# PDE Constrained Optimization – selected Proofs

Jeff Snider

jeff@snider.com

George Mason University
Department of Mathematical Sciences

April, 2014

# Outline

# Preliminary definitions

$X$ denotes a Banach space, and a function $y : [a, b] \to X$ is called *vector-valued*.

We specifically consider $X := L^2(\Omega)$, where $\Omega \subset \mathbb{R}^N$ is defined on the next slide.

**Definition**

$C\big([a, b]; X\big)$ is the space of functions $y(x, t)$ continuous at every point $t \in [a, b]$.

**Definition**

$L^p(a, b; X), 1 \le p < \infty$, is the space of measurable vector-valued functions $y : [a, b] \to X$ such that $\|y\|_{L^p(a,b;X)} := \left(\int_a^b \|y(t)\|_X^p \, dt\right)^{1/p} < \infty$.

**Definition**

$L^\infty(a, b; X)$ is the space of measurable vector-valued functions $y : [a, b] \to X$ such that $\|y\|_{L^\infty(a,b;X)} := \operatorname{ess\,sup}_{[a,b]} \|y(t)\|_X < \infty$.

# Problem Statement

## Problem 3.23

$$\begin{aligned} y_t - \Delta y + c_0 y &= f && \text{in } Q = \Omega \times (0, T) \\ \partial_\nu y + \alpha y &= g && \text{on } \Sigma = \Gamma \times (0, T) \\ y(0) &= y_0 && \text{in } \Omega. \end{aligned} \tag{3.23}$$

*where* $f \in L^2(Q)$, $g \in L^2(\Sigma)$, *and* $y_0 \in L^2(\Omega)$.

## Assumption 3.8

*Let* $\Omega \subset \mathbb{R}^N$ *be a bounded Lipschitz domain with boundary* $\Gamma$, *and let* $T > 0$ *be a fixed final time. Moreover, assume that functions* $c_0 \in L^\infty(Q)$ *and* $\alpha \in L^\infty(\Sigma)$, *where* $\alpha(x,t) \geq 0$ *for almost every* $(x,t) \in \Sigma$, *are prescribed.*

Q: does this imply that $\Omega$ is connected?

# Definitions

**Definition**

$$W_2^{1,0}(Q) := \left\{ y \in L^2(Q) \;\middle|\; \frac{\partial}{\partial x_i} y \in L^2(Q), i = 1, \dots, N \right\}.$$

**Definition**

$$W(0,T) := \{ y \in L^2(0,T;V) \mid y' \in L^2(0,T;V^*) \}$$

**Definition**

$V = H^1(\Omega)$, and $H = L^2(\Omega)$ where we identify $H = H^*$.

We have shown that

$$V \subset H \subset V^*,$$

is a sequence of dense and continuous embeddings, implying

$$H^1(\Omega) \subset L^2(\Omega) \subset H^1(\Omega)^*.$$

### Theorem 3.9

*Suppose Assumption 3.8 holds. Then the parabolic initial-boundary value problem (3.23) has a unique weak solution in* $W_2^{1,0}(Q)$. *Moreover there is a constant* $c_p > 0$, *independent of* $f$, $g$, *and* $y_0$, *such that*

$$\max_{t\in[0,T]} \|y(\cdot,t)\|_{L^2(\Omega)} + \|y\|_{W_2^{1,0}(Q)} \leq c_p\big(\|f\|_{L^2(Q)} + \|g\|_{L^2(\Sigma)} + \|y_0\|_{L^2(\Omega)}\big) \tag{3.26}$$

*for all* $f \in L^2(Q)$, $g \in L^2(\Sigma)$, *and* $y_0 \in L^2(\Omega)$.

### Theorem 3.10

*Every* $y \in W(0,T)$ *coincides, possibly after modification on a null set, with an element of* $C([0,T],H)$. *In this sense we have the continuous embedding* $W(0,T) \hookrightarrow C([0,T],H)$.

## Theorem 3.11

*For all $y, p \in W(0,T)$ the formula of* integration by parts *holds:*

$$\int_0^T \big(y'(t), p(t)\big)_{V^*,V}\,dt = \big(y(T), p(T)\big)_H - \big(y(0), p(0)\big)_H - \int_0^T \big(p'(t), y(t)\big)_{V^*,V}\,dt.$$

## Proof (3.11)

Apply the chain rule:

$$\frac{d}{dt}\big(y(t), p(t)\big)_H = \big(y'(t), p(t)\big)_{V^*,V} + \big(p'(t), y(t)\big)_{V^*,V}.$$

XXX Much detail is missing from the above equality. XXX

Now integrate both sides:

$$\int_0^T \frac{d}{dt}\big(y(t), p(t)\big)_H\,dt = \int_0^T \big(y'(t), p(t)\big)_{V^*,V}\,dt + \int_0^T \big(p'(t), y(t)\big)_{V^*,V}\,dt$$

$$\big(y(T), p(T)\big)_H - \big(y(0), p(0)\big)_H = \int_0^T \big(y'(t), p(t)\big)_{V^*,V}\,dt + \int_0^T \big(p'(t), y(t)\big)_{V^*,V}\,dt.$$

□

**Theorem 3.12**

*Let $y \in W_2^{1,0}(Q)$ be the weak solution to problem (3.23), which exists according to Theorem 3.9. Then $y$ belongs, possibly after a modification on a set of zero measure, to $W(0,T)$.*

The proof covers the next several slides.

You should recall the inner product,

$$(y, v)_{L^2(\Omega)} := \int_\Omega y\, v\, dx.$$

# Proof (3.12)

It follows from the problem statement, when looking for a weak solution via (3.25), that for all $v \in W_2^{1,1}(Q)$ with $v(T) = 0$,

$$-\iint_Q y\, v_t \, dx\, dt = - \iint_Q \nabla y \cdot \nabla v \, dx\, dt - \iint_Q c_0\, y\, v\, dx\, dt - \iint_\Sigma \alpha\, y\, v\, ds\, dt + \int_\Omega y_0\, v(0)\, dx + \iint_Q f\, v\, dx\, dt + \iint_\Sigma g\, v\, ds\, dt. \quad (1)$$

In particular we may insert any function of the form $v(x,t) := v(x)\phi(t)$, where $\phi \in C_0^\infty(0,T)$ and $v \in V = H^1(\Omega)$. Setting $H = L^2(\Omega)$ and $H^N = H \times H \times \cdots \times H$ ($N$ times), we find, first

$$-\iint_Q y\, v_t\, dx\, dt = -\int_0^T \int_\Omega y\, v_t\, dx\, dt = -\int_0^T \left(y, v_t\right)_{L^2(\Omega)} dt$$
$$= -\int_0^T \left(y, v\phi'(t)\right)_{L^2(\Omega)} dt = -\int_0^T \left(y(t)\phi'(t), v\right)_{L^2(\Omega)} dt.$$

Applying this technique to each term in (1),

# Proof (3.12) contd.

$$-\int_0^T \big(y(t)\phi'(t), v\big)_H\, dt = -\int_0^T \big(\nabla y(t), \nabla v\big)_{H^N}\phi(t)\, dt - \int_0^T \big(c_0 y(t), v\big)_H \phi(t)\, dt$$
$$-\int_0^T \big(\alpha(t) y(t), v\big)_{L^2(\Gamma)}\phi(t)\, dt + \int_0^T \big(f(t), v\big)_H \phi(t)\, dt + \int_0^t (g(t), v)_{L^2(\Gamma)}\phi(t)\, dt.$$

Now $y \in L^2(Q)$, by the definition of $W_2^{1,0}(Q)$. Hence, by Fubini's theorem, $y(\cdot, t) \in L^2(\Omega)$ for almost every $t \in (0, T)$. Moreover $D_i y \in L^2(Q)$ for $i = 1, \cdots, N$, and thus $\nabla y(\cdot, t) \in \big(L^2(\Omega)\big)^N = H^N$ for a.e. $t \in (0, T)$. Finally, $y(\cdot, t) \in H^1(\Omega)$, and thus $y(\cdot, t) \in L^2(\Gamma)$ for a.e. $t \in (0, T)$.

On the set of measure zero in $[0, T]$ where one of the above statements possibly does not hold, we put $y(t) = 0$, which does not change the vector-valued function $y$ in the sense of $L^2$ spaces. Hence, we see that for any fixed $t$, the expressions in the integrals on the right-hand side define linear functionals $F_i(t) : H^1(\Omega) \to \mathbb{R}$:

$$F_1(t) : v \mapsto \big(\nabla y(t), \nabla v\big)_{H^N}$$
$$F_2(t) : v \mapsto \big(c_0(t) y(t), v\big)_H$$
$$F_3(t) : v \mapsto \big(\alpha(t) y(t), v\big)_{L^2(\Gamma)}$$
$$F_4(t) : v \mapsto \big(f(t), v\big)_H$$
$$F_5(t) : v \mapsto \big(g(t), v\big)_{L^2(\Gamma)}$$

# Proof (3.12) contd.

We claim that the functionals $F_i(t)$ are bounded and thus continuous on $V$ for every $t$.

First we have for $F_1(t) : v \mapsto (\nabla y(t), \nabla v)_{H^N} = \int_\Omega \nabla y \cdot \nabla v \, dx$,

$$\begin{aligned}
|F_1(t)v| &= \left| \int_\Omega \nabla y(t) \cdot \nabla v \, dx \right| \le \int_\Omega |\nabla y(t) \cdot \nabla v| \, dx \\
&= \|\nabla y(t) \cdot \nabla v\|_{L^1(\Omega)} \le \|\nabla y(t)\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} \qquad (2) \\
&\le \left(\|y(t)\|_{L^2(\Omega)} + \|\nabla y(t)\|_{L^2(\Omega)}\right)\left(\|v\|_{L^2(\Omega)} + \|\nabla v\|_{L^2(\Omega)}\right) \\
&= \|y(t)\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}, \qquad \forall v \in H^1(\Omega). \qquad (3)
\end{aligned}$$

Line (2) is exactly Hölder's inequality, and the final equality is by definition of the $H^1(\Omega)$ norm. Note that the function $t \mapsto \|y(t)\|_{H^1(\Omega)}$ belongs to $L^2(0, T)$, and $\|y(t)\|_{H^1(\Omega)}$ is by construction everywhere finite: $\|y(t)\|_{H^1(\Omega)} = \left(\int_\Omega |y(t)|^2 \, dx\right)^{1/2} < \infty$. Recall, for arbitrary $F \in V^*$, the dual norm is defined by

$$\|F\|_{V^*} := \sup_{v \in V} \frac{|F\, v|}{\|v\|_V}.$$

(3) gives us $\frac{|F(t)v|}{\|v\|_{H^1(\Omega)}} \le \|y(t)\|_{H^1(\Omega)}$, for all $v \in H^1(\Omega)$. Thus $F_1(t)$ is bounded: $\|F_1(t)\|_{H^1(\Omega)^*} \le \|y(t)\|_{H^1(\Omega)}$.

# Proof (3.12) contd.

Similarly, for $F_3(t) : v \mapsto \left(\alpha(t)y(t), v\right)_{L^2(\Gamma)} = \int_\Gamma \alpha(t)y(t)v\,ds$,

$$|F_3(t)v| \le \int_\Gamma |\alpha(t)||y(t)||v|\,ds \le \|\alpha\|_{L^\infty(\Sigma)}\|y(t)\|_{L^2(\Gamma)}\|v\|_{L^2(\Gamma)}$$
$$\le \tilde{c}\|\alpha\|_{L^\infty(\Sigma)}\|y(t)\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)}.$$

Plenty of detail is omitted from the above line. Repeated application of Cauchy-Schwartz and/or Hölder, followed by the Trace Theorem which gives us the $\tilde{c}$ multiple when going from the boundary to the space.

Hence $F_3(t)$ is bounded: $\|F_3(t)\|_{H^1(\Omega)^*} \le \tilde{c}\|\alpha\|_{L^\infty(\Sigma)}\|y(t)\|_{H^1(\Omega)}$.

# Proof (3.12) contd.

These proofs are left out of the text as "easy exercises for the reader:"

- $F_2(t) : v \mapsto \big(c_0(t)y(t), v\big)_H = \int_\Omega c_0(t)y(t)v\,dx,$

$$|F_2(t)v| \;\le\; \int_\Omega |c_0(t)|\,|y(t)|\,|v|\,dx \;\le\; \|c_0\|_{L^\infty(Q)}\|y(t)\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}$$

so that

$$\|F_2(t)\|_{H^1(\Omega)^*} \;\le\; \|c_0\|_{L^\infty(Q)}\|y(t)\|_{L^2(\Omega)}.$$

- $F_4(t) : v \mapsto \big(f(t), v\big)_H = \int_\Omega f(t)v\,dx$

$$|F_4(t)v| \;\le\; \int_Q |f(t)|\,|v|\,dx \qquad \Longrightarrow \qquad \|F_4(t)\|_{H^1(\Omega)^*} \;\le\; \|f(t)\|_{L^2(\Omega)}.$$

- $F_5(t) : v \mapsto \big(g(t), v\big)_{L^2(\Gamma)} = \int_\Gamma g(t)v\,ds,$

$$|F_5(t)v| \;\le\; \int_\Gamma |g(t)|\,|v|\,ds \;\le\; \|g(t)\|_{L^2(\Gamma)}\|v\|_{H^1(\Gamma)} \;\le\; \tilde{c}\|g(t)\|_{L^2(\Gamma)}\|v\|_{H^1(\Omega)}$$

so that

$$\|F_5(t)\|_{H^1(\Omega)^*} \;\le\; \tilde{c}\|g(t)\|_{L^2(\Gamma)}.$$

Since each $F_i$ is bounded, we know each $F_i(t) \in H^1(\Omega)^*$ for every $t$.

# Proof (3.12) contd.

Since each $\|F_i(t)\|_{V^*}$ is bounded by some constant multiple of $\|y(t)\|_{H^1(\Omega)}$, $\|f(t)\|_{L^2(\Omega)}$, or $\|g(t)\|_{L^2(\Gamma)}$, there is some constant $c > 0$ such that

$$\sum_{i=1}^{5} \|F_i(t)\|_{V^*} \le c\Big(\|y(t)\|_{H^1(\Omega)} + \|f(t)\|_{L^2(\Omega)} + \|g(t)\|_{L^2(\Gamma)}\Big). \tag{3.32}$$

Since the expression on the right-hand side belongs to $L^2(0,T)$, so does the expression on the left-hand side, showing that $F_i \in L^2(0,T;V^*)$ for each $i$. But then the functional $F$ on the right-hand side of the variational formulation, being just the sum of of the $F_i$, also belongs to $L^2(0,T;V^*)$.

Rewriting the variational formulation in terms of $F$, we obtain that for all $v \in V$ we have the chain of inequalities

$$\begin{aligned}\Big(-\int_0^T y(t)\,\phi'(t)\,dt, v\Big)_{L^2(\Omega)} &= -\int_0^T \Big(y(t)\,\phi'(t), v\Big)_{L^2(\Omega)} dt \\ &= \int_0^T \Big(F(t)\,\phi(t), v\Big)_{V^*,V} dt = \Big(\int_0^T F(t)\,\phi(t)\,dt, v\Big)_{V^*,V}\end{aligned}$$

and therefore as an equation in the space $V^*$

$$-\int_0^T y(t)\,\phi'(t)\,dt = \int_0^T F(t)\phi(t)\,dt, \qquad \forall \phi \in C_0^\infty(0,T).$$

# Proof (3.12) contd.

$$-\int_0^T y(t)\,\phi'(t)\,dt = \int_0^T F(t)\phi(t)\,dt, \qquad \forall \phi \in C_0^\infty(0,T)$$

means that $y' = F$ in the sense of vector-valued distributions; hence $y' \in L^2(0,T;V^*)$.

We conclude that, for any $y \in W_2^{1,0}(Q)$ which is a weak solution of (3.23), we have

$$y \in W(0,T) := \{y \in L^2(0,T;V) \mid y' \in L^2(0,T;V^*)\}.$$

□

**Theorem 3.13**

*The weak solution $y$ to the problem (3.23) satisfies an estimate of the form*

$$\|y\|_{W(0,T)} \le c_w \big(\|f\|_{L^2(Q)} + \|g\|_{L^2(\Sigma)} + \|y_0\|_{L^2(\Omega)}\big),$$

*with some constant $c_w > 0$ that does not depend on $(f, g, y_0)$. In other words, the mapping $(f, g, y_0) \mapsto y$ defines a continuous linear operator from $L^2(Q) \times L^2(\Sigma) \times L^2(\Omega)$ into $W(0, T)$ and, in particular, into $C\big([0, T], L^2(\Omega)\big)$.*

# Proof (3.13)

We estimate the square of the norm $\|y\|_{W(0,T)}$:

$$\|y\|^2_{W(0,T)} = \|y\|^2_{L^2(0,T;H^1(\Omega))} + \|y'\|^2_{L^2(0,T;H^1(\Omega)^*)}.$$

For the first summand, we obtain from Theorem 3.9 on page 140, with a generic constant $c > 0$, the estimate

$$\|y\|^2_{L^2(0,T;H^1(\Omega))} = \|y\|^2_{W_2^{1,0}(Q)} \leq c\big(\|f\|^2_{L^2(Q)} + \|g\|^2_{L^2(\Sigma)} + \|y_0\|^2_{L^2(\Omega)}\big). \tag{3.33}$$

The second summand requires only a little more effort. Indeed, with the functionals $F_i$ defined previously, we have

$$\|y'\|_{L^2(0,T;H^1(\Omega)^*)} = \left\|\sum_{i=1}^{5} F_i\right\|_{L^2(0,T;H^1(\Omega)^*)} \leq \sum_{i=1}^{5} \|F_i\|_{L^2(0,T;H^1(\Omega)^*)}.$$

Using the above estimates, in particular (3.32), we find, with generic constants $c > 0$, that

# Proof (3.13) contd.

$$\|F_1\|^2_{L^2(0,T;V^*)} = \int_0^T \|F_1(t)\|^2_{V^*}\,dt \le \int_0^T c\|y(t)\|^2_{H^1(\Omega)}\,dt \le c\|y\|^2_{W_2^{1,0}(Q)}$$

$$\|F_2\|^2_{L^2(0,T;V^*)} = \int_0^T \|F_2(t)\|^2_{V^*}\,dt \le \int_0^T \|c_0\|^2_{L^\infty(Q)}\|y(t)\|^2_{L^2(\Omega)}\,dt \le c\|y\|^2_{W_2^{1,0}(Q)}$$

$$\|F_3\|^2_{L^2(0,T;V^*)} = \int_0^T \|F_3(t)\|^2_{V^*}\,dt \le \int_0^T \tilde{c}\|\alpha\|^2_{L^\infty(Q)}\|y(t)\|^2_{L^2(\Omega)}\,dt \le c\|y\|^2_{W_2^{1,0}(Q)}$$

$$\|F_4\|^2_{L^2(0,T;V^*)} = \int_0^T \|F_4(t)\|^2_{V^*}\,dt \le \int_0^T \|f(t)\|^2_{L^2(\Omega)}\,dt \le c\|f\|^2_{L^2(Q)}$$

$$\|F_5\|^2_{L^2(0,T;V^*)} = \int_0^T \|F_5(t)\|^2_{V^*}\,dt \le \int_0^T \tilde{c}\|g(t)\|^2_{L^2(\Gamma)}\,dt \le c\|g\|^2_{L^2(\Sigma)}$$

Since, from (3.33),

$$\|y\|^2_{W_2^{1,0}(Q)} \le c\big(\|f\|_{L^2(Q)} + \|g\|_{L^2(\Sigma)} + \|y_0\|_{L^2(\Omega)}\big)^2,$$

we have

$$\|y'\|^2_{L^2(0,T;H^1(\Omega)^*)} \le \sum_{i=1}^{5} \|F_i\|^2_{L^2(0,T;H^1(\Omega)^*)} \le c\big(\|f\|_{L^2(Q)} + \|g\|_{L^2(\Sigma)} + \|y_0\|_{L^2(\Omega)}\big)^2,$$

and the assertion is proved. $\square$

# Example 3.5.1 – Optimal nonstationary boundary temperature

$$\min J(y,u) := \frac{1}{2}\int_\Omega |y(x,T) - y_\Omega(x)|^2 \, dx + \frac{\lambda}{2}\iint_\Sigma |u(x,t)|^2 \, ds(x)\, dt,$$

subject to

$$\boxed{\begin{aligned} y_t - \Delta y &= 0 && \text{in } Q \\ \partial_\nu y + \alpha y &= \beta u && \text{on } \Sigma \\ y(0) &= 0 && \text{in } \Omega \end{aligned}}$$

and

$$u_a(x,t) \le u(x,t) \le u_b(x,t) \qquad \text{for a.e. } (x,t) \in \Sigma.$$

This is a *parabolic boundary control problem* with final-value cost functional.

# Assumptions

**Assumption 3.14**

*Let $\Omega \subset \mathbb{R}^N$ be a domain with Lipschitz boundary $\Gamma$, and let $\lambda \geq 0$ be a fixed constant. Assume that we are given functions $y_\Omega \in L^2(\Omega)$, $y_Q \in L^2(Q)$, $y_\Sigma \in L^2(\Sigma)$, $\alpha, \beta \in L^\infty(E)$, and $u_a, u_b, v_a, v_b \in L^2(E)$ with $u_a(x,t) \leq u_b(x,t)$ and $v_a(x,t) \leq v_b(x,t)$ for almost every $(x,t) \in E$. Here, depending on the specific problem under study, $E = Q$ or $E = \Sigma$.*

$E$ is the domain of the control, which depends on the problem.

# Example 3.5.1 contd.

Theorems 3.12 and 3.13 guarantee a weak solution $y \in W(0,T)$ for any control $u \in L^2(\Sigma)$, which we represent by

$$y = G_\Sigma(\beta u).$$

We only need information about the solution at time $T$, so we define an *observation operator* $E_T : y \mapsto y(T)$; this is a continuous linear mapping from $W(0,T)$ into $L^2(\Omega)$ since the embedding $W(0,T) \hookrightarrow C\big([0,T], L^2(\Omega)\big)$ has these properties. Hence for some constant $c > 0$ the bound applies:

$$\|y(T)\|_{L^2(\Omega)} \leq \max_{t \in [0,T]} \|y(t)\|_{L^2(\Omega)} =: \|y\|_{C([0,T],L^2(\Omega))} \leq c\|y\|_{W(0,T)}.$$

The first inequality holds by nature of the max operator. The second inequality holds by consequence of Theorem 3.10. Hence, we have

$$y(T) = E_T G_\Sigma(\beta u) =: Su.$$

In summary, $u \mapsto y \mapsto y(T)$ is a continuous linear mapping

$$S : u \mapsto y(T)$$

from the control space $L^2(\Sigma)$ into the space $L^2(\Omega)$ which contains $y(T)$.

# Example 3.5.1 contd.

Using $S$ we can rewrite the problem as a quadratic optimization problem in the Hilbert space $U = L^2(\Sigma)$:

$$\min_{u \in U_{ad}} f(u) := \frac{1}{2}\|Su - y_\Omega\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)}, \tag{3.37}$$

where

$$U_{ad} = \{u \in L^2(\Sigma) : u_a(x,t) \le u(x,t) \le u_b(x,t) \quad \text{for a.e. } (x,t) \in \Sigma\}.$$

The functional $f$ is continuous in $u$ because $S$ and the norms are continuous. For any $h \in [0,1]$, Hölder's inequality gives us

$$\|hz_1 + (1-h)z_2\|^2_{L^2} \le \|hz_1\|^2_{L^2} + \|(1-h)z_2\|^2_{L^2} \le h\|z_1\|^2_{L^2} + (1-h)\|z_2\|^2_{L^2}.$$

Setting $z = Su - y_\Omega$ in one case and $z = u$ in the other and adding, we see $f(u)$ is convex.

The admissible set $U_{ad}$ is nonempty, closed, bounded, and convex subset of the Hilbert space $L^2(\Sigma)$. Hence we can infer from Theorem 2.14 on page 50 the following existence result.

**Theorem 3.15**

*Suppose that Assumption 3.14 holds with $E := \Sigma$. Then the optimization problem (3.37) and hence the optimal nonstationary boundary temperature problem* (3.1)-(3.3) *has at least one optimal control $\bar{u} \in U_{ad}$. If $\lambda > 0$ then $\bar{u}$ is unique.*

# Example 3.5.2 – Optimal nonstationary heat source

$$\min J(y,u) := \frac{1}{2}\iint_{\Sigma} |y(x,t) - y_{\Sigma}(x,t)|^2 \, ds(x)\, dt + \frac{\lambda}{2}\iint_{Q} |u(x,t)|^2 \, dx\, dt, \tag{3.38}$$

subject to

$$\boxed{\begin{aligned} y_t - \Delta y &= \beta u && \text{in } Q \\ \partial_\nu y &= 0 && \text{on } \Sigma \\ y(0) &= 0 && \text{in } \Omega \end{aligned}} \tag{3.39}$$

and

$$u_a(x,t) \le u(x,t) \le u_b(x,t) \qquad \text{for a.e. } (x,t) \in Q. \tag{3.40}$$

This can be seen as an *inverse problem*: an unknown heat source $u$ has to be recovered from measurements of the temperature at the boundary.


7:40
INSTANT ACTION
1 2 3
4 5 6
7 8 9
POWER LEVEL
0
TIMER CLOCK
STOP CLEAR
START

# Example 3.5.2 contd.

Theorems 3.12 and 3.13 guarantee a weak solution $y \in W(0,T)$ for any $u \in L^2(Q)$. The control-to-state operator is

$$y = G_Q(\beta u).$$

The cost functional only requires the values at the boundary, $y(x,t)\big|_\Sigma$.

Since the trace operator $y \mapsto y_\Gamma$ maps $H^1(\Omega)$ continuously into $L^2(\Gamma)$, the mapping $E_\Sigma : y \mapsto y_\Sigma$ defines a continuous linear operator from $L^2\big(0,T;H^1(\Omega)\big)$ into $L^2\big(0,T;L^2(\Gamma)\big)$. Consequently the mapping $u \mapsto y \mapsto y\big|_\Sigma$, i.e., the operator

$$S : u \mapsto y\big|_\Sigma$$

maps the control space $L^2(Q)$ continuously into the space $L^2\big(0,T;L^2(\Gamma)\big) \cong L^2\big((0,T)\times\Gamma\big) = L^2(\Sigma)$ to which $y\big|_\Sigma$ belongs. Thus

$$Su = E_\Sigma G_Q(\beta u).$$

# Example 3.5.2 contd.

Substituting $y = Su$ in the objective we arrive at the quadratic minimization problem in the Hilbert space $U = L^2(Q)$:

$$\min_{u \in U_{ad}} f(u) := \frac{1}{2}\|Su - y_\Sigma\|^2_{L^2(\Sigma)} + \frac{\lambda}{2}\|u\|^2_{L^2(Q)}. \tag{3.42}$$

Theorem 2.14 gives us the existence of an optimal control, hence

**Theorem 3.16**

*Suppose that Assumption 3.14 holds with $E = Q$. Then the optimal nonstationary heat source problem* (3.38)*-(3.40) has at least one optimal control $\overline{u} \in U_{ad}$. If $\lambda > 0$ then $\overline{u}$ is unique.*

# Necessary Optimality Conditions

We will derive the necessary optimality conditions for the prior two examples.

First a variational inequality will be derived that involves the state $y$.

Then $y$ will be eliminated by means of the adjoint state to deduce a variational inequality for the control $u$.

# Adjoint Problem

Consider the parabolic problem

$$\begin{aligned} -p_t - \Delta p + c_0 p &= a_Q \\ \partial_\nu p + \alpha p &= a_\Sigma \\ p(\cdot, T) &= a_\Omega \end{aligned} \tag{3.43}$$

with bounded and measurable coefficient functions $c_0$ and $\alpha$, and prescribed functions $a_Q \in L^2(Q)$, $a_\Sigma \in L^2(\Sigma)$, and $a_\Omega \in L^2(\Omega)$.

This is the *adjoint problem.*

# Lemma 3.17

Using the bilinear form

$$a[t; y, v] := \int_\Omega \left(\nabla y \cdot \nabla v + c_0(\cdot, t)\, y\, v\right) dx + \int_\Gamma \alpha(\cdot, t)\, y\, v\, ds,$$

## Lemma 3.17

*The parabolic problem* (3.43) *has a unique weak solution* $p \in W_2^{1,0}(Q)$*, which is the solution to the variational problem*

$$\iint_Q p\, v_t\, dx\, dt \;+\; \int_0^T a[t; p, v]\, dt \;=\; \int_\Omega a_\Omega v(T)\, dx \;+\; \iint_Q a_Q\, v\, dx\, dt \;+\; \iint_\Sigma a_\Sigma\, v\, dx\, dt,$$

$\forall v \in W_2^{1,1}(Q)$ *with* $v(\cdot, 0) = 0$*.*

*We have* $p \in W(0, T)$*, and there is a constant* $c_a > 0$*, which does not depend on the given functions, such that*

$$\|p\|_{W(0,T)} \le c_a \big(\|a_Q\|_{L^2(Q)} + \|a_\Sigma\|_{L^2(\Sigma)} + \|a_\Omega\|_{L^2(\Omega)}\big).$$

## Proof (3.17)

Let $\tau \in [0, T]$, and put $\tilde{p}(\tau) := p(T - \tau)$ and $\tilde{v}(\tau) := v(T - \tau)$. Then $\tilde{p}(0) = p(T)$, $\tilde{p}(T) = p(0)$, $\tilde{v}(0) = v(T)$, $\tilde{v}(T) = v(0)$, as well as $\tilde{a}_Q(\cdot, t) := a_Q(\cdot, T - \tau)$, etc., and also

$$\iint_Q p\, v_t\, dx\, dt = -\iint_Q \tilde{p}\, \tilde{v}_\tau\, dx\, d\tau,$$

and so on. Consequently, the asserted variational formulation is equivalent to the definition of the weak solution to the (forward) parabolic initial-boundary value problem

$$\begin{aligned} \tilde{p}_\tau - \Delta \tilde{p} + c_0 \tilde{p} &= \tilde{a}_Q \\ \partial_\nu \tilde{p} + \alpha \tilde{p} &= \tilde{a}_\Sigma \\ \tilde{p}(0) &= a_\Omega \end{aligned}$$

By Theorem 3.9 there is a unique weak solution $\tilde{p}$, which by Theorem 3.12 belongs to $W(0, T)$. The assertion now follows from reversing the time transformation. □

Since $p \in W(0,T)$ we can, in analogy to (3.35) rewrite after integration by parts the variational formulation of the adjoint equation in the following shorter form:

$$\boxed{\begin{aligned} &\int_0^T \Big\{ -(p_t, v)_{V^*,V} + a[t;p,v] \Big\}\, dt = \iint_Q a_Q\, v \, dx\, dt + \iint_\Sigma a_\Sigma \, v\, ds\, dt, \\ &\hspace{20em} \forall v \in L^2(0,T;V) \\ &p(T) = a_\Omega. \end{aligned}} \qquad (3.44)$$

Like in the elliptic case, for the derivation for the adjoint system we need the following somewhat technical result.

## Theorem 3.18

*Let $y \in W(0,T)$ be the solution to the parabolic problem*

$$\begin{aligned} y_t - \Delta y + c_0 y &= b_Q v \\ \partial_\nu y + \alpha y &= b_\Sigma u \\ y(0) &= b_\Omega w, \end{aligned}$$

*with coefficient functions $c_0, b_Q \in L^\infty(Q)$, $\alpha, b_\Sigma \in L^\infty(\Sigma)$, and $b_\Omega \in L^\infty(\Omega)$, and controls $v \in L^2(Q)$, $u \in L^2(\Sigma)$, and $w \in L^2(\Omega)$. Moreover, let square integrable functions $a_\Omega$, $a_Q$, and $a_\Sigma$, be given, and let $p \in W(0,T)$ be the weak solution to the adjoint problem, (3.43). Then we have*

$$\iint_Q a_Q y \, dx\, dt + \iint_\Sigma a_\Sigma \, y \, ds\, dt + \int_\Omega a_\Omega \, y(\cdot, T)\, dx = \iint_Q b_Q \, p \, v \, dx\, dt + \iint_\Sigma b_\Sigma \, p \, u \, ds\, dt + \int_\Omega b_\Omega \, p(\cdot, 0)\, w \, dx.$$

## Proof (3.18)

The variational formulation for $y$, using the test function $p$, we have

$$\int_0^T \left\{(y_t, p)_{V^*,V} + a[t; y, p]\right\} dt = \iint_Q b_q\, p\, v\, dx\, dt + \iint_\Sigma b_\Sigma\, p\, u\, ds\, dt, \tag{3.45}$$

with the initial condition $y(0) = b_\Omega w$. Integrating by parts we obtain

$$\int_0^T \left\{ - (p_t, y)_{V^*,V} + a[t; y, p]\right\} dt = -(y(T), a_\Omega)_{L^2(\Omega)} + \left(b_\Omega w, p(0)\right)_{L^2(\Omega)} + \iint_Q b_q\, p\, v\, dx\, dt + \iint_\Sigma b_\Sigma\, p\, u\, ds\, dt. \tag{3.47}$$

Analogously, taking $y$ as test function in the equation for $p$, we find that

$$\int_0^T \left\{ - (p_t, y)_{V^*,V} + a[t; p, y]\right\} dt = \iint_Q a_Q\, y\, dx\, dt + \iint_\Sigma a_\Sigma\, y\, ds\, dt, \tag{3.46}$$

with the final condition $p(T) = a_\Omega$. Since thet left-hand sides of (3.47) and (3.46) coincide, the right hand sides of (3.46) and (3.47) must also be equal, which gives us

$$\iint_Q a_Q\, y\, dx\, dt + \iint_\Sigma a_\Sigma\, y\, ds\, dt + (y(T), a_\Omega)_{L^2(\Omega)} = \iint_Q b_q\, p\, v\, dx\, dt + \iint_\Sigma b_\Sigma\, p\, u\, ds\, dt + \left(b_\Omega w, p(0)\right)_{L^2(\Omega)}$$

□

# Necessary Conditions for Example 1

In this section we determine the necessary conditions for (3.1)-(3.3), the *parabolic boundary control problem with final-value cost functional*:

$$\min J(y,u) := \frac{1}{2}\|y(T) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)},$$

subject to

$$\boxed{\begin{aligned} y_t - \Delta y &= 0 \\ \partial_\nu y + \alpha y &= \beta u \\ y(0) &= y_0 \end{aligned}}$$

and

$$u_a \le u \le u_b.$$

Here the initial condition may be nonzero, which has been avoided previously. It is a straightforward exercise to show the previous conclusions apply also to this case. (Next slide.)

**Problem 3.2**

*Prove the existence of an optimal control for (3.1)-(3.3) given an inhomogeneous initial condition $y_0$.*

The problem is

$$\min J(y,u) := \frac{1}{2}\int_\Omega |y(x,T) - y_\Omega(x)|^2 \, dx + \frac{\lambda}{2}\iint_\Sigma |u(x,t)|^2 \, ds(x)\, dt, \tag{3.1}$$

subject to

$$\boxed{\begin{aligned} y_t - \Delta y &= 0 && \text{in } Q \\ \partial_\nu y + \alpha y &= \beta u && \text{on } \Sigma \\ y(0) &= y_0 && \text{in } \Omega \end{aligned}} \tag{3.2}$$

and

$$u_a(x,t) \le u(x,t) \le u_b(x,t) \qquad \text{for a.e. } (x,t) \in \Sigma. \tag{3.3}$$

The arguments are identical to those leading to Theorem 3.15, and the supporting theorems and equations all allow for any $y_0 \in L^2(\Omega)$. Hence the same arguments go through nearly unmodified:

Theorem 3.9 gives us that for any control, including the initial condition, there is a unique weak solution $y \in W_2^{1,0}(Q)$, and by Theorem 3.12 it is in the equivalence class of a function in $W(0,T)$. Adapting Theorem 3.13, we represent the unique solution for the zero initial condition $y(0) = 0$ using the continuous linear operator $y = G_\Sigma(\beta u)$, and the unique solution for homogeneous boundary conditions $u = 0$ with nonzero initial condition $y(0) = y_0$ using the continuous linear operator $y = G_0(y_0)$. The superposition principle for linear problems gives us a continuous linear mapping

$$y = G_\Sigma(\beta u) + G_0(y_0),$$

We need only the value at final time $T$, so the linear continuous *observation operator* $E_T$ gives us:

$$y(T) = E_T\big(G_\Sigma(\beta u) + G_0(y_0)\big).$$

which can be written as the linear continuous mapping

$$S : (u, y_0) \mapsto y(T).$$

Thus the objective is

$$\min J(y, u, y_0) := \frac{1}{2}\|S(u, y_0) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)},$$

subject to

$$U_{ad} = \{u \in L^2(\Sigma) : u_a(x,t) \le u(x,t) \le u_b(x,t) \quad \text{for a.e. } (x,t) \in \Sigma\}.$$

Since $f$ is convex and continuous, and $U_{ad}$ is nonempty, closed, bounded and convex, Theorem 2.14 gives us existence of an optimal control. □

# Adjoint problem for Example 1

Each term in the derivative of the objective w.r.t. $y$,

$$\nabla_y J(y,u) = y(T) - y_\Omega,$$

appears in the right-hand side of the adjoint. Since $y(T) - y_\Omega \in L^2(\Omega)$, for optimal $\bar{y}$, the *adjoint system* must be

$$\boxed{\begin{aligned} -p_t - \Delta p &= 0 && \text{in } Q \\ \partial_\nu p + \alpha p &= 0 && \text{on } \Sigma \\ p(T) &= \bar{y}(T) - y_\Omega && \text{in } \Omega. \end{aligned}} \tag{3.48}$$

This conclusion comes from experience with previous problems. The formal Lagrange method reaches the same system, as detailed in Chapter 6.

An incomplete attempt to use the formal Lagrange method follows, highlighting some of its difficulties.

# Lagrangian

We will use $p^{(i)}$ as our Lagrange multipliers – what space they are from is as yet unknown, a difficulty only overcome in Chapter 6 in the case of elliptic problems with the statement "the theory for parabolic problems is quite similar." However, for this formal approach that detail is unecessary. The Lagrangian for Example 1 is

$$\mathcal{L}(y,u,p) := \frac{1}{2}\|y(T) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)} - \underbrace{\iint_Q p^{(1)}(y_t - \Delta y)\,dx\,dt}_{\text{1st constraint}} - \underbrace{\iint_\Sigma p^{(2)}(\partial_\nu y + \alpha y - \beta u)\,ds\,dt}_{\text{2nd constraint}}$$

where we have left out the box constraints for now; they are applied later in a variational inequality. Similarly the initial condition is left out of this formulation, to be used later. Integrating by parts, we have (which can also be inferred from (3.25))

$$\mathcal{L}(y,u,p) := \frac{1}{2}\|y(T) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)} - \iint_Q \left[\nabla y \cdot \nabla p^{(1)} - p^{(1)}_t y\right] dx\,dt - \iint_\Sigma \left[y\,p^{(1)} + \alpha y\,p^{(2)} - \beta u\,p^{(2)}\right] ds\,dt + \int_\Omega y_0 p^{(1)}(0)\,dx - \int_\Omega y(T)p^{(1)}(T)\,dx$$

Note that there is a cancelation between two $\partial_\nu y$ terms, one multiplied by $p^{(1)}$ and the other by $p^{(2)}$. We proceed under the hope that our result will show, as it had previously, that the two $p$ terms are equal (more precisely, $p^{(2)}$ is the restriction of $p^{(1)}$ to the boundary).

# Lagrangian, contd.

Now that we have the Lagrangian,

$$\mathcal{L}(y,u,p) := \frac{1}{2}\|y(T) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)} - \iint_Q \left[\nabla y \cdot \nabla p^{(1)} - p_t^{(1)} y\right] dx\, dt$$
$$- \iint_\Sigma \left[y\, p^{(1)} + \alpha y\, p^{(2)} - \beta u\, p^{(2)}\right] ds\, dt + \int_\Omega y_0 p^{(1)}(0)\, dx - \int_\Omega y(T) p^{(1)}(T)\, dx$$

integrating the $\nabla p$ term by parts, then taking the $y$ derivative we have

$$D_y\mathcal{L}(y,u,p) := \iint_Q p_t\, dx\, dt - \iint_\Sigma \alpha p\, ds\, dt + \iint_Q \Delta p\, ds\, dt - \iint_\Sigma \partial_\nu p\, ds\, dt$$

Following the technique of Section 2.10, we multiply with test function $v \in H^1(Q)$ and, after integration by parts and waving of hands, by setting equal to zero we achieve the weak formulation of the adjoint problem (see Lemma 3.17):

$$\iint_Q p\, v_t\, dx\, dt + \iint_Q \nabla p \cdot \nabla v\, dx\, dt + \iint_\Sigma \alpha\, p\, v\, ds\, dt = \int_\Omega p(T)v(T)\, dx - \int_\Omega p(0)v(0)\, dx.$$

In Theorem 3.9 the assumption is made that $v(0) = 0$. The leap to $p(T) = y(T) - y_\Omega$ comes from pp. 120–122, where also the $p^{(1)} = p^{(2)}$ argument can be found.

Here, we end our trip down the rabbit hole of the formal Lagrangian method and continue with the remaining material of this section.

**Theorem 3.19**

*Let $\overline{u} \in U_{ad}$ be a control with associated state $\overline{y}$, and let $p \in W(0,T)$ be the corresponding adjoint state that solves* (3.48)*. Then $\overline{u}$ is an optimal control for the optimal nonstationary boundary temperature problem* (3.1)-(3.3) *if and only if the variational boundary inequality*

$$\iint_{\Sigma} \left(\beta(x,t)p(x,t) + \lambda\overline{u}(x,t)\right)\left(u(x,t) - \overline{u}(x,t)\right) ds(x)\, dt \geq 0 \qquad (3.49)$$

*holds for all $u \in U_{ad}$.*

Observe that this is where the box constraints come into play in the formal Lagrange method, and shows why they can be omitted from the Lagrangian.

$\beta p + \lambda\overline{u}$ is the $u$ derivative of the Lagrangian at the point $\overline{u}$. This variational inequality is one way of writing that the constrained problem has no direction of descent from this point.

# Proof (3.19)

Overview of the proof: We are going to use Lemma 2.21 (p.63) to show equivalence between the optimal solution and the solution of the variational inequality, then by the result of Theorem 3.18 transform the inequality into the necessary form.

Lemma 2.21 tells us that, for a convex $f$, and nonempty convex $U_{ad}$ a subset of a Banach space, $\overline{u}$ solves

$$\min_{u \in U_{ad}} f(u) := \frac{1}{2}\|Su - y_\Omega\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)} \tag{2.43}$$

if and only if it solves the variational inequality

$$f'(\overline{u})(u - \overline{u}) \geq 0, \qquad \forall u \in U_{ad}. \tag{2.44}$$

Since $f'(u) = S^*(S\overline{u} - y_\Omega) + \lambda\overline{u}$, the variational inequality can be written

$$\left(S^*(S\overline{u} - y_\Omega) + \lambda\overline{u}, u - \overline{u}\right)_{L^2(\Sigma)} \geq 0, \quad \forall u \in U_{ad} \tag{2.45}$$

which is equivalently, and more conveniently, written

$$(Su - y_\Omega, Su - S\overline{u})_{L^2(\Sigma)} + \lambda(\overline{u}, u - \overline{u})_{L^2(\Sigma)} \geq 0, \quad \forall u \in U_{ad}. \tag{2.47}$$

# Proof (3.19) contd.

Let $S : L^2(\Sigma) \to L^2(\Omega)$ be the continuous linear operator that, for the homogeneous initial condition $y_0 = 0$, assigns to each control $u$ the final value $y(T)$ of the weak solution $y$ to the state equation. Moreover, let $\overline{y} = G_0 y_0$ denote the weak solution corresponding to $y_0 \neq 0$ and $u = 0$. Then it follows from the superposition principle for linear equations that

$$y(T) - y_\Omega = Su + (G_0 y_0)(T) - y_\Omega = Su - z,$$

where $z := y_\Omega - (G_0 y_0)(T)$. The above control problem then takes the form

$$\min_{u \in U_{ad}} f(u) := \frac{1}{2}\|Su - z\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)}.$$

The general variational inequality (2.47) yields, for all $u \in U_{ad}$,

$$\begin{aligned} 0 &\leq \big(S\overline{u} - z, S(u - \overline{u})\big)_{L^2(\Omega)} + \lambda\big(\overline{u}, u - \overline{u}\big)_{L^2(\Sigma)} \\ &= \int_\Omega \big(\overline{y}(T) - y_\Omega\big)\big(y(T) - \overline{y}(T)\big)\, dx + \lambda \iint_\Sigma \overline{u}(u - \overline{u})\, ds\, dt. \end{aligned} \tag{3.50}$$

Here, we have used the identity

$$Su - S\overline{u} = Su + (G_0 y_0)(T) - (G_0 y_0)(T) - S\overline{u} = y(T) - \overline{y}(T)$$

and, once more, that $z = y_\Omega - (G_0 y_0)(T)$.

## Proof (3.19) contd.

Now put $\tilde{y} := y - \overline{y}$ and apply Theorem 3.18 with the specifications $a_\Omega = \overline{y}(T) - y_\Omega$, $a_Q = 0$, $a_\Sigma = 0$, $b_\Sigma = \beta$, $v = 0$, $w = 0$, $y := \tilde{y}$, and $\tilde{u} := u - \overline{u}$. Also $c_0 = 0$. Note that $w = 0$ in this situation since, by definition, $Su(0) = 0$; the part originating from $y_0$ is incorporated into $z$. It follows that (in the interest of space, the first line without differentials)

$$\iint_Q a_Q y + \iint_\Sigma a_\Sigma\, y + \int_\Omega a_\Omega\, y(\cdot, T) = \iint_Q b_Q\, p\, v + \iint_\Sigma b_\Sigma\, p\, u + \int_\Omega b_\Omega\, p(\cdot, 0)\, w$$

$$\int_\Omega (\overline{y}(T) - y_\Omega)\, y(\cdot, T) = \iint_\Sigma \beta\, p\, u\, ds\, dt$$

$$\left(\overline{y}(T) - y_\Omega, \tilde{y}(T)\right)_{L^2(\Omega)} = \iint_\Sigma \beta\, p\, \tilde{u}\, ds\, dt.$$

Substituting this result into inequality (3.50), we find that

$$\begin{aligned} 0 &\le \int_\Omega \left(\overline{y}(T) - y_\Omega\right)\left(y(T) - \overline{y}(T)\right) dx + \lambda \iint_\Sigma \overline{u}(u - \overline{u})\, ds\, dt \\ &= \iint_\Sigma \beta\, p(u - \overline{u})\, ds\, dt + \lambda \iint_\Sigma \overline{u}(u - \overline{u})\, ds\, dt \\ &= \iint_\Sigma (\beta\, p + \lambda \overline{u})(u - \overline{u})\, ds\, dt, \end{aligned}$$

which concludes the proof of the assertion. □

Next, we employ the method described on page 69 for elliptic problems to derive a number of results converning the possible form of optimal controls. This is a *pointwise* view of the control, in contrast to the "$L^2$ view" in Theorem 3.19.

### Theorem 3.20

*A control* $\bar{u} \in U_{ad}$ *with associated state* $\bar{y}$ *is optimal for the problem* (3.1)-(3.3) *if and only if it satisfies, together with the adjoint state* $p$ *from* (3.48)*, the following conditions for almost all* $(x,t) \in \Sigma$*: the* weak minimum principle

$$\big(\beta(x,t)p(x,t) + \lambda\bar{u}(x,t)\big)\big(v - \bar{u}(x,t)\big) \geq 0, \qquad \forall v \in \big[u_a(x,t), u_b(x,t)\big], \tag{3.51}$$

*the* minimum principle

$$\beta(x,t)p(x,t)\bar{u}(x,t) + \frac{\lambda}{2}\bar{u}(x,t)^2 = \min_{v \in [u_a(x,t),u_b(x,t)]} \Big\{\beta(x,t)p(x,t)v + \frac{\lambda}{2}v^2\Big\}, \tag{3.52}$$

*and, in the case of* $\lambda > 0$*, the projection formula*

$$\bar{u}(x,t) = \mathbb{P}_{[u_a(x,t),u_b(x,t)]}\Big\{-\frac{1}{\lambda}\beta(x,t)p(x,t)\Big\}. \tag{3.53}$$

For the proof of this assertion one takes, starting from the variational inequality (3.49), the same series of steps that led from Theorem 2.25 to Theorem 2.27 in the elliptic case.

# Proof/Discussion (3.20)

Theorem 3.19, like Theorem 2.25, shows that, for control $\overline{u}$, associated state $\overline{y}$, and weak solution $p$ to the adjoint equation, satisfying the variational inequality, e.g., (3.49), is equivalent to optimality. Rewrite (3.49) in the form

$$\iint_{\Sigma} \left(\beta(x,t)p(x,t) + \lambda\overline{u}(x,t)\right) \underbrace{\overline{u}(x,t)}\, ds(x)\, dt \\ \leq \iint_{\Sigma} \left(\beta(x,t)p(x,t) + \lambda\overline{u}(x,t)\right) \underbrace{u(x,t)}\, ds(x)\, dt, \quad \forall u \in U_{ad}.$$

The difference between the LHS and RHS has been marked for clarity.

Hence

$$\iint_{\Sigma} \left(\beta(x,t)p(x,t) + \lambda\overline{u}(x,t)\right) \underbrace{\overline{u}(x,t)}\, ds(x)\, dt \\ = \min_{u \in U_{ad}} \iint_{\Sigma} \left(\beta(x,t)p(x,t) + \lambda\overline{u}(x,t)\right) \underbrace{u(x,t)}\, ds(x)\, dt,$$

and we can conclude that, assumping the expression inside the first parentheses is known, we obtain $\overline{u}$ as the solution to a linear optmization problem in a function space.

This observation forms the basis of the *conditioned gradient method*, discussed in the various sections on numerical methods.

# Proof/Discussion (3.20) contd.

*"It is intuitively clear..."*

We present a lemma providing insight in the direction of a "pointwise" form of solution. It is an exact parallel to the elliptic case's Lemma 2.26.

**Lemma 22 (c.f. Lemma 2.26)**

*A necessary and sufficient condition for the variational inequality* (3.49) *to be satisfied is that for almost every* $(x,t) \in Q$,

$$\overline{u}(x,t) = \begin{cases} u_a(x,t) & \text{if } \beta(x,t)p(x,t) + \lambda\overline{u}(x,t) > 0 \\ \in [u_a(x,t), u_b(x,t)] & \text{if } \beta(x,t)p(x,t) + \lambda\overline{u}(x,t) = 0 \\ u_b(x,t) & \text{if } \beta(x,t)p(x,t) + \lambda\overline{u}(x,t) < 0. \end{cases} \tag{4}$$

*An equivalent condition is given by the pointwise variational inequality in* $\mathbb{R}$,

$$\big(\beta(x,t)p(x,t)+\lambda\overline{u}(x,t)\big)\big(v-\overline{u}(x,t)\big) \geq 0, \quad \forall v \in [u_a(x,t), u_b(x,t)], \text{ for a.e. } (x,t) \in Q. \tag{5}$$

The proof of Lemma 2.26 goes through for this one, solely with the change of $x$ to $(x,t)$ and $\Omega$ to $Q$. It shows that (3.49) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (3.49).

The most illuminating portion of that proof is the first implication, adapted on the next slide.

# Proof (Lemma 22)

Let $\overline{u}$, $u_a$, and $u_b$ be arbitrary but fixed representatives of the corresponding equivalence classes in $L^\infty(Q)$. Suppose (4) is false (but the constraints are obeyed almost everywhere), and consider the measurable sets

$$A_+(\overline{u}) = \{(x,t) \in Q : \beta(x,t)p(x,t) + \lambda\overline{u}(x,t) > 0\},$$
$$A_-(\overline{u}) = \{(x,t) \in Q : \beta(x,t)p(x,t) + \lambda\overline{u}(x,t) < 0\}.$$

By our assumption, there is a set $E_+ \subset A_+(\overline{u})$ having positive measure such that $\overline{u}(x,t) > u_a(x,t)$ for all $(x,t) \in E_+$, or there is a set $E_- \subset A_-(\overline{u})$ having positive measure such that $\overline{u}(x,t) < u_b(x,t)$ for all $(x,t) \in E_-$.

In the first case, we define the function $u \in U_{ad}$,

$$u(x,t) = \begin{cases} u_a(x) & \text{for } (x,t) \in E_+ \\ \overline{u}(x,t) & \text{for } (x,t) \in Q \setminus E_+. \end{cases}$$

Then

$$\iint_Q \big(\beta(x,t)p(x,t) + \lambda\overline{u}(x,t)\big)\big(u(x,t) - \overline{u}(x,t)\big)\,dx$$
$$= \iint_{E_+} \big(\beta(x,t)p(x,t) + \lambda\overline{u}(x,t)\big)\big(u_a(x,t) - \overline{u}(x,t)\big)\,dx < 0,$$

since the first factor is positive on $E_+$ and the second is negative. This contradicts (3.49).

The other case is handled similarly. We end our incomplete proof of the lemma here.

# Proof/Discussion (3.20) contd.

By a rearrangement of terms the pointwise variational inequality (5) can be rewritten

$$\big(\beta(x,t)p(x,t)+\lambda\overline{u}(x,t)\big)\overline{u}(x,t) \le \big(\beta(x,t)p(x,t)+\lambda\overline{u}(x,t)\big)v, \;\forall v \in [u_a(x,t), u_b(x,t)], \quad (6)$$

for almost every $(x,t) \in Q$. Here $v \in \mathbb{R}$, it is not a function. We can now complete our proof of Theorem 3.20.

## Proof (Theorem 3.20)

The weak minimum principle is nothing but a reformulation of (6).

The minimum principle is easily verified: a real number $\overline{v}$ solves for fixed $(x,t)$ the convex quadratic optimization problem in $\mathbb{R}$,

$$\min_{v\in[u_a(x,t),u_b(x,t)]} g(v) := \beta(x,t)p(x,t)v + \frac{\lambda}{2}v^2, \quad (7)$$

if and only if the variational inequality

$$g'(\overline{v})(v-\overline{v}) \ge 0, \quad \forall v \in [u_a(x,t), u_b(x,t)]$$

is satisfied, that is, if

$$\big(\beta(x,t)p(x,t)+\lambda\overline{v}\big)(v-\overline{v}) \ge 0, \quad \forall v \in [u_a(x,t), u_b(x,t)].$$

The minimum condition follows from taking $\overline{v} = \overline{u}(x,t)$.

The solution to the quadratic optimization problem (7) in $\mathbb{R}$ is given by the projection formula (3.53). □

# Conclusion

## of Necessary Conditions for Example 1

In the $\lambda > 0$ case, the triple $(\overline{u}, \overline{y}, p)$ satisfies the *optimality system*

$$\boxed{\begin{aligned} y_t - \Delta y &= 0 & -p_t - \Delta p &= 0 \\ \partial_\nu y + \alpha y &= \beta u & \partial_\nu p + \alpha p &= 0 \\ y(0) &= y_0 & p(T) &= y(T) - y_\Omega \\ u &= \mathbb{P}_{[u_a, u_b]}\Big\{ -\frac{1}{\lambda}\beta\, p \Big\}. \end{aligned}} \tag{3.54}$$

If $\lambda = 0$ then the projection formula has to be replaced by:

$$u(x,t) = \begin{cases} u_a(x,t) & \text{if } \beta(x,t)p(x,t) > 0 \\ \in [u_a(x,t), u_b(x,t)] & \text{if } \beta(x,t)p(x,t) = 0 \\ u_b(x,t) & \text{if } \beta(x,t)p(x,t) < 0. \end{cases}$$

# Special Case: $u_a = -\infty, u_b = \infty$

In the case there are no control constraints, the projection formula yields $u = -\lambda^{-1}\beta p$. Hence $u$ can be eliminated from the state equation, and we obtain the following *forward-backward system* of two parabolic problems for the unknown functions $y$ and $p$:

$$\boxed{\begin{array}{ll} y_t - \Delta y = 0 & -p_t - \Delta p = 0 \\ \partial_\nu y + \alpha y = -\beta^2\lambda^{-1} p & \partial_\nu p + \alpha p = 0 \\ y(0) = y_0 & p(T) = y(T) - y_\Omega \end{array}} \tag{3.55}$$

The solution of such systems is not an easy task. (See page 162 for brief discussion of methods.)

# Optimal nonstationary heat source

We recall problem (3.38)-(3.40), which in shortened form reads

$$\min J(y,u) := \frac{1}{2}\|y - y_\Sigma\|^2_{L^2(\Sigma)} + \frac{\lambda}{2}\|u\|^2_{L^2(Q)},$$

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

$$\boxed{\begin{aligned} y_t - \Delta y &= \beta u \\ \partial_\nu y &= 0 \\ y(0) &= 0 \end{aligned}}$$

Invoking the operator $G_Q : L^2(Q) \to W(0,T)$ introduced in (3.36), we can express the solution $y$ to the state system in the form

$$y = G_Q(\beta u).$$

The cost functional involves the observation $y\big|_\Sigma$. The observation operator $E_\Sigma : y \mapsto y\big|_\Sigma$ is a continuous linear mapping from $W(0,T)$ into $L^2(0,T;L^2(\Gamma)) \cong L^2(\Sigma)$, which entail that the *control-to-observation operator* $S : u \mapsto y\big|_\Sigma$ defined in (3.41) is continuous from $L^2(Q)$ into $L^2(\Sigma)$. Hence, the problem is equivalent to the problem $\min_{u\in U_{ad}} f(u)$, where $f$ is the reduced functional introduced in (3.42).

GEORGE MASON UNIVERSITY

As in (3.50), we obtain the following as the necessary optimality condition for $\overline{u}$:

$$0 \leq (S\overline{u} - y_\Sigma, Su - S\overline{u})_{L^2(\Sigma)} + \lambda(\overline{u}, u - \overline{u})_{L^2(Q)} \quad \forall u \in U_{ad},$$

that is, upon substituting $\overline{y}\big|_\Sigma = S\overline{u}$ and $y\big|_\Sigma = Su$,

$$0 \leq \iint_\Sigma (\overline{y} - y_\Sigma)(y - \overline{y})\, ds\, dt + \lambda \iint_Q \overline{u}(u - \overline{u})\, dx\, dt \quad \forall u \in U_{ad}. \tag{3.56}$$

It is evident how the adjoint state $p$ must be defined, namely as the weak solution to the parabolic problem

$$\boxed{\begin{aligned} -p_t - \Delta p &= 0 \\ \partial_\nu p &= \overline{y} - y_\Sigma \\ p(T) &= 0 \end{aligned}}$$

By virtue of Lemma 3.17, it has a unique weak solution $p$.

**Theorem 3.21**

*A control* $\overline{u} \in U_{ad}$ *is optimal for the optimal nonstationary heat source problem* (3.38)-(3.40) *if and only if it satisfies, together with the adjoint state p defined above, the variational inequality*

$$\iint_Q (\beta p + \lambda \overline{u})\, dx\, dt \geq 0 \qquad \forall u \in U_{ad}.$$

**Proof (3.21)**

This assertion is again a direct consequence of Theorem 3.18, with the specifications $a_\Sigma = \overline{y} - y_\Sigma$, $a_\Omega = 0$, $a_Q = 0$, $b_Q = \beta$, $b_\Sigma = 0$, and $b_Q = 0$. The steps are similar to those in the proof of Theorem 3.20. □

As in the previous section, the variational inequality just proved can be transformed into an equivalent pointwise minimum principle for $\overline{u}$ or, if $\lambda > 0$, into a projection formula. In particular, if $\lambda > 0$,

$$\overline{u}(x,t) = \mathbb{P}_{[u_a(x,t),u_b(x,t)]}\left\{ -\frac{1}{\lambda}\beta(x,t)p(x,t)\right\} \quad \text{for a.e. } (x,t) \in Q.$$

# Overview of Numerical Methods

In the following slides we sketch the projected gradient method for the optimal nonstationary boundary temperature problem (the "sketch" in the text spans 5 pages). Here we briefly mention other published methods which are computationally more efficient:

- Multigrid methods (for the unconstrained case)
  - In 1- and 2-dimensional problems, many methods for solving (3.55) do not require multigrid
- Primal-dual active set
  - In each iteration we update the active set for the upper and lower constraints
    - using $-\lambda^{-1}f'(u) \lesseqgtr 0$, as in section 2.12.4
  - Then solve the unconstrained forward-backward problem for the remaining values
- Direct solution of the optimality system
  - Substitute $u = \mathbb{P}_{[u_a,u_b]}\{-\lambda^{-1}\beta p\}$ in the state equation
  - The system is differentiable except at $u_a$ and $u_b$
  - Good results have been achieved in solving the nonsmooth system
- Discretize then optimize
  - Fully discretize the parabolic problem and the cost functional
  - Use existing solvers for finite-dimensional quadratic optimization problems

# Projected Gradient Method

As before, let $S : u \mapsto y(T)$ for $y_0 = 0$, and let $\widehat{y}$ be the solution to the homogeneous problem with nonzero initial condition $y_0$, so that

$$y(x, T) = (Su)(x) + \widehat{y}(x, T).$$

In this way the problem becomes a quadratic Hilbert space optimization problem,

$$\min_{u \in U_{ad}} f(u) := \frac{1}{2}\|Su + \widehat{y}(T) - y_\Omega\|^2_{L^2(\Omega)} + \frac{\lambda}{2}\|u\|^2_{L^2(\Sigma)}.$$

We follow the algorithm in iterations, starting with an initial guess $u_0$, which generates $y_0$, then $p_0$, then $u_1$, $y_1$, $p_1$, $u_2$, $y_2$, $p_2$, etc., each in turn. The $y_n$ and $p_n$ iterates are (separately) the solution of the now familiar state and adjoint equations,

$$\begin{aligned} y_t - \Delta y &= 0 & \qquad -p_t - \Delta p &= 0 \\ \partial_\nu y + \alpha y &= \beta u & \qquad \partial_\nu p + \alpha p &= 0 \\ y(0) &= y_0 & \qquad p(T) &= y_n(T) - y_\Omega. \end{aligned}$$

The derivative at an iterate $u_n$ is

$$f'(u_n)v = \iint_\Sigma \big(\beta(x,t)p_n(x,t) + \lambda u_n(x,t)\big)\, v(x,t)\, ds\, dt,$$

By the Riesz representation theorem, we obtain the usual representation of the reduced gradient,

$$f'(u_n) = \beta p_n + \lambda u_n.$$

# Projected Gradient Method contd.

## The Algorithm

The algorithm proceeds as follows, suppose that iterates $u_1, \ldots, u_n$ have been determined.

**S1** *(New State)* Using $u_n$ solve the state equation for $y_n$.

**S2** *(New descent direction)* Using $y_n$ solve the adjoint equation for $p_n$. Take as descent direction the negative gradient

$$v_n = -f'(u_n) = -(\beta p_n + \lambda u_n).$$

**S3** *(Step size control)* Determine the optimal step size $s_n$ by solving

$$f\left(\mathbb{P}_{[u_a,u_b]}\{u_n + s_n v_n\}\right) = \min_{s>0} f\left(\mathbb{P}_{[u_a,u_b]}\{u_n + s_n v_n\}\right).$$

**S4** Put $u_{n+1} := \mathbb{P}_{[u_a,u_b]}\{u_n + s_n v_n\}, \quad n := n+1,$ GO TO **S1**.

The method is completely analogous to that for the elliptic case. The projection step is necessary because $u_n + s_n v_n$ may not be admissible.

Although the method converges only slowly, it is easy to implement and thus very suitable for numerical tests.

If we need to calculate the solution many times for different data problem many times for different data, reducing the problem can save considerable calculation. Section 3.7.2 gives an example of this.

# Bibliography

Optimal Control of Partial Differential Equations; Theory, Methods and Applications. Fredi Trölzsch. Graduate Studies in Mathematics, Volume 112. American Mathematical Society. http://www.ams.org/bookpages/gsm-112