Optimal Control Problems for the Nonlinear Heat Equation

K. Eppler and F. Tröltzsch

[1] Technische Universität Chemnitz, Fakultät für Mathematik, 09107 Chemnitz
[2] Technische Universität Berlin, Fachbereich Mathematik, Straße des 17. Juni, 10623 Berlin

Abstract. Some aspects of numerical analysis are surveyed for the optimal control of the nonlinear heat equation. In the analysis, special emphasis is on second order sufficient optimality conditions. In particular, the case of pointwise state constraints is adressed. Moreover, a numerical technique of instantaneous control type is presented.

Keywords: Nonlinear heat equation, optimal control, necessary optimality conditions, second order sufficient conditions, pointwise state constraints, instantaneous control
AMS subject classification: 49K20, 49M35, 49M37, 90C10

1 Control problem and Optimality Conditions

The optimal control of heating and cooling processes belongs to the core of optimal control theory of parabolic equations. It covers most of the main difficulties of this theory but is not yet overlaid by the technicalities, which are typical for the optimization of other parabolic systems. Therefore, the study of heat control gives also good insight in the methods for the control of other equations such as Burgers equation [14], [23], fuel ignition models [16], Navier-Stokes equations [9], [10], [13], or phase-field models [11], [12].

We report on some applications of control theory to the optimal cooling of steel profiles, which has already been considered in a sequence of papers [20], [24], [30], [31]. Related issues were discussed in [5], [8]. We present the results of our applied research in our second paper in this volume. Here, we give a brief survey on parts of the theory of optimization in semilinear parabolic equations. In real applications to cooling steel, the equation is quasilinear and the results of the semilinear case cannot be applied. However, the study of semilinear problems provides good information on the effects, which should be expected for quasilinear equations as well. To remain simple in the presentation, we begin our short course with the following optimal control problem:

$$\text{(OC)} \qquad \min \; J(y,u) = \beta/2 \int_0^T \int_\Omega (y(x,t) - y_d(x,t))^2 dx dt \tag{1}$$

$$+ \gamma/2 \int \int_\Omega (y(x,T) - y_\Omega(x))^2 \, dx + \nu/2 \int_0^T \int_\Gamma u(x,t)^2 dS_x dt$$

subject to the heat equation with nonlinear boundary condition

$$
\begin{aligned}
\frac{\partial y}{\partial t} &= \Delta y && \text{in } Q, \\
y(x,0) &= y_0(x) && \text{in } \Omega, \\
\frac{\partial y}{\partial n} &= b(x,t,y,u) && \text{in } \Sigma,
\end{aligned}
\tag{2}
$$

and subject to the control constraints

$$
u_a \le u(x,t) \le u_b \tag{3}
$$

to be satisfied a.e. on Σ. Let us consider the state y as the temperature distribution in the bounded domain $\Omega \subset \mathbb{R}^n$ ($n = 1,2,3$ in the applications), while u is the control, acting on $\Sigma = \Gamma \times (0,T)$, where Γ denotes the boundary of Ω and is supposed to be of class $C^{1,s}$.

The control may have various meanings. For instance, it can denote the outer temperature, it may express the intensity of cooling or heating by some surrounding medium, and it might stand for some energy supply. Let us adopt for a while the first view. Then we search an optimal heating strategy $\bar u = \bar u(x,t)$ such that, starting from the initial temperature y_0, the associated temperature in $Q = (0,T) \times \Omega$ evolves in an optimal way, expressed by the functional J in (1).

Here, $y_d \in L^\infty(Q)$ denotes a desired trajectory of temperature, which has to be followed as closely as possible, and $y_\Omega \in L^\infty(\Omega)$ is a desired final temperature distribution. The constants β, γ are positive weights, and $\nu > 0$ can be interpreted as cost of the control u. Moreover, constant bounds $u_a < u_b$ are given.

We assume $y_0 \in C(\bar\Omega)$. The function $b = b(x,t,y,u)$ is assumed to be of class C^2 with respect to $(y,u) \in \mathbb{R}^2$ and measurable w.r. to $(x,t) \in Q$ (the other variables fixed, respectively). In general, b and its first and second order derivatives must satisfy certain Lipschitz conditions on bounded sets with respect to (y,u) and the partial derivative of b with respect to y, denoted by b_y, is assumed to be nonpositive. We refer, for instance, to [3], [27]. To shorten the presentation and to have direct access to the literature we assume for simplicity that

$$
b_y(x,t,y,u) \le 0 \tag{4}
$$

holds a.e. on $Q \times \mathbb{R}^2$, $b(x,t,y,u)$, $b'(x,t,y,u)$, $b''(x,t,y,u)$ are uniformly bounded on $Q \times \mathbb{R}^2$ and uniformly Lipschitz with respect to (y,u) on $Q \times \mathbb{R}^2$. Here, b' and b'' stand for the gradient and the Hessian matrix of the real function b with respect to $(y,u) \in \mathbb{R}^2$. Then the parabolic problem (2) ist well-posed. The assumptions on second order derivatives are not necessary for this. They are needed to establish second order optimality conditions. In the sequel, we fix constants $p > n+1$, $q > n/2 + 1$ and introduce the state space

$$
Y = \{ y \in W(0,T) \,|\, y_t - \Delta y \in L^q(Q),\ \frac{\partial y}{\partial n} \in L^p(\Sigma),\ y(0) \in C(\bar\Omega) \}.
$$

For the definition of $W(0,T)$ we refer to [25] and the concrete choice in [27]. Y is known to be continuously embedded in $C(\bar{Q})$. Moreover, we define the set of admissible controls $U_{ad} = \{u \in L^\infty(\Sigma) \,|\, u_a \le u(x,t) \le u_b \quad \text{a.e. on } \Sigma.\}$

Theorem 1. *([3],[27]) Let b satisfy the assumptions stated above and let a control $u \in U_{ad}$ be given. Then the parabolic initial boundary value problem (2) has a unique solution $y = y(u)$ in Y. There is a positive constant K such that $\|y(u)\|_{C(\bar{Q})} \le K$ holds uniformly for all $u \in U_{ad}$.*

The next question concerns the solvability of the optimal control problem, i.e. the existence of a globally optimal control $\bar{u}$ with associated optimal state $\bar{y} = y(\bar{u})$. To give a practicable answer, we need an additional property of b.

Theorem 2. *Suppose that*

$$b(x,t,y,u) = b_1(x,t,y) + b_2(x,t,y)u\,, \tag{5}$$

i.e., b is affine-linear with respect to u. Then the optimal control problem (OC) admits at least one (globally) optimal control $\bar{u}$.

The well known proof relies on weak compactness of U_{ad} in $L^p(\Sigma)$, because this permits to select a minimizing subsequence of elements $u_n \in U_{ad}$ such that $u_n \rightharpoonup \bar{u}$ in $L^p(\Sigma)$. By uniform boundedness of $\{y(u_n)\}_{n=1}^\infty$, we can select a subsequence of $b_n(x,t) = b(x,t,u_n,y(u_n))$ converging weakly to some function $\bar{b}$ in $L^p(\Sigma)$. Consequently, we have w.l.o.g. $y_n \to \bar{y}$ in $C(\bar{Q})$. The additional assumption (5) is needed to guarantee that $b(\cdot,\cdot,y_n,u_n) \rightharpoonup \bar{b} = b(\cdot,\cdot,\bar{y},\bar{u})$ so that finally $\bar{y} = y(\bar{u})$ holds.

In the numerical analysis, the consideration of global solutions is not the only way to deal with the problem (OC) . Iterates, generated by numerical algorithms, will in general converge to *local* solutions only. Hence an alternative way is to consider a triplet $(\bar{y},\bar{u},\bar{p})$ that satisfies the first order necessary conditions and to ensure local optimality by second order sufficient conditions.

Theorem 3. *([3], [27]) Let $\bar{u}$ be a locally optimal control of (OC) with associated state $\bar{y} = y(\bar{u})$. Then a unique adjoint state $\bar{p} = \bar{p}(x,t)$ exists in $W(0,T)$ such that the adjoint equation*

$$\begin{aligned} -\frac{\partial p}{\partial t} &= \Delta p + \beta\,(\bar{y} - y_d), \\ p(x,T) &= \gamma\,(\bar{y}(x,T) - y_\Omega(x)), \\ \frac{\partial p}{\partial n} &= b_y(x,t,\bar{y},\bar{u})\,p \end{aligned} \tag{6}$$

is satisfied together with the variational inequality

$$\int_0^T \int_\Gamma (\nu\bar{u} + b_u(\bar{y},\bar{u})\,\bar{p})(u - \bar{u})\,dS\,dt \ge 0 \qquad \forall u \in U_{ad}. \tag{7}$$

This result follows, for instance, from the more general *Pontryagin maximum principle* proved in [27], [3] or directly from [28]. The adjoint state $\bar{p}$ is shown to be in $C(\bar{Q})$. Let us discuss the particular case, where $U_{ad} = L^\infty(\Sigma)$ (unrestricted control) and b satisfies (5). Then (7) implies $\bar{u} = -\nu^{-1} b_2(\cdot, \bar{y})\bar{p}$, and u can be eliminated in (2), (6) to obtain a forward-backward coupled system of two parabolic equations for y and p. This system might be solved, for instance, by the Newton method. It may have multiple solutions.

One of the basic difficulties for the numerical solution is the enormous number of variables the system has after discretization. To give an intuitive estimate for this, assume that $\Omega \subset \mathbb{R}^2$ is the unit square with each edge discretized by 100 node points. Adopt the same simple discretization for the time interval $(0,T)$. Then we have to process $2 \cdot 10^6$ variables. For $\Omega \subset \mathbb{R}^3$ this number increases considerably.

Nevertheless, solving the optimality system (2), (6), (7) for the unconstrained case $U_{ad} = L^\infty(\Sigma)$ is one of the core procedures to solve the constrained case as well.

Formally, Theorem 3 can be derived in the following intuitive way. Define the Lagrange function

$$L = L(y,u,p) = J(y,u) - \int_Q (y_t - \Delta y) p \, dx dt - \int_\Sigma \left(\frac{\partial y}{\partial n} - b(y,u)\right) p \, dS dt. \qquad (8)$$

According to well known Lagrange multiplier rules of mathematical programming in Banach spaces, $(\bar{y}, \bar{u})$ must satisfy, together with $\bar{p}$, the relations

$$L_y(\bar{y}, \bar{u}, \bar{p}) y = 0$$

for all $y \in Y$ with $y(0) = 0$ and

$$L_u(\bar{y}, \bar{u}, \bar{p})(u - \bar{u}) \geq 0 \qquad \forall u \in U_{ad}.$$

After some transformations including integration by parts and Greens formulas, these relations imply (6) and (7).

Assume next that $\bar{u} \in U_{ad}$, $\bar{y} = y(\bar{u})$, and $\bar{p}$ satisfy the optimality system (2), (6), (7). What condition can ensure $\bar{u}$ to be optimal, at least locally? To this end, second order sufficient optimality conditions (SSC) can be invoked. We need for their formulation the second order Fréchet-derivative of L w. r. to $(y,u) \in Y \times L^\infty(\Sigma)$,

$$L''(\bar{y}, \bar{u}, \bar{p})[y,u]^2 = \beta \int_Q y^2 dx dt + \gamma \int_\Omega y^2(\cdot, T)\, dx + \int_\Sigma (\nu u^2 + b_{yy}(\bar{y}, \bar{u})\, \bar{p}\, y^2)\, dS dt.$$

Theorem 4. *(SSC) ([26]) Suppose that $\bar{u} \in U_{ad}$ and $(\bar{y}, \bar{u}, \bar{p})$ satisfy (2),(6),(7). Assume the existence of $\delta > 0$ such that*

$$L''(\bar{y}, \bar{u}, \bar{p})[y,u]^2 \geq \delta \, \|u\|^2_{L^2(\Sigma)}, \qquad (9)$$

holds for all $u \in L^\infty(\Sigma)$, $y \in Y$ satisfying the linearized equation

$$\begin{aligned} \frac{\partial y}{\partial t} &= \Delta y, \\ y(0,x) &= 0, \\ \frac{\partial y}{\partial n} &= b_y(\bar{y},\bar{u})\, y + b_u(\bar{y},\bar{u})\, u. \end{aligned} \tag{10}$$

Then there exist constants $\varepsilon > 0$, $\sigma > 0$ such that the quadratic growth condition

$$J(y,u) \geq J(\bar{y},\bar{u}) + \sigma \|u - \bar{u}\|^2_{L^2(\Sigma)} \tag{11}$$

holds for all $u \in U_{ad}$, $y = y(u)$ such that $\|u - \bar{u}\|^2_{L^\infty(\Sigma)} < \varepsilon$. Hence $\bar{u}$ is locally optimal in the norm of $L^\infty(\Sigma)$.

Remarks: (i) If b admits the form (5), then $L^p(\Sigma)$ can be substituted here for $L^\infty(\Sigma)$. This is an essential advantage, since $\|u - \bar{u}\|_{L^\infty(\Sigma)} < \varepsilon$ requires more or less that jumps of $\bar{u}$, if there are any, must be reproduced by u.

(ii) The second order sufficient conditions can be relaxed by considering active sets, [26]. Then $u = 0$ can be assumed in (10) on so-called strongly active sets.

The theory of (SSC) for problems of the type (OC) is well understood. This refers also to the elliptic case, see [4]. The situation is much more complicated, if state constraints are added. In the case of pointwise state constraints the theory is widely open. For elliptic problems, satisfactory results were obtained in two-dimensional domains Ω, [4], while for parabolic problems the one-dimensional case is considered best, [26].

Let us illustrate by a simple example, where the main difficulty appears. Regard, for instance, (OC) with the additional pointwise state constraint

$$y(x_1,t) - y_2(x_2,t) \leq c \qquad \forall t \in [0,T]\,. \tag{12}$$

Constraints of this type will occur in our application to cooling steel. They are well formulated, since the choice of Y guarantees $y \in C(\bar{Q})$, hence the functions $y(x_i,t)$ are well defined and continuous on $[0,T]$. In the theory of optimality conditions, the state constraint (12) is considered by another Lagrange multiplier μ, which is a monotone increasing function of bounded variation on [0,T]. We have to introduce the extended Lagrange function

$$\tilde{L}(y,u,p,\mu) = L(y,u,p) + \int_0^T (y(x_1,t) - y(x_2,t))\, d\mu(t).$$

The associated theory of first order *necessary* conditions is well developed, see [3], [27]. The main difficulty in proving *sufficient* conditions is the appearence of measures like $d\mu$ extending the right hand side of the adjoint equation (6). This makes the adjoint $\bar{p}$ state less regular. Therefore, in the general case we do not have the important property $p \in L^\infty(Q)$, which is helpful to estimate $L''(\bar{y},\bar{u},\bar{p},\bar{\mu})[y,u]^2$ with respect to (y,u) in the appropriate norms.

2 Numerical methods

The numerical solution of optimal control problems for semilinear elliptic and parabolic equations has made considerable progress. Various methods were discussed, and the numerical results provide essential contributions to the fast developing field of large scale optimization. To give the reader an access to further study, we quote [1,2,6,7,11,12,15,17–19,21,22].

Elliptic problems in two-dimensional domains Ω and parabolic problems in domains of dimension one can be solved in short time, since the number of variables after discretization of the problem is still moderate. If the dimension of Ω is larger than one, then the solution of parabolic problems is still time consuming. However, they can be treated succesfully. For the solution of parabolic problems in two-dimensional domains we refer to [6], [11]. One of the favorite techniques is that of (S)equential (Q)uadratic (P)rogramming. Let us briefly describe the main idea for the classical SQP method, which reduces the solution of the nonlinear problem (OC) to a sequence of quadratic optimal control problems with *linear* equation.

Suppose that (y_n, u_n, p_n) has already been determined. Then the next iterate $(y, u) = (y_{n+1}, u_{n+1})$ is found as the solution of the linear-quadratic problem

$$\text{(QP)} \qquad \min \quad J'(y,u)[y-y_n, u-u_n] + \frac{1}{2} L''_{(y,u)}(y_n, u_n, p_n)[y-y_n, u-u_n]^2$$

subject to $u \in U_{ad}$ and

$$\begin{aligned} \frac{\partial y}{\partial t} &= \Delta y, \\ y(x,0) &= y_0(x), \\ \frac{\partial y}{\partial n} &= b_y(y_n,u_n)(y-y_n) + b_u(y_n,u_n)(u-u_n) + b(y_n,u_n)\,. \end{aligned}$$

The new Lagrange multiplier p_{n+1} is obtained from (6), where (y_{n+1}, u_{n+1}) is substituted for $(\bar{y}, \bar{u})$. Under natural assumptions, among them second order sufficient conditions are most essential, this method locally converges q-quadratically to $(\bar{y}, \bar{u}, \bar{p})$, if considered in the infinite dimensional setting [6], [7], [29]. For instance, the second order assumptions (2), (6), (7), (9), (10) of Theorem 4 can be used for this purpose. For semilinear elliptic equations, the convergence analysis was presented in [32]. Discretizing the problem, various approximation errors influence the performance of the method. Moreover, modifications of the standard SQP method can be numerically more effective.

Our computational experience shows that the SQP method converges very fast, i.e. only a few steps are needed to gain high precision. However, each single step of the method can be very expensive, in particular for domains of higher dimension. If the parabolic equation is quasilinear rather than semilinear, then the situation is even more complicated.

Therefore, in our problem of cooling steel we did not apply the SQP method. First we applied a method of feasible directions, [20], [24], [30]. Later, a suboptimal technique was implemented – a method of instantaneous control type. These techniques are considerable cheaper than SQP methods and have been sucessfully applied to find suboptimal solutions in the control of fluid flows. We only quote Hinze [13] and refer the reader to the extensive references therein. We also mention [14] for the case of the Burgers equation.

As a preparation of our report on optimal cooling of steel in this volume, here we explain the main idea of our technique for the following simplified control problem with state constraints. Let points $x_i \in \Omega$ be given fixed, $i = 1, 2, 3$, and assume that $\Gamma = \cup_{i=1}^{P} \bar{\Gamma}_i$, where $\{\Gamma_i\}_{i=1}^{P}$ is a partition of Γ into pairwise disjoint relatively open subsets. Moreover, consider an equidistant partition of $[0, T]$ into subintervals $I_k = (t_{k-1}, t_k]$, $k = 1(1)K$, $0 = t_0 < t_1 < \ldots < t_{K-1} < t_K = T$. Define $\Sigma_{ik} = \Gamma_i \times I_k$. The partition of Γ and $[0, T]$ into subsets should not be viewed as a result of discretization. In our application, it reflects the associated *technical construction.* In cooling steel, Γ_i is the zone influenced by spray nozzle i, and the time interval I_k is associated with passing the cooling segment k. The control function $u = u(x, t)$ is assumed to be constant on Σ_{ik}, i.e. $u(x, t) = u_{ik}$ on Σ_{ik}. Our simplified control problem "steel" is

(OCS) $$\min \quad y(x_0, T)$$

subject to

$$\begin{aligned} \frac{\partial y}{\partial t} &= \Delta y && \text{in } Q, \\ y(x,0) &= y_0(x) && \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= u_{ik}\, \alpha(x,y)[y_{fl} - y] && \text{in } \Sigma_{ik}, \end{aligned} \tag{13}$$

$$y(x_1, t_k) - y(x_2, t_k) \le c\,,\; k = 1(1)K,$$

$$0 \le u_{ik} \le 1, \qquad i = 1(1)P,\; k = 1(1)K.$$

In this setting, $\alpha = \alpha(x, y)$ is the heat exchange coefficient and y_{fl} is the temperature of the cooling fluid. We assume that α is sufficiently smooth with respect to $y \in \mathbb{R}$. The main idea of instantaneous control is as follows: First minimize $y(x_0, t_1)$, i.e. find optimal controls $\bar{u}_{i1}$ on the short time horizon $[t_0, t_1]$. Next insert $y(x, t_1)$ as a new initial condition in (13), to optimize next the process on $[t_1, t_2]$. In this way, we have to solve K single optimal short horizon control problems with P control variables $u_{1k}, \ldots, u_{Pk}$, each. However, the problems are nonlinear, since the boundary condition is nonlinear (notice that $\alpha = \alpha(x, y)$ depends on the state y). Even if the boundary condition would be linear with respect to y, i.e. $\alpha = \alpha(x)$ (or $\alpha = \alpha(x, t)$), the mapping $u \mapsto y$ is still nonlinear (bilinear), because the product $y\,u$ appears in the boundary condition of (13).

We resolve this difficulty by several manipulations. First of all, we introduce the right hand side of the boundary condition in (13) as a new auxiliary control v, i.e. on $[0, T]$ we put

$$v_i(t) := u_i(t)\alpha(x, y(x,t))(y_{fl} - y(x,t)), \tag{14}$$

where $u_i(t)$ denotes the step function being equal to u_{ik} on I_k. From now on, we search controls $v_i(t)$ subject to the linear boundary condition

$$\frac{\partial y}{\partial n} = v_i(t) \quad \text{on } \Gamma_i,$$

$i = 1(1)K$. The $v_i(t)$ are approximated by step functions. To this aim, we consider partitions of $I_k = (t_{k-1}, t_k]$ into L smaller subintervals having equidistant length $\tau = (t_k - t_{k-1})/L$ and define

$$I_{kl} = (t_{k-1} + (l-1)\tau, t_{k-1} + l\tau) = (t_{k-1,l-1}, t_{k-1,l}),\ k = 1(1)K,\ l = 1(1)L.$$

Finally, these are the intervals, where we really apply the idea of instantaneous control. The optimization is started on $I_{11} = (t_0, t_0 + \tau)$ to obtain optimal values $\bar{v}_{i11},\ i = 1(1)P$. Define $y_{01}(x) = y(x, t_0 + \tau)$ as the new initial temperature for I_{12}. Next the $\bar{v}_{i12}$ are determined, and we put $y_{02}(x) := y(x, t_0 + 2\tau)$. Proceeding in this way, linear short-time optimal control problems (OCS_{kl}) are solved for $k = 1(1)K$, $l = 1(1)L$,

(OCS_{kl}) $\qquad \min \quad y(x_0, t_{kl})$

subject to

$$\begin{aligned} \frac{\partial y}{\partial t} &= \Delta y && \text{in } \Omega \times I_{kl}, \\ y(x, t_{k-1,l-1}) &= y_{k-1,l-1}(x) && \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= v_i && \text{on } \Gamma_i \times I_{kl}, \end{aligned} \tag{15}$$

$$y(x_1, t_{k-1,l}) - y(x_2, t_{k-1,l}) \le c\,,$$

$$q_{ikl} \le v_i \le 0,\ i = 1(1)P.$$

The optimal controls of (OCS_{kl}) are denoted by $\bar{v}_{ikl}$. Moreover, we put $y_{k-1,l}(x) := y(x, t_{k-1,l})$, if $l < L$, and $y_{k,0}(x) := y(x, t_{k-1,L})$. It remains to define the values q_{ikl}. We preselect some characteristic points $\hat{x}_i \in \Gamma_i$ (say midpoints of Γ_i in some sense) and regard formula (14) at $y = y_{k-1,l-1}(\hat{x}_i)$ with maximal control value $u = +1$. This should result in the minimum heat flux

$$q_{ikl} := 1 \cdot \alpha(\hat{x}_i, y_{k-1,l-1}(\hat{x}_i))(y_{fl} - y_{k-1,l-1}(\hat{x}_i)). \tag{16}$$

After having exhausted the whole interval $[0,T]$ by the optimization process, we compose the auxiliary controls $\bar{v}_{ikl}$ to suboptimal controls $\bar{u}_{ik}$, $i=1(1)P$, $k=1(1)K$, as follows: Motivated by (14), resolving for $u_i(t)$, we define

$$\begin{aligned} u^-_{kli} &= \bar{v}_{kli}/\alpha(y(\hat{x}_i, t_{k-1,l-1}))(y_{fl} - y(\hat{x}_i, t_{k-1,l-1})), \\ u^+_{kli} &= \bar{v}_{kli}/\alpha(y(\hat{x}_i, t_{k-1,l}))(y_{fl} - y(\hat{x}_i, \tilde{t}_{k-1,l})). \end{aligned}$$

Finally, the mean values

$$\bar{u}_{kli} = \frac{u^+_{kli} + u^-_{kli}}{2},$$

are taken to compose

$$\bar{u}_{ki} = \frac{\sum\limits_{l=1}^{L} l\, \bar{u}_{kli}}{\sum\limits_{l=1}^{L} l}, \qquad i=1(1)P. \tag{17}$$

The principle of superposition can be used to efficiently generate the problems (OCS$_{kl}$). On I_{kl}, the solution y of (15) is represented in the form

$$y(x,t) = y_I(x,t) + \sum_{i=1}^{P} v_{ikl}\, y_i(x,t), \tag{18}$$

where y_I solves the heat equation subject to $\partial y_I/\partial n = 0$ and $y_I(x, t_{k-1,l-1}) = y_{k-1,l-1}(x)$, while the *response functions* y_i solve the heat equation on I_{kl} with homogeneous initial condition and boundary condition $\partial y_i/\partial n = \chi(\Gamma_i)$.

We notice that $y_i(x,t)$ does not depend on k and l, because $y_i(x,t) = z_i(x, t - t_{k-1,l-1})$ holds, where, for $i=1(1)P$,

$$\begin{aligned} \frac{\partial z_i}{\partial t} &= \Delta z_i && \text{in } \Omega \times (0,\tau), \\ z_i(x,0) &= 0 && \text{on } \Omega, \\ \frac{\partial z_i}{\partial n} &= \chi(\Gamma_i) && \text{on } \Gamma_i \times (0,\tau). \end{aligned} \tag{19}$$

After having solved the P parabolic problems (19) at the beginning of the computations, the functions z_i can be taken to define y_i on all I_{kl}. In this way, (OCS$_{kl}$) is given by

$$\min \quad \sum_{i=1}^{P} v_i\, z_i(x_0, \tau)$$

subject to

$$\begin{gathered} \sum_{i=1}^{P} v_i\,(z_i(x_1,\tau) - z_i(x_2,\tau)) \le c + y_I(x_2, t_{k-1,l}) - y_I(x_1, t_{k-1,l}), \\ q_{ikl} \le v_i \le 0, \; i=1(1)P. \end{gathered}$$

As the $z_i(x_j, t)$, j=1,2,3, have been determined at the beginning, only the values $y_I(x_2, t_{k-1,l})$, $y_I(x_1, t_{k-1,l})$, and q_{ikl} must be updated during the optimization process. This drastically reduces the number of PDE solves.

The application to the concrete example of cooling steel is based on the same type of ideas. However, we need some essential modifications since the heat equation will be nonlinear and the constraints are more complex. The suboptimal method, despite of all its heuristics, delivered surprisingly precise results, [31].

References

1. N. Arada, J.-P. Raymond, and F. Tröltzsch. On an augmented Lagrangian SQP method for a class of optimal control problems in Banach spaces. *Submitted to Computational Optimization and Applications*, to appear.
2. M. Bergounioux and K. Kunisch. Primal-dual strategy for state-constrained optimal control problems. Submitted.
3. E. Casas. Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations. *SIAM J. Control and Optimization*, 35:1297–1327, 1997.
4. E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for some state-constrained control problems of semilinear elliptic equations. *SIAM J. Control and Optimization*, 38(5):1369–1391, 2000.
5. J. Chen, H.W. Engl, G. Landl, and K. Zeman. Optimal strategies for the cooling of steel strips in hot strip mills. *Inverse Problems in Engineering*, 2:103–118, 1995.
6. H. Goldberg and F. Tröltzsch. On a SQP–multigrid technique for nonlinear parabolic boundary control problems. In W.W. Hager and P.M. Pardalos, editors, *Optimal Control: Theory, Algorithms and Applications*, pages 154–177, Dordrecht, 1997. Kluwer Academic Publishers B.V.
7. H. Goldberg and F. Tröltzsch. On a Lagrange-Newton method for a nonlinear parabolic boundary control problem. *Optimization Methods and Software*, 8:225–247, 1998.
8. W. Grever, A. Binder, H.W. Engl, and K. Mörwald. Optimal cooling strategies in continuous casting of steel with variable casting speed. *Inverse Problems in Engineering*, 2:289–300, 1997.
9. M. Gunzburger and S. Manservisi. The velocity tracking problem for Navier-Stokes flows with bounded distributed controls. *SIAM J. Control and Optimization*, 37:1913–1945, 1999.
10. M. Gunzburger and S. Manservisi. Analysis and approximation of the velocity tracking problem for Navier-Stokes flows with distributed controls. *SIAM J. Numer. Analysis*, To appear.
11. M. Heinkenschloss and E. W. Sachs. Numerical solution of a constrained control problem for a phase field model. In *Control and Estimation of Distributed Parameter Systems*, volume 118 of *Int. Ser. Num. Math.*, pages 171–188. Birkhäuser–Verlag, 1994.
12. M. Heinkenschloss and F. Tröltzsch. Analysis of the Lagrange-SQP-Newton method for the control of a phase field equation. *Control and Cybernetics*, 28(2):178–211, 1999.
13. M. Hinze. *Optimal and instantaneous control of the instationary Navier-Stokes equations.* Habilitation thesis, TU Berlin, 1999.
14. M. Hinze and S. Volkwein. Instantaneous control for the Burgers equation: Convergence analysis and numerical implementation. Tech. Report, Spezialforschungsbereich Optimierung und Kontrolle 170, Inst. f. Mathematik, Karl-Franzens-Universität Graz, 1999.
15. K. Ito and K. Kunisch. Augmented Lagrangian-SQP methods for nonlinear optimal control problems of tracking type. *SIAM J. Control and Optimization*, 34:874–891, 1996.
16. A. Kauffmann. *Optimal control of the solid fuel ignition model.* Thesis, TU Berlin, 1998.
17. C.T. Kelley and E. Sachs. Multilevel algorithms for constrained compact fixed point problems. *SIAM J. Sci. Comput.*, 15:645–667, 1994.
18. C.T. Kelley and E. Sachs. Solution of optimal control problems by a pointwise projected Newton method. *SIAM J. Control and Optimization*, 33:1731–1757, 1995.
19. C.T. Kelley and E. Sachs. A trust region method for parabolic boundary control problems. *SIAM J. Optimization*, 9:1064–1081, 1999.

20. R. Krengel, R. Standke, F. Tröltzsch, and H. Wehage. Mathematisches Modell einer optimal gesteuerten Abkühlung von Profilstählen in Kühlstrecken. Preprint 98–6, TU Chemnitz, Fakultät für Mathematik, 1996.
21. K. Kunisch and E. Sachs. Reduced SQP-methods for parameter identification problems. *SIAM Journal Numerical Analysis*, 29:1793–1820, 1992.
22. K. Kunisch and S. Volkwein. Augmented Lagrangian-SQP techniques and their approximations. *Contemporary Math. 209*, pages 147–159, 1997.
23. K. Kunisch and S. Volkwein. Control of Burgers equation by a reduced order approach using proper orthogonal decomposition. Preprint, Spezialforschungsbereich Optimierung und Kontrolle 138, Inst. f. Mathematik, Karl-Franzens-Universität Graz, 1998.
24. R. Lezius and F. Tröltzsch. Theoretical and numerical aspects of controlled cooling of steel profiles. In H. Neunzert, editor, *Progress in Industrial Mathematics at ECMI 94*, pages 380–388. Wiley-Teubner, 1996.
25. J. L. Lions. *Contrôle optimal de systèmes gouvernès par des équations aux dérivées partielles.* Dunod, Gauthier–Villars, Paris, 1968.
26. J.-P. Raymond and F. Tröltzsch. Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints. *Discrete and Continuous Dynamical Systems*, 6:431–450, 2000.
27. J.-P. Raymond and H. Zidani. Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations. *Applied Mathematics and Optimization*, 39:143–177, 1999.
28. F. Tröltzsch. *Optimality conditions for parabolic control problems and applications*, volume 62 of *Teubner Texte zur Mathematik.* B.G. Teubner Verlagsgesellschaft, Leipzig, 1984.
29. F. Tröltzsch. An SQP method for the optimal control of a nonlinear heat equation. *Control and Cybernetics*, 23(1/2):267–288, 1994.
30. F. Tröltzsch, R. Lezius, R. Krengel, and H. Wehage. Mathematische Behandlung der optimalen Steuerung von Abkühlungsprozessen bei Profilstählen. In K.H. Hoffmann, W. Jäger, T. Lohmann, and H. Schunck, editors, *Mathematik– Schlüsseltechnologie für die Zukunft, Verbundprojekte zwischen Universität und Industrie*, pages 513–524. Springer-Verlag, 1997.
31. F. Tröltzsch and A. Unger. Fast solution of optimal control problems in the selective cooling of steel. *ZAMM*, 81:447–456, 2001.
32. A. Unger. *Hinreichende Optimalitätsbedingungen 2. Ordnung und Konvergenz des SQP–Verfahrens für semilineare elliptische Randsteuerprobleme.* PhD thesis, Technische Universität Chemnitz, 1997.