\psi_z(1,t) = (-1)^i$ in (s_{i-1},s_i).

The last requirement needs $\psi_z(1,t) < 0$ in $(0,s_1)$.

Moreover, we want to have a switching structure that is stable with respect to certain perturbations of the problem. Therefore, we require

$$\partial_t\psi(1,s_i) \neq 0,\ \ i = 1,\dots k.$$

The auxiliary adjoint state ψ_z is given by the following expansion:

$$\begin{aligned} \psi_z(x,t) &= \int_0^1 G(x,\xi,T-t)(-z(\xi))\,d\xi \\ &= -\int_0^1 \left(1 + 2\sum_{n=1}^\infty \cos(n\pi x)\cos(n\pi\xi)e^{-n^2\pi^2(T-t)}\right) z(\xi)\,d\xi \\ &= -\int_0^1 z(\xi)\,d\xi - \sum_{n=1}^\infty \sqrt{2}\cos(n\pi x)e^{-n^2\pi^2(T-t)} \int_0^1 \sqrt{2}\cos(n\pi\xi)z(\xi)d\xi \\ &= -\sum_{n=0}^\infty (v_n,z)_{L^2(0,1)} v_n(x) e^{-n^2\pi^2(T-t)}, \end{aligned} \tag{13}$$

where

$$v_0 \equiv 1 \quad \textit{and} \quad v_n(x) = \sqrt{2}\cos(n\pi x), \quad n \geq 1,$$

denote the normalized eigenfunctions of ∂_{xx} subject to the given homogeneous Neumann conditions. The system $\{v_n\}_{n\geq 0}$ is complete and orthonormal in $L^2(0,1)$.

In view of these findings, the problem of constructing an optimal control problem with a desired optimal bang-bang control $\bar{u}$ boils down to the construction of a suitable function z. This is the issue of the next section.

4 Construction of z

We use the ansatz

$$z(x) = 1 + \sum_{n=1}^{k} \alpha_n \cos(n\pi x) \tag{14}$$

with $\alpha_1, \ldots, \alpha_k$ to be determined. By orthonormality, we have

$$\int_0^1 1 \cdot z(x)\, dx = 1,$$

and, thanks to $\int_0^1 \cos^2(n\pi x)\, dx = 1/2$,

$$\int_0^1 \cos(n\pi x) z(x)\, dx = \frac{1}{2}\alpha_n.$$

Therefore,

$$\begin{aligned} -\psi_z(1,t) &= \int_0^1 z(\xi)\left(1 + 2\sum_{n=1}^{\infty} \cos(n\pi 1)\cos(n\pi\xi) e^{-n^2\pi^2(T-t)}\right) d\xi \\ &= 1 + 2\sum_{n=1}^{k}(-1)^n \frac{1}{2}\alpha_n e^{-n^2\pi^2(T-t)} \\ &= 1 + \sum_{n=1}^{k}(-1)^n \alpha_n e^{-n^2\pi^2(T-t)}. \end{aligned} \tag{15}$$

Then the conditions $\psi_z(1, s_i) = 0$, $i = 1, \ldots, k$, lead to the following linear inhomogeneous system for $(\alpha_1, \ldots \alpha_k)$:

$$\sum_{n=1}^{k} (-1)^{n+1} \alpha_n e^{-n^2\pi^2(T-s_i)} = 1, \quad i = 1, \ldots, k. \tag{16}$$

Let us define for convenience $\beta_n = (-1)^{n+1}\alpha_n$. Then (16) reads

$$\sum_{n=1}^{k} \beta_n e^{-n^2\pi^2(T-s_i)} = 1, \quad i = 1, \dots k, \tag{17}$$

for the unknowns $\beta_1, \dots, \beta_k$. Notice that the numbers s_i, $i = 1, \dots, k$, are given as desired switching points. The following questions will be answered in the next section:

(i) Does the linear system (17) have a unique solution $(\beta_1, \dots, \beta_k)$?

(ii) If $(\beta_1, \dots, \beta_k)$ is a solution of (17), are $s_1, \dots, s_k$ the only zeros of the function $t \mapsto \sum_{n=1}^{k} \beta_n e^{-n^2\pi^2(T-t)}$?

(iii) Is the derivative of this function nonzero in any of the points $s_1, \dots, s_k$?

Prior to the discussion of existence and uniqueness of $(\beta_1, \dots, \beta_k)$, let us briefly consider the determinant of the system matrix of (17) and define

$$\lambda_i = e^{-\pi^2(T-s_i)}.$$

Then the system (17) can be written in the equivalent form

$$\sum_{n=1}^{k} \beta_n \, \lambda_i^{n^2} = 1, \quad i = 1, \dots, k.$$

The determinant of the matrix $A = [\lambda_i^{n^2}]$, $i, n = 1, \dots, k$, is the generalized Vandermonde determinant

$$\det A = \begin{vmatrix} \lambda_1 \; \lambda_1^4 \;\; \lambda_1^9 \dots \; \lambda_1^{k^2} \\ \lambda_2 \; \lambda_2^4 \;\; \lambda_2^9 \dots \; \lambda_2^{k^2} \\ \dots \\ \lambda_k \; \lambda_k^4 \;\; \lambda_k^9 \dots \; \lambda_k^{k^2} \end{vmatrix}. \tag{18}$$

Many papers were contributed to generalized Vandermonde determinants. Nevertheless, it is difficult to establish a general solution formula for the determinant above as it is known for the classical Vandermonde determinant. While this is fairly easy for $k \leq 4$, it becomes difficult for larger k.

Therefore, to show that the linear system (17) is uniquely solvable, we apply a different but fairly standard method. It was used in [18] for the proof of bang-bang properties of time optimal controls for linear ordinary differential equations. We finally obtain that the generalized Vandermonde determinant (18) does not vanish. This is done in the next section.

5 Discussion of the questions (i)-(iii)

First, we consider the well-posedness of the system (17), i.e. we answer question (i). Let us assume that the system (17) were not uniquely solvable. Then the associated homogeneous system has at least one non-trival solution, i.e. there is a nonzero vector $(\beta_1, \ldots, \beta_k)$, such that

$$\sum_{n=1}^{k} \beta_n e^{-n^2\pi^2(T-s_i)} = 0, \quad i = 1, \ldots, k. \tag{19}$$

Lemma 1 *For every nonzero vector $(\beta_1, \ldots, \beta_k) \in \mathbb{R}^k$, and any sequence of real numbers $\sigma_1 < \ldots < \sigma_k$ the function $t \mapsto \sum_{n=1}^{k} \beta_n e^{-\sigma_n(T-t)}$ has at most $k-1$ zeros.*

Proof. We proceed by induction. For $k = 1$ and $\beta_1 \neq 0$, the function $t \mapsto \beta_1 e^{-\sigma_1(T-t)}$ does not vanish, hence there is no zero. This confirms the claim for $k = 1$.

Assume now that the claim is true up to $k = m$, $m \in \mathbb{N}$, but not for $k = m+1$. Then there are a non-zero vector $(\beta_1, \ldots, \beta_{m+1})$ and a sequence $\sigma_1 < \ldots < \sigma_{m+1}$ such that the function $t \mapsto \sum_{n=1}^{m+1} \beta_n e^{-\sigma_n(T-t)}$ has at least $m+1$ zeros.

Then we multiply this function by $e^{\sigma_1(T-t)}$ that does never vanish. The resulting function

$$f(t) = \beta_1 + \sum_{n=2}^{m+1} \beta_n e^{-(\sigma_n - \sigma_1)(T-t)}$$

has the same zeros as before $\sigma_1 < \ldots < \sigma_{m+1}$. By the theorem of Rolle, the derivative f' must have at least m zeros. We have

$$f'(t) = \sum_{n=2}^{m+1} \beta_n(\sigma_n - \sigma_1) e^{-(\sigma_n - \sigma_1)(T-t)} = \sum_{n=1}^{m} \tilde{\beta}_n e^{-\tilde{\sigma}_n(T-t)}$$

with numbers $\tilde{\beta}_n$, $\tilde{\sigma}_n$ that satisfy the assumption of the induction. The sum has m summands and is covered by our claim for $k = 1, \ldots, m$. So it should have at most $m-1$ zeros. This is a contradiction. □

Lemma 2 *For each positive $k \in \mathbb{N}$ and all numbers $s_1 < \ldots < s_k$, the inhomogeneous linear system (17) has a unique solution $(\beta_1, \ldots, \beta_k)$.*

Proof. Let us assume that (19) has a nontrivial solution. Then the function $t \mapsto \sum_{n=1}^{k} \beta_n e^{-n^2\pi^2(T-t)}$ has the zeros $s_1 < \ldots < s_k$. In view of Lemma 1, this is only possible for $\beta_1 = \ldots = \beta_k = 0$ in contrary to the assumption. Therefore the homogeneous system (19) has only the trivial solution and its system matrix $A = [e^{-n^2\pi^2(T-s_i)}], i, n = 1, \ldots, k$, is regular. Consequently, the linear system (17) has a unique solution $(\beta_1, \ldots, \beta_k)$.

□

Let now $(\beta_1, \ldots, \beta_k)$ be the unique solution of (17). Then the function $t \mapsto -1 + \sum_{n=1}^{k} \beta_n e^{-n^2\pi^2(T-t)}$ has the zeros $s_1, \ldots, s_k$. We show that its derivative in these points does not vanish.

Lemma 3 *Let* $(\beta_1, \ldots, \beta_k)$ *be the unique solution of the system* (17). *Then the derivative of the function* $f : t \mapsto -1 + \sum_{n=1}^{k} \beta_n e^{-n^2\pi^2(T-t)}$ *does not vanish in its zeros* $s_1, \ldots, s_k$. *Therefore,* f *changes the sign in* $s_1, \ldots, s_k$.

Proof. Assume that in at least one of the zeros, say in s_j, the function f' vanishes,

$$f'(s_j) = 0.$$

We have

$$f'(t) = \sum_{n=1}^{k} \beta_n n^2 \pi^2 e^{-n^2\pi^2(T-t)}. \tag{20}$$

Thanks to the theorem of Rolle, f' has at least $k-1$ zeros located strictly between the zeros $s_1, \ldots, s_k$. Moreover, by the assumption of the proof, it has the additional zero s_j. Therefore, f' would have at least k zeros. This is a contradiction to Lemma 1. □

The proof of the previous Lemma also answers question (ii):

Lemma 4 *Let* $(\beta_1, \ldots, \beta_k)$ *be the unique solution of* (17). *Then the points* $s_1, \ldots, s_k$ *are the only zeros of the function* $f : t \mapsto -1 + \sum_{n=1}^{k} \beta_n e^{-n^2\pi^2(T-t)}$.

Proof. Assume that f has an additional zero s^*. Then its derivative f', given by (20) would have at least k zeros. However, thanks to Lemma 1, f' can only have $k-1$ zeros (notice that the form of f' fits to Lemma 1). This is a contradiction. □

Summary of the construction. For given desired numbers $0 < s_1 < \ldots s_k < T$, we want to find a desired state y_Ω such that the optimal boundary control $\bar{u}$ of (1)-(3) has the form (6). Then we proceed as follows:

1. Compute the vector $(\alpha_1, \ldots, \alpha_k)$ as the solution of the inhomogeneous linear system (16). Thanks to Lemma 2, this vector exists, is unique and nonzero.

   (Lemma 2 ensures this for $(\beta_1, \ldots, \beta_k)$; we set $\alpha_n = (-1)^{n+1}\beta_n$, $n = 1, \ldots, k$.)

2. Define z by (14).

   (Then the solution ψ_z of the adjoint equation (11) is given by the expansion (13) and its negative trace at the boundary $x = 1$ has the form (15). By construction, the function $f : t \mapsto \psi_z(1,t)$ has the zeros $s_1, \ldots, s_k$. Thanks to Lemma 4, these are the only zeros of $t \mapsto \psi_z(1,t)$ in $[0,T]$. Moreover, Lemma 3 ensures that $\partial_t\psi_z(1,s_i) \neq 0$, $i = 1, \ldots, k$. Since $s_1, \ldots, s_k$ are the only zeros of $\psi_z(1,t)$, this function has constant sign in (s_{i-1}, s_i) and alternates the sign in the zeros.)

3. Compute $y_\Omega^0 = y_{\bar{u}}(\cdot,T)$ and fix y_Ω by

$$y_\Omega = y_\Omega^0 + z.$$

4. If $\psi_z(1,t) > 0$ on $(0,s_1)$, then substitute $-z$ for z in (14).

   (Otherwise, $\bar{u}$ would not obey (8).)

The adjoint state associated with $\bar{u}$ is $\bar{\varphi}$ that solves the adjoint equation (4) with terminal data $y_{\bar{u}}(\cdot,T) - y_\Omega$. By (12), the auxiliary adjoint state ψ_z solves the same equation with the same terminal datum, hence we have $\bar{\varphi} = \psi_z$. Therefore, $\bar{u}$ obeys the necessary optimality conditions with the constructed y_Ω. Because the conditions are also sufficient, $\bar{u}$ is optimal.

This construction is numerically implementable, at least for small numbers k. For large n, the terms $e^{-n^2\pi^2(T-t)}$ become very small for $t < T$. This is an obstacle for the numerical use. However, the construction provides the following existence result:

Theorem 2 *For each positive integer k and every selection of numbers $0 < s_1 < \ldots < s_k < T$, there is a desired target $y_\Omega \in C[0,1]$, such that the optimal control $\bar{u}$ of the problem (1)-(3) is bang-bang with switching points $s_1, \ldots, s_k$ and has the form (7),(8). In view of the bang-bang property, the optimal control is unique.*

The proof of uniqueness is standard, we refer to [18] or [30].

Example 1 *We select $T = 1$, $k = 2$, $s_1 = 3/4$, $s_2 = 7/8$. Then the system* (16) *is equivalent to*

$$\begin{aligned} \alpha_1 e^{5-\frac{\pi^2}{4}} - \alpha_2 e^{5-\pi^2} &= e^5 \\ \alpha_1 e^{-\frac{\pi^2}{8}} - \alpha_2 e^{-\frac{\pi^2}{2}} &= 1. \end{aligned} \tag{21}$$

To avoid cancellations, we multiplied the first equation of (16) *by e^5.*

The numerical solution of (21) *is $\alpha_1 = 12.00339596$, $\alpha_2 = 346.9945248$. The associated adjoint function $\psi_z(1,t)$ is illustrated in Fig. 1*

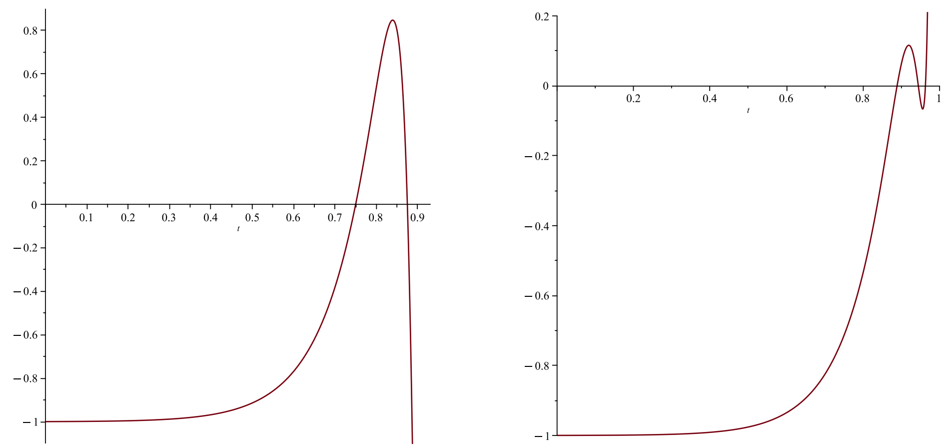

Figure 1: Function $t \mapsto \psi_z(1,t)$ for Example 1 (left) and Example 2 (right)

To complete the example, we have to set up $\bar{u}$ by (7)*, to compute $y_{\bar{u}}$ by solving the state equation* (2) *and to set $y_\Omega = y_{\bar{u}}(x,T) + z(x)$. Then the solution of* (1)-(3) *is $\bar{u}$.*

The need to solve the partial differential equation (2) *for the given $\bar{u}$ is certainly a disadvantage, because it is the source of an additional numerical error. In Example 4, we use a slightly different idea avoiding that problem.*

Theorem 1 allows the existence of infinitely many switching points. The question arises, if there exists an optimal control problem of the form (1)-(3) where the optimal control has infinitely many switching points. This is equivalent to the question of existence of $z \in L^2(0,1)$ such that the function $t \mapsto \psi_z(1,t)$ defined by (11) has infinitely many zeros in $[0,T]$ (accumulating necessarily at $t = T$). To the best knowledge of the author, an associated example or even the proof of existence is not yet known.

Example 2 *We take* $T = 1$*,* $k = 3$*,* $s_1 = 8/9$*,* $s_2 = 17/18$*,* $s_3 = 26/27$*. After equilibrating the linear* 3×3*-system* (16)*, we find* $\alpha_1 = 3.375290066$*,* $\alpha_2 = 10.35044926$*,* $\alpha_3 = 28.36129127$*. The function* $\psi_z(1, t)$ *is presented in the right-hand side of Fig. 1.*

6 Extension to distributed and boundary control in higher dimensions

We considered the simplest 1D boundary control problem, where the control naturally depends on t only. In distributed control problems with $N = dim(\Omega) \geq 1$ or boundary control problems with $N \geq 2$, the control depends on (x, t) so that the concept of switching points does not directly apply. However, it is useful for controls of the form

$$u(x, t) := \sum_{i=1}^{m} w_i(x) u_i(t)$$

with fixed functions $w_i : \Omega \to \mathbb{R}$ (for distributed control) or $w_i : \partial\Omega \to \mathbb{R}$ (for boundary control). Then the switching structure of the controls u_i is of interest. We mention, e.g., Glashoff and Weck [13] for boundary control or the recent paper [22] for time-optimal distributed control. Both papers investigate the switching structure of optimal controls.

In principle, the constructive method of the preceding sections can be extended to such problems. We do not aim at a detailed discussion of this issue. Rather than, let us concentrate on distributed control problems in a smooth bounded domain $\Omega \subset \mathbb{R}^N$, $N \geq 1$, with control

$$u(x, t) := w(x) u(t)$$

with fixed $w \in L^2(\Omega)$ and control $u \in L^\infty(0, T)$. We consider the problem

$$\min \int_\Omega |y(x, T) - y_\Omega(x)|^2 \, dx \tag{22}$$

subject to

$$\begin{aligned} \partial_t y(x,t) - \Delta_x y(x,t) &= w(x) u(t) && \text{in } \Omega \times (0, T) \\ y(x,t) &= 0 && \text{in } \partial\Omega \times (0, T) \\ y(x,0) &= 0 && \text{in } \Omega, \end{aligned} \tag{23}$$

and

$$|u(t)| \leq 1 \quad \text{a.e. in } (0, T). \tag{24}$$

Here, we impose homogeneous Dirichlet boundary conditions on y. The treatment for Neumann or Robin boundary conditions is analogous.

For each $u \in L^\infty(0,T)$ and fixed $w \in L^2(\Omega)$, the initial-boundary value problem (23) has a unique solution $y_u \in W(0,T)$. It has the semigroup representation

$$y(x,t) = \int_0^t \int_\Omega G(x,\xi,t-s)w(\xi)u(s)d\xi\, ds$$

with the Green's function

$$G(x,\xi,t) = \sum_{n=1}^{\infty} v_n(x)v_n(\xi)e^{-\mu_n^2 t}.$$

The numbers μ_n, $n = 1,\ldots,\infty$, are the eigenvalues of the Laplace operator $-\Delta_x$ subject to homogeneous Dirichlet boundary conditions; $\{v_n\}_{n=1,\ldots,\infty}$ is the associated orthonormal system of eigenfunctions in $L^2(\Omega)$. Therefore, we have

$$\begin{aligned} y(x,T) &= \sum_{n=1}^{\infty} v_n(x) \underbrace{\int_\Omega v_n(\xi)w(\xi)\, d\xi}_{w_n} \int_0^T e^{-\mu_n^2(T-s)}u(s)\, ds \\ &= \sum_{n=1}^{\infty} w_n\, v_n(x) \int_0^T e^{-\mu_n^2(T-s)}u(s)\, ds =: (Su)(x), \end{aligned}$$

where $S : L^2(0,T) \to L^2(\Omega)$, $u \mapsto y$, is the linear and continuous control-to-state operator.

Its adjoint operator $S^* : L^2(\Omega) \to L^2(0,T)$ is represented by

$$(S^*d)(t) = \sum_{n=1}^{\infty} w_n\, e^{-\mu_n^2(T-t)} \int_\Omega d(\xi)v_n(\xi)\, d\xi.$$

Switching points of an optimal control $\bar{u}$ with state $\bar{y}$ are now related to the zeros of the function

$$f : t \mapsto \sum_{n=1}^{\infty} w_n\, e^{-\mu_n^2(T-t)} \int_\Omega d(x)v_n(\xi)\, d\xi, \tag{25}$$

where $d(x) = \bar{y}(x,T) - y_\Omega(x)$.

Following the method of Section 4, we set $y_\Omega = y_{\bar{u}}(\cdot,T) + z$ and consider the auxiliary adjoint function

$$\psi_z(t) = (S^*(-z))(t).$$

For z, we use the ansatz

$$z(x) = \sum_{j=1}^{k} \alpha_j v_{n_j}(x),$$

where $k \geq 2$ and $\{v_{n_j}\}$ is a selection of eigenfunctions such that all associated eigenvalues are mutually different. Instead of developing an associated theory that is similar to that of the previous sections, we explain the method by an example.

Example 3 *We consider the case $N = 2$ in the rectangle $\Omega = (0, \pi) \times (0, \pi)$. The normalized eigenfunctions of $-\Delta = -\partial_{x_1}^2 - \partial_{x_2}^2$ are $\{\sin(nx_1)\sin(mx_2)\}$ with associated eigenvalues $n^2 + m^2$, $n, m = 1, \ldots \infty$. With $x = (x_1, x_2)$, the first 4 eigenfunctions and eigenvalues are*

$$\begin{array}{rcll} v_1(x) &=& \sin(x_1)\sin(x_2), & \mu_1 = 2, \\ v_2(x) &=& \sin(2x_1)\sin(x_2), & \mu_2 = 5, \\ v_3(x) &=& \sin(x_1)\sin(2x_2), & \mu_3 = 5, \\ v_4(x) &=& \sin(2x_1)\sin(2x_2), & \mu_3 = 8, \end{array}$$

and the next are $\sin(3x_1)\sin(x_2)$, $\sin(x_1)\sin(3x_2)$ *with eigenvalue* 10. *We refer e.g. to [20] or [28, Sect. 19.5.2].*

We fix $T = 1$, $w(x) = x_1 x_2$, the control having the form $x_1 x_2\, u(t)$. Our aim is to find a desired final state y_Ω such that the optimal control $\bar{u}$ has switching points in $s_1 = 3/4, s_2 = 7/8$, i.e.

$$\bar{u}(t) = \begin{cases} 1, & 0 \leq t \leq 3/4, \\ -1, & 3/4 < t \leq 7/8, \\ 1, & 7/8 < t \leq 1. \end{cases} \tag{26}$$

We use the ansatz

$$z(x) = 1 \cdot v_1(x) + \alpha_1 v_2(x) + \alpha_2 v_4(x),$$

because $\mu_2 = \mu_3$ is a multiple eigenvalue. To establish the linear system for (α_1, α_2), we first need the numbers w_1, w_2, w_4. From

$$\int_0^\pi \xi \sin(n\xi)\, d\xi = \frac{(-1)^{n+1}\pi}{n},$$

we find

$$w_1 = \int_0^\pi \int_0^\pi x_1 x_2 \sin(x_1)\sin(x_2)\, dx = \left(\int_0^\pi r \sin(r) dr\right)^2 = \pi^2$$

and analogously

$$w_2 = -\frac{\pi^2}{2}, \quad w_4 = \frac{\pi^2}{4}.$$

Thanks to $s_1 = 3/4$, $s_2 = 7/8$, *the system for* (α_1, α_2) *is*

$$\begin{aligned} w_2 e^{-25/4}\alpha_1 + w_4 e^{-64/4}\alpha_2 &= -w_1 e^{-4/4} \\ w_2 e^{-25/8}\alpha_1 + w_4 e^{-64/8}\alpha_2 &= -w_1 e^{-4/8}. \end{aligned}$$

We insert the values for w_i, $i = 1, 2, 4$, *and multiply the first equation by* $-4e^8/\pi^2$, *the second by* $-4/\pi^2$. *In this way, we arrive at*

$$\begin{aligned} 2e^{1.75}\alpha_1 - e^{-8}\alpha_2 &= 4e^7 \\ 2e^{-3.125}\alpha_1 - e^{-8}\alpha_2 &= 4e^{-0.5}. \end{aligned}$$

The numerical solution is $\alpha_1 = 383.8524883$, $\alpha_2 = 93317.33956$. *From* (25) *we find that the function* $(S^*z)(t)$ *should have the zeros* 3/4 *and* 7/8. *By orthogonality, we find*

$$(-S^*z)(t) = -\pi^2 e^{-4(1-t)} + \frac{\pi^2}{2} e^{-25(1-t)}\alpha_1 - \frac{\pi^2}{4} e^{-64(1-t)}\alpha_2.$$

In view of the very different growth of the appearing exponential functions, we plot the function $\frac{4}{\pi^2} e^{32(1-t)}(-S^*z)(t)$ *in* $[0.73, 0.9]$ *in Fig. 2. The function* $t \mapsto (-S^*z)(t)$ *is negative on* $[0, 0.75)$, *hence it complies with the requirement* $\bar{u}(t) = 1$ *on* $[0, 0.75)$. *We have*

$$z(x_1, x_2) = \sin(x_1)\sin(x_2) + \alpha_1 \sin(2x_1)\sin(x_2) + \alpha_2 \sin(2x_1)\sin(2x_2). \quad (27)$$

In view of (9) *and* (10), *the desired final target* y_Ω *is*

$$y_\Omega(x_1, x_2) = y_{\bar{u}}(x_1, x_2, T) + z(x_1, x_2),$$

where $y_{\bar{u}}$ *is the state associated with* $\bar{u}(t) = sign\,(S^*z)(t)$.

The construction of Example 3 has one difficulty: y_Ω is not given explicitely. To find $y_{\bar{u}}$, we have to solve the partial differential equation (23) for $\bar{u}$ defined in (26). This is the source of a numerical error that can be avoided by the following trick: Instead of the state equation (23), we consider the optimal control problem for the inhomogeneous equation

$$\begin{aligned} \partial_t y(x,t) - \Delta_x y(x,t) &= e(x,t) + w(x)u(t) && \text{in } \Omega \times (0,T) \\ y(x,t) &= 0 && \text{in } \partial\Omega \times (0,T) \\ y(x,0) &= y_0(x) && \text{in } \Omega \end{aligned} \quad (28)$$

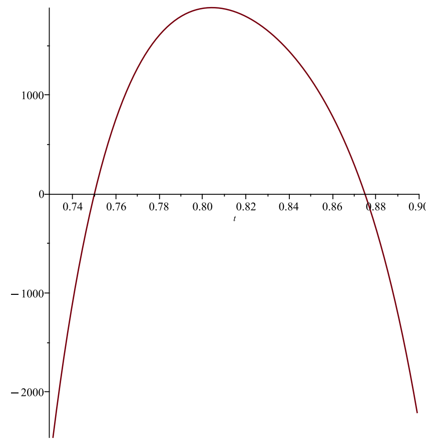

Figure 2: Example 3, function $t \mapsto -\frac{4}{\pi^2} e^{32(1-t)}(S^*z)(t)$ in $[0.73, 0.9]$

with fixed L^2-functions e and y_0. The theory of necessary optimality conditions for the problem (22), (28), (24) remains unchanged, because e and y_0 do neither appear in the adjoint equation nor in the variational inequality. They only influence the computation of $y_{\bar{u}}$. Therefore, we can proceed as follows: In addition to the desired $\bar{u}$, we fix a desired $\bar{y}$. Then we adapt e and y_0 such that $\bar{y}$ and $\bar{u}$ fit together, i.e. $\bar{y} = y_{\bar{u}}$. This avoids the numerical solution of the state equation.

Example 4 *We adopt all data from Example 3, but we consider the distributed optimal control problem of this section with the state equation* (28) *substituted for* (23)*. For the desired optimal state we fix*

$$\bar{y}(x_1, x_2, t) = e^t \sin(x_1)\sin(x_2).$$

Clearly, $\bar{y}$ obeys the homogeneous Dirichlet boundary conditions. To match the initial condition, we have to set $y_0 = \bar{y}(\cdot, 0)$, i.e.

$$y_0(x_1, x_2) = \sin(x_1)\sin(x_2).$$

Finally, we find e by inserting $\bar{y}$ and $\bar{u}$ in (28)*,*

$$e(x_1, x_2, t) = (\partial_t - \Delta)\bar{y}(x_1, x_2, t) - x_1 x_2 \bar{u}(t),$$

where $\bar{u}$ is defined by (26)*. For y_Ω, we obtain*

$$\begin{aligned} y_\Omega(x_1, x_2) &= \bar{y}(x_1, x_2, T) + z(x_1, x_2) \\ &= e^T \sin(x_1)\sin(x_2) + \sin(x_1)\sin(x_2) + \alpha_1 \sin(2x_1)\sin(x_2) \\ &\quad + \alpha_2 \sin(2x_1)\sin(2x_2). \end{aligned}$$

Remark 2 *Compared with the 1D boundary control problem of Section 1, distributed control problems of higher-dimensional boundary control problems are more difficult to investigate. In particular, we can have vanishing coefficients w_n and multiple eigenvalues. Therefore, an extension of the theory of the preceding sections becomes more technical. Here, we do not aim at proving associated counterparts of Theorem 2.*

Boundary control. Boundary control problems for the equation

$$\begin{array}{rcll} \partial_t y(x,t) - \Delta_x y(x,t) & = & 0 & \text{in } \Omega \times (0,T) \\ \partial_\nu y(x,t) + \lambda y(x,t) & = & w(x)u(t) & \text{in } \partial\Omega \times (0,T) \\ y(x,0) & = & 0 & \text{in } \Omega \end{array}$$

can be treated in a similar way (∂_ν denotes the outward normal derivative on $\partial\Omega$). We only mention the integral representation

$$y(x,t) = \int_0^t \int_{\partial\Omega} G(x,\xi,t-s)w(\xi)u(s)dS(\xi)ds,$$

where dS is the surface measure on $\partial\Omega$. Moreover, we refer to [13].

7 Short survey on related references

In the early paper [12], a proof of the bang-bang principle for problems of the type (1)-(3) is given, we refer also to [36] for the minimization of the supremum norm $\|y(\cdot,T) - y_\Omega\|_{L^\infty(\Omega)}$. An extension to higher-dimensional parabolic boundary control problems is discussed in [13]. Numerical methods are investigated in [10, 14, 32]. One of the first numerical examples was presented in [24]. The computed optimal control had one switching point. In [6], it was proved that the optimal control has exactly one switching point, indeed.

If in the problem (1)-(3) the supremum norm $\|y(\cdot,T) - y_\Omega\|_{L^\infty(\Omega)}$ is minimized for given $y_\Omega \in L^\infty(\Omega)$, then the optimal control has a finite number of switching points. This quite spectacular result was proved in [15]. Later, an alternative and slightly simpler proof was given in [11].

The bang-bang principle for time-optimal control problems with evolution equations was discussed in many papers, we only mention [7, 25, 17, 19, 16, 34]. For a certain class of distributed parabolic time-optimal control problems, a result on the number of switching points was recently published in [22].

Many facts on the bang-bang principle for different types of optimal and time-optimal parabolic control problems are presented in the books [8, 9, 27, 33, 21, 30]. The bang-bang principle for semilinear parabolic optimal control problems is discussed in [23, 26, 35].

For certain mixed control-state constraints, a generalized bang-bang principle was shown, first in [29] for the problem (1)-(2) with $\lambda = 1$ and the constraints $0 \leq u(t) \leq c + y(1,t)$. It was extended to other problems in [1, 2, 17].

In [31], bang-bang properties are discussed for sparse optimal controls. Here, the term $\mu \|u\|_{L^1(0,T)}$, $\mu > 0$, is added to the objective functional (1). Another interesting direction of research are second order sufficient optimality conditions for optimal controls of problems with semilinear elliptic equations, where optimal bang-bang controls can be expected, see [3, 5, 4]. They can be applied to the perturbation analysis of associated optimal control problems.

Acknowledgement. The author thanks A.F.M. ter Elst (University of Auckland) and J. Rehberg (WIAS Berlin) for a helpful discussion on the problem of the infinite bang-bang principle.

References

[1] M. Bergounioux, D. Tiba. Some examples of optimality conditions for convex control problems with general constraints. In *Control of partial differential equations and applications (Laredo, 1994)*, volume 174 of Lecture Notes in Pure and Appl. Math., pages 23-30, Dekker, 1996.

[2] J. F. Bonnans, D. Tiba. Control problems with mixed constraints and application to an optimal investment problem. *Math. Rep. (Bucur.)* 11:293-306, 2009.

[3] E. Casas. Second order analysis for bang-bang control problems of PDEs. *SIAM J. Control Optim.* 50:2355-2372, 2012.

[4] E. Casas, K. Kunisch. Optimal control of the two-dimensional evolutionary Navier-Stokes equations with measure valued controls. *SIAM J. Control Optim.* 59:2223-2246, 2021.

[5] E. Casas, D. Wachsmuth, G. Wachsmuth. Sufficient second-order conditions for bang-bang control problems. *SIAM J. Control Optim.* 55:3066-3090, 2017.

[6] V. Dhamo, F. Tröltzsch. Some aspects of reachability for parabolic boundary control problems with control constraints. *Comput. Optim. Appl.* 50:75-110, 2011.

[7] H.O. Fattorini. Time-optimal control of solutions of operational differential equations. *SIAM J. on Control* 2:54-59, 1964.

[8] H.O. Fattorini. *Infinite-dimensional optimization and control theory.* Volume 62 of Encyclopedia of Mathematics and its Applications, Cambridge, 1999.

[9] H.O. Fattorini. *Infinite dimensional linear control systems.* Volume 201 of North-Holland Mathematics Studies, Amsterdam, 2005.

[10] K. Glashoff, E. Sachs. On theoretical and numerical aspects of the bang-bang-principle. *Numer. Math.* 29:93-113, 1977/78.

[11] K. Glashoff. Restricted approximation by strongly sign-regular kernels: the finite bang-bang principle. *J. Approx. Theory* 29:212-217, 1980.

[12] K. Glashoff, W. Krabs. Dualität und Bang-Bang-Prinzip bei einem parabolischen Rand-Kontrollproblem. In *Numerische Behandlung von Variations- und Steuerungsproblemen (Tagung, Sonderforschungsber. 72 "Approximation und Optimierung", Inst. Angew. Math., Univ. Bonn, Bonn, 1974)* Bonn. Math. Schr. 77:1-8, 1975.

[13] K. Glashoff, N. Weck. Boundary control of parabolic differential equations in arbitrary dimensions: supremum-norm problems. *SIAM J. Control Optim.* 14:662-681, 1976.

[14] W. Hackbusch, Th. Will. A numerical method for a parabolic bang-bang problem. In *Optimal control of partial differential equations (Oberwolfach, 1982) Internat. Schriftenreihe Numer. Math.* 68:92-109, Birkhäuser, 1984.

[15] A. Karafiat. The problem of the number of switches in parabolic equations with control. *Ann. Polon. Math.* 34:289-316, 1977.

[16] K. Kunisch and L. Wang. Bang-bang property of time optimal controls of semilinear parabolic equation. *Discrete Contin. Dyn. Syst.* 36:279-302, 2016.

[17] U. Mackenroth. Time-optimal parabolic boundary control problems with state constraints. *Numer. Funct. Anal. Optim.* 3:285-300, 1981.

[18] J. Macki, A. Strauss. *Introduction to Optimal Control Theory.* Berlin, 1982.

[19] V.J. Mizel, T.I. Seidman. An abstract bang-bang principle and time-optimal boundary control of the heat equation. *SIAM J. Control Optim.* 35:1204-1216, 1997.

[20] T. Myint-U. *Partial differential equations of mathematical physics.* New York-London-Amsterdam,1973.

[21] P. Neittaanmäki, D. Tiba. *Optimal control of nonlinear parabolic systems*, volume 179 of Monographs and Textbooks in Pure and Applied Mathematics, New York, 1994.

[22] S. Qin, G. Wang, H. Yu. Switching properties of time optimal controls for systems of heat equations coupled by constant matrices. *SIAM J. Control Optim.* 59:1420-1442, 2021.

[23] E. Sachs. A parabolic control problem with a boundary condition of the Stefan-Boltzmann type. *Z. Angew. Math. Mech.* 58:443-449, 1978.

[24] K. Schittkowski. Numerical solution of a time-optimal parabolic boundary value control problem. *J. Optim. Theory Appl.* 27:271-290, 1979.

[25] E.J.P.G. Schmidt. The "bang-bang" principle for the time-optimal problem in boundary control of the heat equation. *SIAM J. Control Optim.* 18:101-107, 1980.

[26] E.J.P.G. Schmidt. Boundary control for the heat equation with nonlinear boundary condition. *J. Differential Equations* 78:89-121, 1989.

[27] D. Tiba. *Optimal control of nonsmooth distributed parameter systems.* Volume 1459 of Lecture Notes in Mathematics, Berlin, 1990.

[28] H. Triebel. *Analysis und mathematische Physik.* Leipzig, 1981.

[29] F. Tröltzsch. A minimum principle and a generalized bang-bang-principle for a distributed optimal control problem with constraints on the control and the state. *ZAMM* 59:737-739, 1979.

[30] F. Tröltzsch. *Optimal Control of Partial Differential Equations. Theory, Methods and Applications.* American Math. Society, Providence, 2010.

[31] F. Tröltzsch, D. Wachsmuth. On the switching behavior of sparse optimal controls for the one-dimensional heat equation. *Math. Control Relat. Fields* 8:135-153, 2018.

[32] N. von Daniels, M. Hinze. Variational discretization of a control-constrained parabolic bang-bang optimal control problem. *J. Comput. Math.* 38:14-40, 2020.

[33] G. Wang, L. Wang, Y. Xu, Y. Zhang. *Time optimal control of evolution equations*, volume 92 of Progress in Nonlinear Differential Equations and their Applications Birkhäuser/Springer, 2018.

[34] G. Wang and L. Wang. The bang-bang principle of time optimal controls for the heat equation with internal controls. *Systems Control Lett.*, **56** (2007), 709–713.

[35] L. Wang, Q. Yan. Bang-bang property of time optimal null controls for some semilinear heat equation. *SIAM J. Control Optim.* 54:2949-2964, 2016.

[36] N. Weck. Über Existenz, Eindeutigkeit und das "Bang-Bang-Prinzip" bei Kontrollproblemen aus der Wärmeleitung. In *Numerische Behandlung von Variations und Steuerungsproblemen (Tagungsband, Sonderforschungsber. 72 "Approximation und Optimierung", Inst. Angew. Math., Univ. Bonn, Bonn, 1974)* Bonn. Math. Schriften 77: 9-19, 1975.