Optimal Control of Crystallization Processes

Thomas Götz
Department of Mathematics
University of Kaiserslautern
Erwin–Schrödinger–Str. 48
D–67663 Kaiserslautern
Germany

Rene Pinnau
Department of Mathematics
University of Kaiserslautern
Erwin–Schrödinger–Str. 48
D–67663 Kaiserslautern
Germany

Jens Struckmeier
Department of Mathematics
University of Hamburg
Bundesstr. 55
D–20146 Hamburg
Germany

August 10, 2005

Abstract

In this paper an optimal control problem for polymer crystallization is investigated. The crystallization is described by a non–isothermal Avrami–Kolmogorov model and the temperature at the boundary of the domain serves as control variable. The cost functional takes into account the spatial variation of the crystallinity and the final degree of crystallization. This results in a boundary control problem for a parabolic equation coupled with two ordinary differential equations, which is treated by an adjoint variable approach. We prove the existence and uniqueness of solutions to the state system as well as the existence of a minimizer for the cost functional under consideration. The adjoint system is derived and we use a steepest descent algorithm to solve the problem numerically. Numerical simulations illustrate the applicability and performance of the optimization algorithm.

Keywords: Crystallization, Optimal Control, Numerics.

MSC:

1 Introduction

In many technological applications like semiconductor production or polymer processing, crystallization processes play an important role. A recent overview on mathematical models for crystallization can be found in [1] and detailed analysis of some of these models can be found e.g. in [2, 3, 4]. However in industrial applications the control and optimization of these processes is an important task [5]. In this context, the main question is how to control the boundary temperature of the crystallizing material in order to obtain a good quality of the final product. Similar questions also appear in glass production [6] or in semiconductor design [7, 8]. In this paper, we want to apply tools from optimal control to a simplified crystallization model. In [9] the following model is used to describe the non–isothermal crystallization of polymers:

$$\partial_t \theta = \varepsilon \Delta \theta + \lambda \exp(-\varphi)\, \alpha_t(\theta) \qquad \text{in } (0,T)\times\Omega\,, \tag{1.1a}$$
$$\theta(0,x) = \theta_0(x) \qquad \text{in } \Omega\,, \tag{1.1b}$$
$$-\partial_n \theta = \beta(\theta - u) \qquad \text{on} (0,T)\times\partial\Omega\,, \tag{1.1c}$$

where θ is the dimensionless temperature and $\theta_0 \geq 0$ is the initial temperature and u is the exterior temperature. The function $\alpha_t(\theta)$ is defined as

$$\alpha_t(\theta) = \gamma a(\theta) \int_0^t a(\theta)\, b(\theta)\, ds$$

and

$$\varphi = \int_0^t \alpha_s(\theta)\, ds$$

denotes the degree of crystallinity. This parabolic heat equation with memory is considered inside a bounded domain $\Omega \subset \mathbb{R}^n$ with smooth boundary $\partial\,\Omega$ and for times $t \in (0,T]$.

Remark 1.1. Strictly speaking, the equations describe the crystallization only for dimension $n = 2$. However, the subsequent analysis of the equations is independent of n.

In order to apply tools from optimal control it is advantageous to rewrite (1.1) as a system of ordinary differential equations coupled to a nonlinear heat equation which reads

$$\partial_t \psi = a(\theta)\, b(\theta)\,, \tag{1.2a}$$
$$\partial_t \varphi = \gamma a(\theta)\, \psi\,, \tag{1.2b}$$
$$\partial_t \theta = \varepsilon \Delta \theta + \lambda\gamma a(\theta)\, e^{-\varphi}\, \psi\,, \tag{1.2c}$$

and which is supplemented with the following initial and boundary data

$$\psi(0,x) = 0\,, \tag{1.2d}$$
$$\varphi(0,x) = 0\,, \tag{1.2e}$$
$$\theta(0,x) = \theta_0(x)\,, \tag{1.2f}$$
$$-\partial_n \theta = \beta(\theta - u)\,. \tag{1.2g}$$

In technological applications, one is interested in controlling the applied boundary temperature u is such a way, that the crystallinity φ is as uniform as possible and reaches a maximal value at the final time T. On the other hand, the energy for generating that boundary control should be as low as possible. To model this, we consider the cost functional

$$\begin{aligned} J(\varphi,u) &= J_0(\varphi) + J_1(\varphi) + J_2(u) \\ &= \frac{\omega_0}{2}\int_0^T \int_\Omega |\nabla\varphi|^2\, dx\, dt + \frac{\omega_1}{2}\int_\Omega \left|e^{-\varphi(T,x)}\right|^2 dx + \frac{\omega_2}{2}\int_0^T \int_{\partial\Omega} |u|^2\, dx\, dt\,, \end{aligned} \tag{1.3}$$

where ω_0, ω_1 and $\omega_2 \geq 0$ are some weights which allow to adjust the objectives J_0 and J_1.

In this paper we discuss the minimization problem:

$$J(\varphi,u) \to \min \qquad \text{such that (1.2) holds.} \tag{1.4}$$

The paper is organized as follows: In Section 2 we proof the existence and uniqueness solutions for the state problem (1.2). The existence of a minimizer for the optimization problem (1.4) in shown in Section 3. Section 4 deals with the derivation of the adjoint equations to solve the minimization task. The numerical method and some simulations are presented in Section 5.

2 Analysis of the State System

Let $\Omega \subset \mathbb{R}^n$ denote an open domain with smooth boundary and let $0 < T < \infty$. We introduce $\Omega_T = (0,T] \times \Omega$ and $\Sigma_T = (0,T] \times \partial\Omega$. We use the standard notation for Sobolev spaces (see [10]), denoting the norm of $W^{m,p}(\Omega)$ ($m \in \mathbb{N}, p \in [1,\infty]$) by $\|\cdot\|_{W^{m,p}(\Omega)}$. In the special case $p = 2$ we use $H^m(\Omega)$ instead of $W^{m,2}(\Omega)$. Further, let $H_0^m(\Omega)$ be the closure of $C_0^\infty(\Omega)$ with respect to the $H^m(\Omega)$–norm. Its dual space $(H_0^m(\Omega))^*$ is denoted by $H^{-m}(\Omega)$. The duality pairing of a Banach space X with its dual space X^* is given by $\langle \cdot, \cdot \rangle_{X^*,X}$. For a Hilbert space H the inner product is denoted by $(\cdot,\cdot)_H$, if $H = L^2(0,1;L^2(\Omega))$ we just write $(\cdot,\cdot)$. Moreover, for any Banach space B we define the space $L^p(0,T;B)$ with $p \in [1,\infty]$ consisting of all measurable functions $\varphi : (0,T) \to B$ for which the norm

$$\|\varphi\|_{L^p(0,T;B)} \stackrel{\text{def}}{=} \left(\int_0^T \|\varphi(t)\|_B^p \, dt \right)^{1/p}, \quad p \in [1,\infty),$$

$$\|\varphi\|_{L^\infty(0,T;B)} \stackrel{\text{def}}{=} \sup_{t \in (0,T)} \|\varphi(t)\|_B \,, \quad p = \infty,$$

is finite. If the time interval is clear we write shortly $\|\cdot\|_{L^p(B)}$. We impose the following assumptions.

A.1 The parameters ε, λ, γ and β are assumed to be positive. For the functions a and b we require that $a, b : \mathbb{R}^+ \to [0,1]$ are continuously differentiable and monotonically decreasing.

A.2 The exterior temperature fulfills $u \geq 0$ and $u \in L^2(0,T;\mathbb{R})$.

A.3 The initial condition fulfills $\theta_0 \geq 0$ and $\theta_0 \in L^2(\Omega)$.

Remark 2.1. A typical choice in **A.1** is $a(\theta) = b(\theta) = \exp(-\kappa\theta)$, for some positive constant κ. The upper bound for the functions a and b can be any arbitrary positive number. However, for the sake of simplicity we use for the subsequent analysis the upper bound equal to 1.

2.1 Existence of Solutions

In the following we prove the existence of at least one weak solution to system (1.2). For the proof we employ fixed point arguments and energy estimates.

Theorem 2.2. *Assume* **A.1**–**A.3**. *Then, there exists a weak solution* $\theta \in L^2(0,T;H^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))$ *of* (1.1). *Further, there exists a constant* $C = C(\Omega,T) > 0$ *such that*

$$\|\theta\|_{L^2(H^1)} + \|\theta\|_{L^\infty(L^2)} + \|\partial_t\theta\|_{L^2(H^{-1})} \leq C \left(\|\theta_0\|_{L^2(\Omega)} + \|u\|_{L^2(0,T;\mathbb{R})} \right).$$

Proof. For the proof we utilize the fixed point theorem of Leray–Schauder, see [11]. Let $\sigma \in [0,1]$ be arbitrary. We introduce the fixed point mapping $G : L^2(0,T;L^2(\Omega)) \times [0,1] \to L^2(0,T;L^2(\Omega))$ where $G(w,\sigma) = \theta$ is defined by the following steps

1. Given $w \in L^2(0,T;L^2(\Omega))$ solve

$$\partial_t\psi = a(w)b(w), \quad \psi(0,x) = 0$$

for ψ.

2. Given w and ψ solve
$$\partial_t\varphi = \gamma a(w)\psi, \quad \varphi(0,x) = 0$$
for φ.

3. Given w, ψ and φ solve
$$\begin{aligned} \partial_t\theta &= \varepsilon\Delta\theta + \sigma\lambda\gamma a(w)e^{-\varphi}\psi && \text{(2.1a)} \\ -\partial_n\theta &= \beta(\theta - \sigma u) && \text{(2.1b)} \\ \theta(0,x) &= \sigma\theta_0(x) && \text{(2.1c)} \end{aligned}$$
for θ.

Here, step 1 is well defined, since we have $a(w), b(w) \in L^\infty(\Omega_T)$ which yields $\psi \in H^1(0,T;L^\infty(\Omega))$. Analogously, we can argue that from step 2 we get $\varphi \in H^1(0,T;L^\infty(\Omega))$. Finally, standard results from parabolic theory [12, Sec. II.3, Thm. 3.1] yield the unique solvability of step 3 for $\theta \in L^2(0,T;H^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))$. Hence, the fixed point mapping G is well–defined.

Now, let θ be a fixed point of G. We use energy methods to derive apriori estimates for θ. Testing (2.1) with θ we get
$$\begin{aligned} \frac{1}{2}\partial_t \int_\Omega |\theta|^2\ dx &= -\varepsilon \int_\Omega |\nabla\theta|^2\ dx + \varepsilon\beta \int_{\partial\Omega} (\sigma u - \theta)\theta\ ds + \sigma\lambda\gamma \int_\Omega a(\theta)e^{-\varphi}\psi\theta\ dx \\ &\leq -\varepsilon \int_\Omega |\nabla\theta|^2\ dx - \frac{\varepsilon\beta}{2}\int_{\partial\Omega} |\theta|^2\ ds + \frac{\varepsilon\beta\sigma^2}{2}\int_{\partial\Omega} |u|^2\ ds + \sigma\lambda\gamma t \int_\Omega |\theta|\ dx \\ &\leq \frac{\varepsilon\beta}{2}\int_{\partial\Omega} |u|^2\ ds + \lambda\gamma t \int_\Omega |\theta|\ dx\ . \end{aligned}$$

Here, we applied Young's inequality and the assumptions$a \leq 1$. Further, we need $\psi \leq t$, $\varphi \geq 0$ which is a direct consequence of the ODE and $\sigma \leq 1$. Integration with respect to t yields
$$\frac{1}{2}\int_\Omega |\theta(t)|^2\ dx \leq \frac{1}{2}\int_\Omega |\theta(0)|^2\ dx + \frac{\varepsilon\beta}{2}\int_0^t \int_{\partial\Omega} \left|u(\tilde t)\right|^2\ ds\, d\tilde t + \lambda\gamma\, c(\Omega) \int_0^t \tilde t \left(\int_\Omega \left|\theta(\tilde t)\right|^2\right)^{1/2}\ dx\, d\tilde t\ .$$

The generalized Gronwall lemma implies
$$\int_\Omega |\theta(t)|^2 \leq \left(\|\theta_0\|^2_{L^2(\Omega)} + \frac{\varepsilon\beta}{2}\|u\|^2_{L^2(\Sigma_T)}\right)^{1/2} + \frac{\lambda\gamma\, c(\Omega)}{2} t\ .$$

Altogether, we have
$$\|\theta\|_{L^\infty(L^2)} \leq K_1(\Omega, \theta_0, u, T)\ ,$$
where $K_1 > 0$ is especially independent of σ. Further, we get immediately
$$\|\theta\|_{L^2(H^1)} \leq K_2(\Omega, \theta_0, u, T)\ ,$$
where again $K_2 > 0$ is independent of σ. To use compactness arguments we need additional information on the regularity of θ with respect to time. To prove an estimate on the time derivative $\partial_t\theta$ we supply $H^{-1}(\Omega)$ with the norm $\left\|\nabla\Delta^{-1}\cdot\right\|_{L^2(\Omega)}$, where $\Delta^{-1} : H^{-1}(\Omega) \to H^1_0(\Omega)$ is the inverse Laplacian. Note, that for $f \in H^{-1}(\Omega)$ we have $\left\|\nabla\Delta^{-1}f\right\|_{L^2(\Omega)} = \|\varphi\|_{H^1_0(\Omega)}$, where $\varphi \in H^1_0(\Omega)$ solves $\Delta\varphi = f$. Using $\varphi = -\Delta^{-1}\partial_t\theta$ as a test function in (2.1) we get
$$\left\|\nabla\Delta^{-1}\partial_t\theta\right\|^2_{L^2(\Omega)} = \varepsilon \int_\Omega \nabla\theta\ \nabla\Delta^{-1}\partial_t\theta\, dx - \sigma\lambda\gamma \int_\Omega a(\theta)e^{-\varphi}\psi\ \Delta^{-1}\partial_t\theta\, dx$$

and applying the Poincare–Friedrichs inequality we derive

$$\leq \varepsilon \|\nabla\theta\|_{L^2(\Omega)} \left\|\nabla\Delta^{-1}\partial_t\theta\right\|_{L^2(\Omega)} + c(\Omega)\,\lambda\gamma\,\|\psi\|_{L^2(\Omega)} \left\|\nabla\Delta^{-1}\partial_t\theta\right\|_{L^2(\Omega)} ,$$

from which we deduce using the previous estimates that

$$\left\|\nabla\Delta^{-1}\partial_t\theta\right\|_{L^2(\Omega_T)} \leq K_3(\Omega,\theta_0,u,T) ,$$

independently of σ. Hence

$$\begin{aligned}\|\theta\|^2_{H^1(H^{-1})} &= \|\theta\|^2_{L^2(H^{-1})} + \|\partial_t\,\theta\|^2_{L^2(H^{-1})} = \left\|\nabla\Delta^{-1}\theta\right\|^2_{L^2(\Omega_T)} + \left\|\nabla\Delta^{-1}\,\partial_t\,\theta\right\|^2_{L^2(\Omega_T)}\\ &\leq c(\Omega_T)^2\;\|\theta\|^2_{L^2(\Omega_T)} + K_3(\Omega,\theta_0,u,T)^2 .\end{aligned}$$

Finally, there exists a constant $K > 0$, independent of θ and σ, such that each fixed point of G fulfills

$$\|\theta\|_{L^2(H^1)} + \|\theta\|_{H^1(H^{-1})} \leq K .$$

It is easy to verify that the operator G is continuous. From Aubin's Lemma [13, Corollary 4] we deduce the compactness of the embedding $L^2(0,T;H^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega)) \hookrightarrow L^2(\Omega_T)$, which implies the compactness of the fixed point operator G. Furthermore, $G(w,0) = 0$ for all $w \in L^2(\Omega_T)$. Now the existence of at least one solution follows from Leray–Schauder's fixed point theorem. □

Corollary 2.3. *Assume* **A.1**–**A.3**. *Then, the system* (1.2) *has a solution*

$$(\theta,\psi,\varphi) \in \left(L^2(0,T;H^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))\right) \times L^\infty(0,T;L^\infty(\Omega)) \times L^\infty(0,T;L^\infty(\Omega))$$

and it holds that

$$0 \leq \psi \leq T, \qquad 0 \leq \varphi \leq \frac{\gamma}{2}T^2.$$

2.2 Uniform Bounds on the Solution

In the next theorem we state some bounds on solutions to (1.2), which correspond to the physical interpretation of the solution.

Theorem 2.4. *Assume* **A.1**–**A.3** *and additionally $u \in L^\infty(0,T;\mathbb{R})$ as well as $\theta_0 \in L^\infty(\Omega)$. Then any solution (θ,φ,ψ) of* (1.2) *defined on Ω_T fulfills:*
There exists a constant $M = M(\Omega,u,\theta_0,T) > 0$ such that

$$0 \leq \theta \leq M(T), \quad 0 \leq \psi \leq T, \quad 0 \leq \varphi \leq \frac{\gamma}{2}T^2.$$

Proof. Note, that $0 \leq a,b \leq 1$ and hence the estimates on ψ and φ follow directly from their equations. To derive the estimates for θ, we first test the equation with θ^- and due to the assumption $u \geq 0$, integration by parts yields

$$\begin{aligned}\frac{1}{2}\partial_t \int_\Omega \left|\theta^-(t)\right|^2\,dx = -\varepsilon \int_\Omega \left|\nabla\theta^-(t)\right|^2\,dx + \varepsilon\beta \int_{\partial\Omega} (u-\theta)(t)\theta^-(t)\,ds\\ + \lambda\gamma \int_\Omega a(\theta(t))e^{-\varphi(t)}\psi(t)\theta^-(t)\,dx \leq 0,\end{aligned}$$

and hence

$$\int_\Omega \left|\theta^-(t)\right|^2\,dx \leq \int_\Omega \left|\theta^-(0)\right|^2\,dx = 0, \quad \forall t \in [0,T),$$

i.e. $\theta^- = 0$ a.e. in Ω_T, or equivalently $\theta \geq 0$.

To derive the upper bound we use Moser's iteration technique. For this purpose we test the equation with θ^p, $p \geq 1$ and get

$$\begin{aligned}
\frac{1}{p+1}\partial_t \int_\Omega \theta^{p+1}(t)\, dx &= -\varepsilon\, p \int_\Omega \theta^{p-1} \left|\nabla\theta(t)\right|^2\, dx + \varepsilon\beta \int_{\partial\Omega} (u-\theta)(t)\theta^p(t)\, ds \\
&\quad + \lambda\gamma \int_\Omega a(\theta(t)) e^{-\varphi(t)} \psi \theta^p(t)\, dx \\
&\leq \varepsilon\beta \left\|u(t)\right\|_{L^{p+1}(\partial\Omega)} \left\|\theta(t)\right\|^p_{L^{p+1}(\partial\Omega)} - \varepsilon\beta \left\|\theta(t)\right\|^{p+1}_{L^{p+1}(\partial\Omega)} \\
&\quad + \lambda\gamma \left|\Omega\right| T \left\|\theta(t)\right\|^p_{L^{p+1}(\Omega)}
\end{aligned}$$

and using Young's inequality

$$\begin{aligned}
&\leq \varepsilon\beta \frac{1}{p+1} \left\|u(t)\right\|^{p+1}_{L^{p+1}(\partial\Omega)} + \varepsilon\beta \left(\frac{p}{p+1} - 1\right) \left\|\theta(t)\right\|^{p+1}_{L^{p+1}(\partial\Omega)} \\
&\quad + \lambda\gamma \frac{1}{p+1} \left(\left|\Omega\right|^{p+1} T^{p+1} + p \left\|\theta(t)\right\|^{p+1}_{L^{p+1}(\Omega)} \right) \\
&\leq \frac{1}{p+1} \left(\varepsilon\beta c(\Omega) \left\|u(t)\right\|^{p+1}_{L^\infty(\partial\Omega)} + \lambda\gamma \left|\Omega\right|^{p+1} T^{p+1} \right) + \lambda\gamma \frac{p}{p+1} \left\|\theta(t)\right\|^{p+1}_{L^{p+1}(\Omega)}.
\end{aligned}$$

Now the Gronwall lemma yields

$$\int_\Omega \theta^{p+1}(t)\, dx \leq \left(\varepsilon\beta c(\Omega) \left\|u\right\|^{p+1}_{L^\infty(\Sigma_T)} + \lambda\gamma \left|\Omega\right|^{p+1} T^{p+1} + \int_\Omega \theta^{p+1}(0)\, dx \right) \exp\left(\lambda\gamma p t\right).$$

Taking the $(p+1)$–th root we get

$$\left\|\theta(t)\right\|_{L^{p+1}(\Omega)} \leq K(\Omega, \theta_0, u, T) e^{\lambda\gamma T}.$$

Since the right hand side does not depend on p we can perform the limit $p \to \infty$ and get

$$\left\|\theta\right\|_{L^\infty(\Omega_T)} \leq M(\Omega, \theta_0, u, T).$$

□

Remark 2.5. From the physical point of view we would expect that the bound M is independent of the final time T, i.e. that we can find a global uniform bound. But this cannot be proved using the above arguments.

2.3 Uniqueness of the Solution

The physical process suggests that the solution to this problem is unique, which is indeed the case as the following theorem shows.

Theorem 2.6. *Assume* **A.1**–**A.3** . *Then the solution* $\theta \in L^2(0,T;H^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))$ *of* (1.1) *is unique.*

Proof. Assume that there exist two solutions θ_1 and θ_2. Taking the difference of the respective equations and testing with $\theta_1 - \theta_2$ we have

$$\begin{aligned}
\frac{1}{2}\partial_t \int_\Omega \left|\theta_1 - \theta_2\right|^2\, dx &= -\varepsilon \int_\Omega \left|\nabla(\theta_1 - \theta_2)\right|^2\, dx - \varepsilon\beta \int_{\partial\Omega} \left|\theta_1 - \theta_2\right|^2\, ds \\
&\quad + \lambda\gamma \underbrace{\int_\Omega \left(a(\theta_1) e^{-\varphi_1} \psi_1 - a(\theta_2) e^{-\varphi_2} \psi_2 \right) (\theta_1 - \theta_2)\, dx}_{\stackrel{\text{def}}{=} I_1} \\
&\leq I_1.
\end{aligned}$$

Using the assumption that a is a decreasing function of θ and $0 \leq a(\theta) \leq 1$, we can estimate further

$$I_1 = \underbrace{\int_\Omega (a(\theta_1) - a(\theta_2))\, e^{-\varphi_1} \psi_1 (\theta_1 - \theta_2)\, dx}_{\leq 0} + \underbrace{\int_\Omega a(\theta_2) \left(e^{-\varphi_1}\psi_1 - e^{-\varphi_2}\psi_2 \right) (\theta_1 - \theta_2)\, dx}_{\stackrel{\text{def}}{=} I_2} \leq I_2.$$

Now we get

$$\begin{aligned} I_2 &\leq \int_\Omega \left| e^{-\varphi_1} - e^{-\varphi_2} \right| |\psi_1| \, |\theta_1 - \theta_2| \; dx + \int_\Omega e^{-\varphi_2} \, |\psi_1 - \psi_2| \, |\theta_1 - \theta_2| \; dx \\ &\leq \int_\Omega |\varphi_1 - \varphi_2| \, |\psi_1| \, |\theta_1 - \theta_2| \; dx + \int_\Omega |\psi_1 - \psi_2| \, |\theta_1 - \theta_2| \; dx \\ &\leq t \left\| \varphi_1 - \varphi_2 \right\|_{L^2(\Omega)} \left\| \theta_1 - \theta_2 \right\|_{L^2(\Omega)} + \left\| \psi_1 - \psi_2 \right\|_{L^2(\Omega)} \left\| \theta_1 - \theta_2 \right\|_{L^2(\Omega)}. \end{aligned}$$

Additionally, one can prove that

$$\begin{aligned} \left\| \psi_1 - \psi_2 \right\|_{L^\infty(L^p)} &\leq K(p,T) \left\| \theta_1 - \theta_2 \right\|_{L^1(L^p)} \\ \left\| \varphi_1 - \varphi_2 \right\|_{L^\infty(L^p)} &\leq K(p,T) \left\| \theta_1 - \theta_2 \right\|_{L^1(L^p)}. \end{aligned}$$

Altogether, we have

$$\begin{aligned} \frac{1}{2} \int_\Omega |\theta_1 - \theta_2|^2 (t) &\leq \lambda\gamma\, K(T) \left(\int_0^t \left\| \theta_1 - \theta_2 \right\|_{L^2(\Omega)} \right)^2 \\ &\leq \lambda\gamma\, K(T)\, T \int_0^t \left\| \theta_1 - \theta_2 \right\|^2_{L^2(\Omega)}. \end{aligned}$$

Finally, the Gronwall lemma implies

$$\left\| (\theta_1 - \theta_2)(t) \right\|_{L^2(\Omega)} \leq 0 \qquad \forall t \in (0,T),$$

which yields directly $\theta_1 = \theta_2$ a.e. in Ω_T. This also implies $\psi_1 = \psi_2$ and $\varphi_1 = \varphi_2$, which ends the proof. □

3 The Optimal Control Problem

Now, that the existence and uniqueness of solutions to the state system (1.2) is settled, we introduce the state variable $z = (\theta, \varphi, \psi) \in Z = Z_\theta \times Z_\varphi \times Z_\psi$ where

$$Z_\theta = \left\{ \theta \in L^2(0,T; H^1(\Omega)) \, : \theta_t \in L^2(0,T; H^{-1}(\Omega)) \right\}$$

and $Z_\varphi = Z_\psi = H^1(0,T; L^\infty(\Omega))$. Additionally we introduce the set of admissible controls $U = L^2(0,T; \mathbb{R}_+)$ and the space for the adjoint variables $\xi = (\xi_\theta, \xi_\varphi, \xi_\psi, \xi^b_\varphi, \xi^b_\psi, \xi_{\theta_0}) \in W$, where

$$W = W_\theta \times W^2_\varphi \times W^2_{\varphi^b} \times L^2(\Omega)$$

and $W_\theta = H^1(0,T; H^1(\Omega))$, $W_\varphi = H^1(0,T; L^\infty(\Omega))$ and $W_{\varphi^b} = H^1(0,T; L^\infty(\partial\,\Omega))$.

Remark 3.1. The additional adjoint variables ξ^b_φ and ξ^b_ψ, which are defined on the boundary of the domain, are introduced due to the PDE–ODE coupling. They are later on needed to identify the boundary conditions in the strong form of the adjoint system.

For the state system (1.2) we introduce the operator $P : Z \times U \to W'$ and denote its weak form as $P(z,u) = 0$ in W'. Then $\langle P(z,u)\,,\,\xi\rangle = 0$ for $\xi \in W$ is defined as

$$
\begin{aligned}
-\langle \theta\,,\,\partial_t\,\xi_\theta\rangle - \varepsilon\langle \theta\,,\,\Delta\,\xi_\theta\rangle + \lambda\gamma\langle a(\theta)e^{-\varphi}\psi\,,\,\xi_\theta\rangle + \int_\Omega \theta\xi_\theta|_0^T + \varepsilon\langle \beta(\theta - u)\,,\,\xi_\theta\rangle_\Sigma + \varepsilon\langle \theta\,,\,\partial_n\,\xi_\theta\rangle_\Sigma &= 0\\
-\langle \varphi\,,\,\partial_t\,\xi_\varphi\rangle - \gamma\langle a(\theta)\psi\,,\,\xi_\varphi\rangle + \int_\Omega \varphi\xi_\varphi|_0^T &= 0\\
-\langle \psi\,,\,\partial_t\,\xi_\psi\rangle - \gamma\langle a(\theta)b(\theta)\,,\,\xi_\psi\rangle + \int_\Omega \psi\xi_\psi|_0^T &= 0\\
-\langle \varphi\,,\,\partial_t\,\xi^b_\varphi\rangle_\Sigma - \gamma\langle a(\theta)\psi\,,\,\xi^b_\varphi\rangle_\Sigma + \langle \varphi\,,\,\xi^b_\varphi\rangle_\Sigma|_0^T &= 0\\
-\langle \psi\,,\,\partial_t\,\xi^b_\psi\rangle_\Sigma - \gamma\langle a(\theta)b(\theta)\,,\,\xi^b_\psi\rangle_\Sigma + \langle \psi\,,\,\xi^b_\psi\rangle_\Sigma|_0^T &= 0\,,\\
\langle \theta(0) - \theta_0\,,\,\xi_{\theta_0}\rangle &= 0,
\end{aligned}
$$

where $\langle f\,,\,g\rangle = \int_{\Omega_T} fg$ and $\langle f\,,\,g\rangle_\Sigma = \int_{\Sigma_T} fg$.

3.1 Existence of a Minimizer

Theorem 3.2. *Assume* **A.1**–**A.3**. *Then, there exists a minimizer* $(\overline{z},\overline{u}) \in Z \times U$ *of* (1.4).

Proof. We consider the cost functional (1.3). Clearly, $J \geq 0$ holds and J is radially unbounded with respect to u. Let $J_0 \stackrel{\text{def}}{=} \inf_{Z\times U} J \geq 0$. We choose a minimizing sequence $(z_k, u_k) = (\theta_k, \varphi_k, \psi_k, u_k)$ in $Z \times U$ such that $P(z_k, u_k) = 0$ in W'. The radial unboundedness of J with respect to u implies $\|u_k\|_U \leq C$. Hence there exists a weakly convergent subsequence, again denoted by $(u_k)_{k\in\mathbb{N}}$, such that $u_k \rightharpoonup \overline{u}$ weakly in U for $k \to \infty$. The apriori estimates of Theorem 2.2 and Corollary 2.3 imply

$$
\begin{gathered}
\|\theta_k\|_{L^\infty(L^2)} + \|\theta_k\|_{L^2(H^1)} + \|\theta_k\|_{H^1(H^{-1})} \leq C\\
\|\psi_k\|_{L^\infty(L^\infty)} \leq C, \qquad \|\varphi_k\|_{L^\infty(L^\infty)} \leq C\\
\|\partial_t\psi_k\|_{L^\infty(L^\infty)} \leq C, \qquad \|\partial_t\varphi_k\|_{L^\infty(L^\infty)} \leq C.
\end{gathered}
$$

Hence, there exist weakly convergent subsequences such that

$$
\begin{aligned}
\theta_k &\rightharpoonup \overline{\theta} \quad \text{weakly in } L^2(0,T;H^1(\Omega))\\
\partial_t\theta_k &\rightharpoonup \overline{\rho} \quad \text{weakly in } L^2(0,T;H^{-1}(\Omega))\\
\psi_k &\rightharpoonup \overline{\psi} \quad \text{weak–}* \ L^\infty(\Omega_T)\\
\varphi_k &\rightharpoonup \overline{\varphi} \quad \text{weak–}* \ L^\infty(\Omega_T)\,.
\end{aligned}
$$

These convergences are by far sufficient to pass to the limit in the equations. □

Remark 3.3. In general, we cannot expect the uniqueness of a minimizer since the set given by the nonlinear constraint is nonconvex. Hence, uniqueness could only be ensured for large values of ω_2 introducing a high convexity into the problem.

4 Adjoint Equations and First Order Optimality

In this section we derive the adjoint equations and state the first–order optimality conditions. We write the first–order optimality system using the Lagrangian $\mathcal{L} : Z \times U \times W \to \mathbb{R}$ associated to (1.4) defined by

$$\mathcal{L}(z,u,\xi) \stackrel{\text{def}}{=} J(z,u) + \langle P(z,u), \xi\rangle_{W',W}\,.$$

Then, the first-order optimality system is given by

$$\mathcal{L}'(z,u,\xi) = 0. \tag{4.1}$$

We rewrite (4.1) in a more concise form.

$$\begin{aligned} P(z,u) &= 0 \quad \text{in } W', \\ P_z^*(z,u)\xi + J_z(z,u) &= 0 \quad \text{in } Z', \\ P_u^*(z,u)\xi + J_u(z,u) &= 0 \quad \text{in } U. \end{aligned}$$

Clearly, the first equation corresponds to the state system. Assuming enough regularity one can identify the strong form of the second equation as the adjoint system for $\xi_\theta, \xi_\varphi, \xi_\psi, \xi_\varphi^b, \xi_\psi^b$ given by

$$-\partial_t \xi_\theta = \varepsilon\,\Delta\,\xi_\theta + \gamma a'(\theta)\,\psi\,\left(\lambda e^{-\varphi}\xi_\theta + \xi_\varphi\right) + \left(a'(\theta)b(\theta) + a(\theta)b'(\theta)\right)\xi_\psi \tag{4.2a}$$

$$-\partial_t \xi_\varphi = -\lambda\gamma\, a(\theta)\, e^{-\varphi}\,\psi\,\xi_\theta + \omega_0\,\Delta\,\varphi \tag{4.2b}$$

$$-\partial_t \xi_\psi = \gamma a(\theta)\left(\lambda e^{-\varphi}\xi_\theta + \xi_\varphi\right)\ , \tag{4.2c}$$

subject to the boundary conditions on Σ_T

$$-\varepsilon\,\partial_n\,\xi_\theta = \varepsilon\beta\,\xi_\theta - \gamma a'(\theta)\psi\,\xi_\varphi^b - \left(a'(\theta)b(\theta) + a(\theta)b'(\theta)\right)\,\xi_\psi^b \tag{4.2d}$$

$$-\partial_t\,\xi_\varphi^b = -\omega_0\,\partial_n\,\varphi \tag{4.2e}$$

$$-\partial_t\,\xi_\psi^b = \gamma a(\theta)\,\xi_\varphi^b\ , \tag{4.2f}$$

and the final conditions at $t = T$

$$\xi_\theta(T,0) = 0 \tag{4.2g}$$

$$\xi_\varphi(T,0) = \omega_1 \exp\left(-2\varphi(T,x)\right) \tag{4.2h}$$

$$\xi_\varphi^b(T,0) = 0 \tag{4.2i}$$

$$\xi_\psi(T,0) = 0 \tag{4.2j}$$

$$\xi_\psi^b(T,0) = 0\ . \tag{4.2k}$$

Remark 4.1. Formally, this adjoint system can be reinterpreted again as a parabolic equation with memory for the adjoint variable ξ_θ. However, for the sake of shorter notation and easier numerical implementation, we will keep the system (4.2).

The final equation is the optimality condition which reads

$$\left\langle \omega_2 u - \varepsilon\beta\xi_\theta|_{\Sigma_T}\ , v\right\rangle_U \geq 0 \qquad \forall v \in U\ .$$

4.1 The Reduced Problem

Thanks to Theorems 2.2 and 2.6 there exists a unique solution $z = z(u) \in Z$ and we are able to introduce the reduced problem $\hat{P}(u) = P(z(u),u) = 0$ and the reduced cost functional $\hat{J}(u) = J(z(u),u)$. Then the optimization problem can be formulated as

$$\min_U \hat{J}(u) \text{ subject to } \hat{P}(u) = 0. \tag{4.3}$$

Next, we compute the derivative of the functional $\hat{J}$. Assuming sufficient regularity we have that the operator $P_z(z,u) : Z \times U \to W'$ is a homeomorphism. Hence, the implicit function theorem implies that the first derivative of the mapping $u \mapsto z(u)$ at u in a direction $v \in U$ is given by

$$z'(u)[v] = -P_z^{-1}(z(u),u)P_u(z(u),u)v.$$

Using the chain rule one obtains

$$\left\langle \hat{J}'(u), v \right\rangle_U = \left\langle J_u(z(u), u) - P_u^*(z(u), u) P_z^{-*}(z(u), u) J_z(z(u), u), v \right\rangle_U .$$

Using the adjoint variable

$$\xi = -P_z^{-*} J_z(u) \in W$$

the Riesz representative of the mapping $u \mapsto \hat{J}'(u)$ is given by

$$\hat{J}'(u) = J_u(z(u), u) + P_u^*(z(u), u)\xi$$

or

$$\hat{J}'(u) = \omega_2 u - \varepsilon\beta\xi_\theta|_{\Sigma_T} \ . \tag{4.4}$$

5 Numerical Results

Now we want to use the results obtained for the reduced problem to construct a projected gradient algorithm for the solution of our constrained optimization problem (1.4).

5.1 The Projected Gradient Algorithm

For the solution of problem (1.4) we apply the following gradient method.

Algorithm 5.1.

1. *Choose an initial guess* $u^{(0)} \in U$ *for the control.*
2. *Solve the state system* $P\left(z^{(k+1)}, u^{(k)}\right) = 0$ *for* $z^{(k+1)} \in Z$.
3. *Solve the adjoint problem* $P_z^*\left(z^{(k+1)}, u^{(k)}\right)\xi^{(k+1)} = -J_z\left(z^{(k+1)}, u^{(k)}\right)$ *for* $\xi^{(k+1)} \in W$.
4. *Update* $u^{(k+1)} = \Pi_U\left(u^{(k)} - \delta\hat{J}'\left(u^{(k)}\right)\right)$, *where* $\delta > 0$ *and* $\Pi_U : L^2(0,T;\mathbb{R}) \mapsto U$ *is the projection onto* U.

For the numerical simulations we discretize the boundary control $u \in L^2(0,T;\mathbb{R}_+)$ as a constant in space and piecewise constant in time, i. e. $u(t,x) = u_l$ for all $x \in \partial\Omega$ and $t_l < t \le t_{l+1}$, where $t_0 = 0$ and $t_L = T$. The state system (1.2) is discretized with standard finite differences and solved forward in time using an explicit scheme. On the same grid the adjoint system (4.2) is discretized and solved backward in time. The information between the state system and the adjoint system is interchanged at the discrete time points t_l. Also the gradient of the cost functional J is computed at these intermediate times according to Eqn. (4.4).

The convergence of the proposed algorithm clearly depends on the choice of the step size δ in step 4 of Algorithm 5.1. The best choice would be the result of a line search

$$\delta^* = \operatorname{argmin}_{\delta>0} \hat{J}\left(\Pi_U\left(u^{(k)} - \delta\hat{J}'\right)\right) .$$

Unfortunately, this is by far too expensive, since each evaluation of the cost functional $\hat{J}$ requires the solution of the nonlinear state problem. Therefore, we resort to a heuristics and use a small but fixed step size $\delta = \tilde{\delta}$. For suitably small values of $\tilde{\delta}$ the cost functional will be reduced in each step.

Remark 5.2. Alternatively we could also choose the step size according to

$$\delta := \min \left\{ \tilde{\delta} \ , \ \left\| \hat{J}'(u^{(k)}) \right\|_{\infty}^{-1} \right\}$$

with some limiting parameter $\tilde{\delta}$. The idea behind this heuristics is to restrict the step size long as the gradient is large in such a way that only changes of 1K are allowed for the new control. Near the optimal solution, where the gradient is small, take steps with a fixed step size $\tilde{\delta}$.

Other possibilities, avoiding the rather expensive line search, are the implementation of an Armijo–heuristics.

5.2 Simulations

We consider a rectangular domain $\Omega = [0,1] \times [0,2] \subset \mathbb{R}^2$, the time interval $[0,T]$, where $T = 800$ and the following set of parameters in our crystallization model: $\varepsilon = 10^{-3}$, $\lambda = 1/3$, $\gamma = 2$, $\beta = 10$, $a(\theta) = b(\theta) = \exp\left[-\kappa(\theta - \theta_{\text{Ref}})\right]$, $\kappa = 3$, $\theta_{\text{Ref}} = 0.4$, $\theta_0 = 2$. The domain is discretized using a uniform space grid with step size $h = 1/20$ and a time step $\tau = 0.25$ ensuring the stability condition $\varepsilon\tau/h^2 < 1/8$.

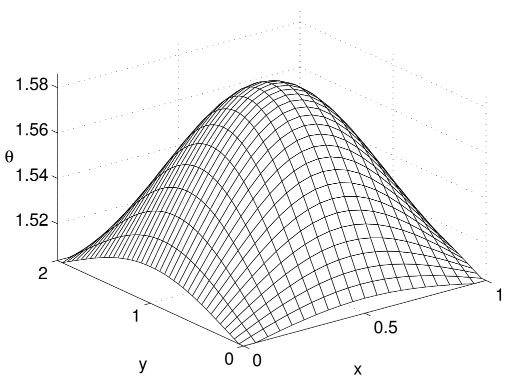

Figure 5.1: Temperature θ at $T = 800$ for constant boundary control $u^{(0)} = 1.5$.

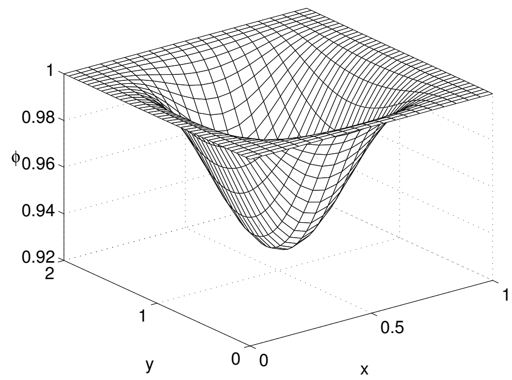

Figure 5.2: Crystallinity φ at $T = 800$ for constant boundary control $u^{(0)} = 1.5$.

Figures 5.1 and 5.2 show the final temperature profile and the degree of crystallinity for the case of a constant boundary temperature $u = u^{(0)} = 1.5$. At the boundaries of the domain, the crystallinity φ has already reached its maximum of $\varphi = 1$, whereas the inner part is not yet fully crystallized.

Figure 5.3 shows the cost functional, as well as the individual contributions to it, for a constant boundary control $u^{(0)}$ varying between $u^{(0)} = 1$ and $u^{(0)} = 2$ and weights $\omega_0 = 10^{-5}$, $\omega_1 = 10^5$ and $\omega_2 = 10^{-2}$. At low boundary temperatures the J_0–term dominates the cost functional indicating high gradients in the crystallinity due to the strong cooling at the boundaries. For high boundary temperatures, the domain will not fully crystallize, therefore the J_1–term has major influence on the cost functional.

We find the optimal *time constant* boundary control $u \in \mathbb{R}_+$ after 20 gradient steps with $\tilde{\delta} = 2 \cdot 10^{-5}$. Figures 5.4 and 5.5 show the temperature and the crystallinity for the optimized boundary control $u = 1.395$. Note, that compared to Figure 5.2 the crystallinity has everywhere reached its maximal value of $\varphi = 1$. The advance of the optimization procedure is given in Figure 5.6. The cost functional is reduced from $J = 116.3$ for $u = 1.50$ to $J = 51.45$ for $u = 1.395$ and the optimization is close to stationary already after 10 gradient steps.

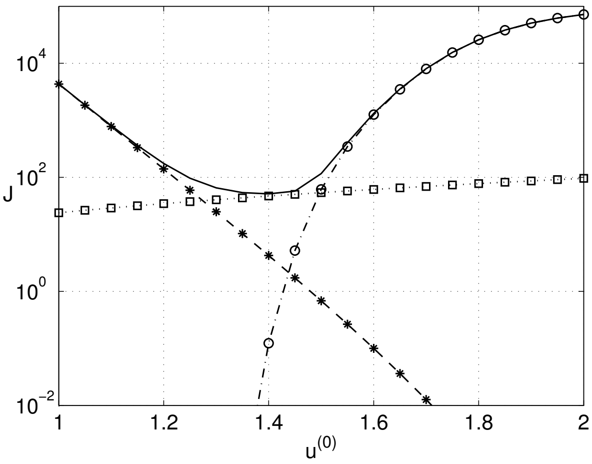

Figure 5.3: Cost functional J for $T = 800$ and $u^{(0)} \in [1, 2]$. The solid line ('—') shows J, the stars ('$*- -$') correspond to J_0, ('$\circ- \cdot$') is the J_1–term and ('$\Box \ldots$') visualizes the contribution of J_2 to the cost functional. The weights are chosen as $\omega_0 = 10^{-5}$, $\omega_1 = 10^5$ and $\omega_2 = 10^{-2}$.

As a final numerical result, we consider a *time dependent* boundary control which is piecewise constant over $L = 40$ time intervals. For the optimization procedure we use again $u^{(0)} = 1.5$ as an initial guess. The optimum is found after $k = 80$ gradient steps. The transient behavior of the boundary temperature is shown in Figure 5.7. The oscillation of the optimized boundary temperature for small times t is due to the fact that we only penalize the L^2–norm of u. However, the additional degrees of freedom in the case of the time dependent boundary temperature only lead to minor improvements of the cost functional, see Table 5.1.

In Table 5.1 we summarize the considered scenarios and the resulting values of the cost functional J.

Control Type	u	J	J_0	J_1	J_2
Initial guess	1.50	116.3	0.687	61.570	54.00
Time constant	1.395	51.45	4.701	0.074	46.68
Time dependent		51.40	4.068	0.191	47.14

Table 5.1: Values of the cost functional for the different considered scenarios.

6 Conclusions

We studied an optimal boundary control problem for a polymer crystallization process with the aim to maximize the final degree of crystallization and the uniformity of the crystallization rate, which correlates with the local gradients of the crystallinity. The Avrami–Kolmogorov crystallization model gives rise to a parabolic equation with memory. We proved the existence and uniqueness of solution for this state equation as well as the existence of a minimizer for the considered cost functional. Based on the first order optimality system, we proposed a projected gradient method to solve the optimization problem numerically.

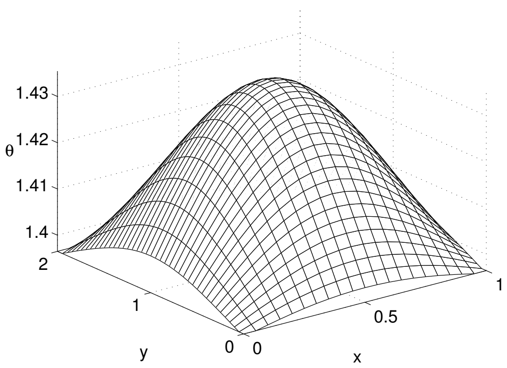

Figure 5.4: Temperature θ at $T = 800$ for the optimized constant boundary control.

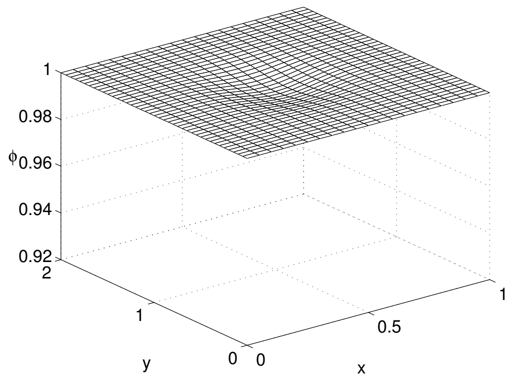

Figure 5.5: Crystallinity φ at $T = 800$ for the optimized constant boundary control.

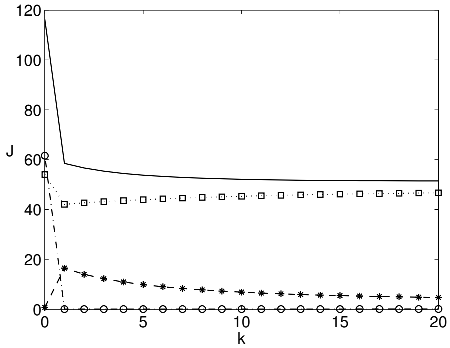

Figure 5.6: Advance of the optimization procedure for a time constant control versus the number k of the gradient step. The solid line ('—') shows J, the stars ('$*$– –') correspond to J_0, ('$\circ$– ·') is the J_1–term and ('$\square$ ...') visualizes the contribution of J_2 to the cost functional.

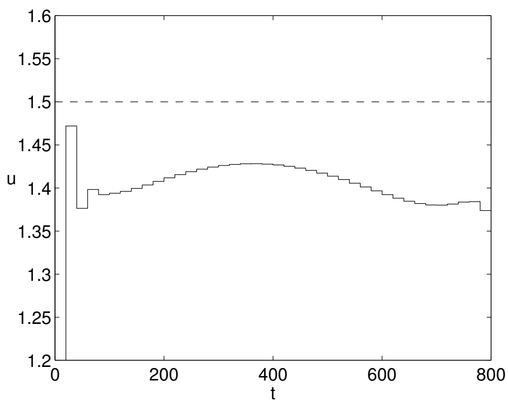

Figure 5.7: Results of the optimization for the case of a time dependent boundary temperature u. The dashed line ('– –') shows the initial guess $u^{(0)} = 1.5$ and the solid line ('—') is the final result $u^{(80)}$ after 80 gradient steps.

The algorithm was tested in a prototypic example and showed a significant improvement of the final degree of crystallization even for a boundary temperature which is constant in time and space. Introducing time variable boundary temperatures yielded only minor improvements. Whether the observation holds true for more industry relevant examples or not, shall be left to future investigations.

References

[1] V. Capasso. *Mathematical modelling for polymer processing.* Springer, New York, 2002.

[2] E. Ratajski and H. Janeschitz-Kriegl. How to determine high growth speeds in polymer crystallization. *Colloid Polym. Sci.*, 274:938–951, 1996.

[3] G. Eder. Mathematical modelling of crystallization processes as occuring in polymer processing. *Nonlinear Analysis, Theory, Methods & Applications*, 30(6):3807–3815, 1997.

[4] M. Burger and V. Capasso. Mathematical modelling and simulation of non–isothermal crystallization of polymers. *Math. Models Meth. Appl. Science*, 11(6):1029–1053, 2001.

[5] M. Burger, V. Capasso, and A. Micheletti. Optimal control of polymer morphologies. *J. Eng. Math.*, 49(4):339–358, 2004.

[6] R. Pinnau and G. Thömmes. Optimal boundary control of glass cooling processes. *Math. Methods Appl. Sci.*, 27(11):1261–1281, 2004.

[7] M. Hinze and R. Pinnau. An optimal control approach to semiconductor design. *Math. Mod. Meth. Appl. Science*, 12(1):89–107, 2002.

[8] K.H. Hoffmann, A. Voigt, and M. Metzger. Numerical simulation and control of industrial crystal growth processes. In W. Jäger, editor, *Mathematics - key technology for the future. Joint projects between universities and industry*, pages 331–342. Springer, Berlin, 2003.

[9] T. Götz and J. Struckmeier. A simplified model for non–isothermal crystallization of polymers. *Mathematical and Computer Modelling*, accepted.

[10] R. A. Adams. *Sobolev spaces*. Academic Press, Boston, 6. edition, 1990.

[11] E. Zeidler. *Nonlinear functional analysis and its applications: V.1: Fixed–point theorems*. Springer, New York, 1986.

[12] R. Temam. *Infinite–dimensional dynamical systems in mechanics and physics*. Springer, New York, 2nd edition, 1997.

[13] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl.*, 146:65–96, 1987.