“Notes on the simulation of the heat equation in Firedrake”

October 10, 2017; rev. October 27, 2017

Simon Pirkelmann

1 Setting

Consider the following equation

$$y_t - a\Delta y = 0 \text{ on } \Omega \tag{1}$$

where $y : \Omega \times [0, \infty) \to \mathbb{R}$ is the temperature, $a \in \mathbb{R}$ is the diffusion coefficient. We use the shorthand $y_t = \frac{\partial y}{\partial t}$ to denote the time derivative.

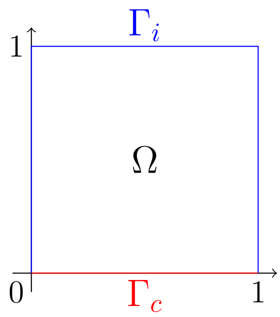
As a domain we consider the unit square. The boundary is partitioned in an isolating boundary Γ_i and a control boundary Γ_c. On the control part of the boundary heating/cooling is applied which is described by a Dirichlet condition

$$y = u \text{ on } \Gamma_c. \tag{2}$$

On the isolating part we have

$$-\frac{\partial y}{\partial n} = 0 \text{ on } \Gamma_i \tag{3}$$

where $\frac{\partial y}{\partial n}$ is the derivative of y in normal direction.



For simplicity we use a single Robin type boundary condition instead

$$-\frac{\partial y}{\partial n} = \gamma(y - u) \text{ on } \Gamma. \tag{4}$$

By choosing $\begin{cases} \gamma = 0 & \text{on } \Gamma_i \\ \gamma = 10^3 & \text{on } \Gamma_c \end{cases}$, we can approximate both types of boundary conditions in a uniform way. This will also allow us to extend the setting more easily in the future.

2 Weak Form

For the weak for of the equation we replace y_t by $\frac{y_{k+1} - y_k}{h}$ and y by y_{k+1} using backward Euler with sampling rate $h > 0$. Multiplying with a test function v and integrating yields

$$\int_\Omega \frac{y_{k+1} - y_k}{h} v \, dx - a \int_\Omega \Delta y_{k+1} v \, dx = 0 \tag{5}$$

Using partial integration on the second integral we get

$$\int_\Omega \frac{y_{k+1} - y_k}{h} v \; dx - a \int_\Gamma \frac{\partial y_{k+1}}{\partial n} v \; ds + a \int_\Omega \nabla y_{k+1} \cdot \nabla v \; dx = 0 \tag{6}$$

Substituting the boundary condition (4) we obtain

$$\int_\Omega \frac{y_{k+1} - y_k}{h} v \; dx + a\gamma \int_\Gamma (y_{k+1} - u) v \; ds + a \int_\Omega \nabla y_{k+1} \cdot \nabla v \; dx = 0 \tag{7}$$

We can group the terms in the previous equations by terms that depend on y_{k+1}, and terms that do not depend on y_{k+1} (an thus only depend on the external data (y_k and u):

$$\underbrace{\int_\Omega y_{k+1} v \; dx + ha\gamma \int_\Gamma y_{k+1} v \; ds + ha \int_\Omega \nabla y_{k+1} \cdot \nabla v \; dx}_{a(y_{k+1}, v)} = \underbrace{\int_\Omega y_k v \; dx + ha\gamma \int_\Gamma uv \; ds}_{F(v)} \tag{8}$$

3 Implementation in Firedrake

To solve the equation $a(y_{k+1}, v) = F(v)$ in Firedrake we first need to define the bilinear-form (?) a and the functional F.

```
# Definitions
S = FunctionSpace(self.mesh, "CG", 1)
v = TestFunction(S)
y1 = Function(S)
y0 = Function(S)

h = Constant(0.001)
gamma = [0.0, 0.0, 1.0e3, 0.0] # gamma for different parts of the the
                                       boundary
ac = 0.2
u = [0.0, 0.0, 25.0, 0.0]
```

```
a = (y1 * v + h * ac * inner(grad(y1), grad(v))) * dx
for i in range(1,5):
    a += h * ac * Constant(gamma[i-1]) * y1 * v * ds(i)
```

```
F = inner(y0, v) * dx
for i in range(1, 5):
    F += h * ac * Constant(gamma[i-1]) * Constant(u[i-1]) * v * ds(i)
```

4 Optimal Control Problem

Now that the simulation is running we want to implement an optimal control problem on top of it. Our goal is to solve the following problem:

$$\begin{aligned}
\min_{y,u} J(y,u) =& \frac{1}{2}\int_\Omega (y(x,T)-y_\Omega(T,x))^2\,dx + \frac{1}{2}\int_0^T\int_\Omega (y(x,t)-y_\Omega(t,x))^2\,dx\,dt\\
&+\frac{\lambda}{2}\int_0^T\int_\Gamma (u(x,t))^2\,ds\,dt\\
&\text{s.t.}(1),(4)\\
&\quad \underline{u}(x,t)\le u(x,t)\le \overline{u}(x,t)\\
&\quad \underline{y}(x,t)\le y(x,t)\le \overline{y}(x,t)
\end{aligned}$$

While the constraints on the control can be dealt with rather straightforwardly, the state constraints present some challenges. We will use Lavrentiev regularisation to handle the state constraints. The Lavrentiev regularisation replaces the state constraint by a mixed state-control constraint. For this we introduce an additional control variable $v:\Omega\times[0,\infty)\to\mathbb{R}$ defined on the whole domain. This control variable will also be penalized in the cost functional, and so the cost functional is modified to

$$\begin{aligned}
\min_{u,v} J(y,u,v) =& \frac{1}{2}\int_\Omega (y(x,T)-y_\Omega(T,x))^2\,dx + \frac{1}{2}\int_0^T\int_\Omega (y(x,t)-y_\Omega(t,x))^2\,dx\,dt\\
&\frac{\sigma}{2}\int_0^T\int_\Omega (v(x,t))^2\,dx\,dt + \frac{\lambda}{2}\int_0^T\int_\Gamma (u(x,t))^2\,ds\,dt
\end{aligned}$$

The state constraint is replaced by the auxiliary control constraint

$$\begin{aligned}
&\underline{y}(x,t)\le y(x,t)+\varepsilon v(x,t)\le \overline{y}(x,t)\\
\Leftrightarrow\ &\underbrace{\frac{1}{\varepsilon}(\underline{y}(x,t)-y(x,t))}_{\underline{v}(x,t)}\le v(x,t)\le \underbrace{\frac{1}{\varepsilon}(\overline{y}(x,t)-y(x,t))}_{\overline{v}(x,t)}
\end{aligned}$$

This kind of optimal control problem (with both a control on the boundary and a control in the domain) is considered in [Tröltzsch, p.221ff]. We look at the general cost functional

$$\begin{aligned}
\min_{y,u,v} J(y,u,v) =& \int_\Omega \phi(x,y(T))\,dx + \int\int_Q \varphi(x,t,y,v)\,dx\,dt\\
&+\int\int_\Sigma \psi(x,t,y,u)\,ds\,dt
\end{aligned}$$

subject to

$$\begin{aligned}
y_t-\Delta y+d(x,t,y,v)&=0 && \text{in } Q\\
\partial_n y+b(x,t,y,u)&=0 && \text{in } \Sigma\\
y(\cdot,0)&=y_0 && \text{on } \Omega\\
v_a(x,t)\le v(x,t)&\le v_b(x,t) && \text{in } Q\\
u_a(x,t)\le u(x,t)&\le u_b(x,t) && \text{in } \Sigma
\end{aligned}$$

In our case we have the following identities:

$$\begin{aligned}\phi(x,y(x,T)) &= \frac{1}{2}(y(x,T)-y_\Omega(x,T))^2\\ \phi_y(x,y(x,T)) &= y(x,T)-y_\Omega(x,T)\end{aligned}$$

$$\begin{aligned}\varphi(x,t,y(x,t),v(x,t)) &= \frac{1}{2}(y(x,t)-y_\Omega(x,t))^2+\frac{\sigma}{2}(v(x,t))^2\\ \varphi_y(x,t,y(x,t),v(x,t)) &= y(x,t)-y_\Omega(x,t)\\ \varphi_v(x,t,y(x,t),v(x,t)) &= \sigma v(x,t)\end{aligned}$$

$$\begin{aligned}\psi(x,t,y(x,t),u(x,t)) &= \frac{\lambda}{2}(u(x,t))^2\\ \psi_y(x,t,y(x,t),u(x,t)) &= 0\\ \psi_u(x,t,y(x,t),u(x,t)) &= \lambda u(x,t)\end{aligned}$$

$$\begin{aligned}d(x,t,y(x,t),v(x,t)) &= 0\\ d_y(x,t,y(x,t),v(x,t)) &= 0\\ d_v(x,t,y(x,t),v(x,t)) &= 0\end{aligned}$$

$$\begin{aligned}b(x,t,y(x,t),u(x,t)) &= \gamma(x,t)(y(x,t)-u(x,t))\\ b_y(x,t,y(x,t),u(x,t)) &= \gamma(x,t)\\ b_u(x,t,y(x,t),u(x,t)) &= -\gamma(x,t)\end{aligned}$$

5 Derivation of the adjoint system

We introduce the adjoint states p_1 and p_2 to remove the PDE equality constraints. The Lagrangian for the problem is given by

$$\begin{aligned}\mathcal{L}(y,u,v,p_1,p_2) =& J(y,u,v) - \int_0^T\int_\Omega (\frac{\partial y}{\partial t} - \Delta y + d(x,t,y,v))p_1\ dx\ dt\\ &- \int_0^T\int_\Gamma (\partial_n y + b(x,t,y,u))p_2\ dx\ dt\end{aligned}$$

For optimality we need

$$\begin{aligned}D_y\mathcal{L}(\bar y,\bar u,\bar v,p_1,p_2)y &= 0, \text{ for all } y \text{ with } y(\cdot,0)=0\\ D_u\mathcal{L}(\bar y,\bar u,\bar v,p_1,p_2)(u-\bar u) &\geq 0, \qquad \text{for all } u\in U_{ad}\\ D_v\mathcal{L}(\bar y,\bar u,\bar v,p_1,p_2)(v-\bar v) &\geq 0, \qquad \text{for all } v\in V_{ad}\end{aligned}$$

In the following we compute $D_y\mathcal{L}$, $D_u\mathcal{L}$ and $D_v\mathcal{L}$. We start with $D_y\mathcal{L}$:

$$\begin{aligned}D_y\mathcal{L}(y,u,v,p_1,p_2)h =& D_yJ(y,u,v)h - \int_0^T\int_\Omega (\frac{\partial h}{\partial t} - \Delta h + d_y(x,t,y,v)h)p_1\ dx\ dt\\ &- \int_0^T\int_\Gamma (\partial_n h + b_y(x,t,y,u)h)p_2\ dx\ dt,\end{aligned}$$

where for $D_yJ(y,u,v)$ it holds that

$$D_yJ(y,u,v)h = \int_\Omega \phi_y(x,y)h(T)\ dx + \int_0^T \int_\Omega \varphi_y(x,t,y,v)h\ dx\ dt + \int_0^T \int_\Gamma \psi_y(x,t,y,u)h\ ds\ dt. \tag{9}$$

First we consider the term

$$\begin{aligned}
&-\int_0^T \int_\Omega (\frac{\partial h}{\partial t} - \Delta h + d_y(x,t,y,v)h)p_1\ dx\ dt \\
&= -\int_\Omega \int_0^T \frac{\partial h}{\partial t} p_1\ dx\ dt + \int_0^T \int_\Omega \Delta h p_1\ dx\ dt - \int_0^T \int_\Omega d_y(x,t,y,v)hp_1\ dx\ dt.
\end{aligned}$$

Using partial integration (in time) for the first term and Green's second formula (in space) for the second term we obtain

$$\begin{aligned}
&-\int_\Omega \int_0^T \frac{\partial h}{\partial t} p_1\ dx\ dt + \int_0^T \int_\Omega \Delta h p_1\ dx\ dt - \int_0^T \int_\Omega d_y(x,t,y,v)hp_1\ dx\ dt \\
&= -\int_\Omega [hp_1]_0^T\ dx + \int_\Omega \int_0^T \frac{\partial p_1}{\partial t} h\ dx\ dt + \int_0^T \int_\Omega \Delta p_1 h\ dx\ dt \\
&\quad + \int_0^T \int_\Gamma p_1 \frac{\partial h}{\partial n}\ ds\ dt - \int_0^T \int_\Gamma \frac{\partial p_1}{\partial n} h\ ds\ dt - \int_0^T \int_\Omega d_y(x,t,y,v)p_1 h\ dx\ dt \\
&= \int_\Omega p_1(0)h(0)\ dx - \int_\Omega p_1(T)h(T)\ dx + \int_0^T \int_\Omega \frac{\partial p_1}{\partial t} h\ dx\ dt + \int_0^T \int_\Omega \Delta p_1 h\ dx\ dt \\
&\quad + \int_0^T \int_\Gamma p_1 \frac{\partial h}{\partial n}\ ds\ dt - \int_0^T \int_\Gamma \frac{\partial p_1}{\partial n} h\ ds\ dt - \int_0^T \int_\Omega d_y(x,t,y,v)p_1 h\ dx\ dt \\
&\overset{h(0)=0}{=} -\int_\Omega p_1(T)h(T)\ dx + \int_\Omega \int_0^T \frac{\partial p_1}{\partial t} h\ dx\ dt + \int_0^T \int_\Omega \Delta p_1 h\ dx\ dt \\
&\quad + \int_0^T \int_\Gamma p_1 \frac{\partial h}{\partial n}\ ds\ dt - \int_0^T \int_\Gamma \frac{\partial p_1}{\partial n} h\ ds\ dt - \int_0^T \int_\Omega d_y(x,t,y,v)p_1 h\ dx\ dt
\end{aligned}$$

In the last step we used that $h(0) = 0$. This is explained in [Tröltzsch, p.96] and follows from a substitution $y := y - \bar{y}$ for the state.
By substituting the above equation in the original equation and ordering by integration domains we obtain

$$\begin{aligned}
D_yJ(y,u,v)h = &\int_\Omega (\phi_y(x,y) - p_1(T))h(T)\ dx \\
&+ \int_0^T \int_\Omega (\frac{\partial p_1}{\partial t} + \Delta p_1 - d_y(x,t,y,v)p_1 + \varphi_y(x,t,y,v))h\ dx\ dt \\
&+ \int_0^T \int_\Gamma (p_1 - p_2)\frac{\partial h}{\partial n}\ ds\ dt \\
&+ \int_0^T \int_\Gamma (-\frac{\partial p_1}{\partial n} - b_y(x,t,y,u)p_2 + \psi_y(x,t,y,u))h\ ds\ dt
\end{aligned}$$

Now we make special choices of h, first $h \in C_0^\infty(Q)$, then with arbitrary $h(T)$, arbitrary $h|_\Sigma$ and finally with arbitrary $\frac{\partial h}{\partial n}$, and we set $p := p_1$, $p_2 = p$ on Σ. From this we obtain the following adjoint equation:

$$\begin{aligned} -p_t - \Delta p + d_y(x,t,\bar{y},\bar{v})p &= \varphi_y(x,t,\bar{y},\bar{v}) \quad \text{in } Q \\ \partial_n p + b_y(x,t,\bar{y},\bar{u})p &= \psi_y(x,t,\bar{y},\bar{u}) \quad \text{in } \Sigma \\ p(x,T) &= \phi_y(x,\bar{y}(x,T)) \text{ in } \Omega \end{aligned} \tag{10}$$

This can also be found in [Tröltzsch, p.225f].
Finally we also compute $D_u\mathcal{L}$ and $D_v\mathcal{L}$:

$$\begin{aligned} D_u\mathcal{L}(y,u,v,p)h &= D_uJ(y,u,v)h - \int_0^T \int_\Gamma b_u(x,t,y,u)ph\; ds\; dt \\ &= \int_0^T \int_\Gamma (\psi_u(x,t,y,u) - b_u(x,t,y,u)p)h\; ds\; dt \end{aligned}$$

$$\begin{aligned} D_v\mathcal{L}(y,u,v,p)h &= D_vJ(y,u,v)h - \int_0^T \int_\Omega d_v(x,t,y,v)ph\; dx\; dt \\ &= \int_0^T \int_\Omega (\varphi_v(x,t,y,v) - d_v(x,t,y,v)p)h\; dx\; dt \end{aligned}$$

6 Weak form of the adjoint system

To derive the weak form of the adjoint system (10) we use backward Euler and replace p_t by $\frac{p_{k+1}-p_k}{h}$. We multiply with the test function v and integrate:

$$-\int_\Omega \frac{p_{k+1}-p_k}{h} v\; dx - \int_\Omega \Delta p_{k+1} v\; dx + \int_\Omega d_y(x,t_k,y_k,v_k)p_{k+1}v\; dx = \int_\Omega \varphi_y(x,t,y_k,v_k)v\; dx$$

Using partial integration on the second integral and inserting the boundary condition yields

$$\begin{aligned} -\int_\Omega \Delta p_{k+1} v\; dx &= -\int_\Gamma \frac{\partial p_{k+1}}{\partial n} v ds + \int_\Omega \nabla p_{k+1} \nabla v\; dx \\ &= \int_\Gamma b_y(x,t_k,y_k,u_k)p_{k+1}v\; ds - \int_\Gamma \psi_y(x,t_k,y_k,u_k)v\; ds + \int_\Omega \nabla p_{k+1}\nabla v\; dx \end{aligned}$$

We insert this in the first equation again and order by integration domain

$$\begin{aligned} \int_\Omega (-\frac{p_{k+1}-p_k}{h} + d_y(x,t_k,y_k,v_k)p_{k+1} - \varphi_y(x,t,y_k,v_k))v + \nabla p_{k+1}\nabla v dx \\ + \int_\Gamma (b_y(x,t_k,y_k,u_k)p_{k+1} - \psi_y(x,t_k,y_k,u_k))v\; ds = 0 \end{aligned}$$

6.1 Adjoint equation for our case

In our case the equation simplifies to:

$$\int_\Omega (-\frac{p_{k+1}-p_k}{h} - (y_k - y_{\Omega,k}))v + \nabla p_{k+1}\nabla v\; dx + \int_\Gamma \gamma p_{k+1} v\; ds = 0$$

Additionally, we have the initial (or rather terminal) condition:

$$p_{N+1} = \phi_y(x,y_{N+1}) = y_{N+1} - y_{\Omega,N+1}$$

7 Solution by Projected Gradient Method

Let $N \in \mathbb{N}$ be the MPC horizon and let $u^n := (u_0^n, u_1^n, \dots, u_{N-1}^n)$, $v^n := (v_0^n, v_1^n, \dots, v_{N-1}^n)$ be the iterates of the optimization algorithm.
The gradient of the reduced cost functional $f(v,u) = J(y(v,u),v,u)$ is given by

$$\begin{aligned} f'(v^n,u^n)(v,u) = &\int \int_Q (\varphi_v(x,t,y^n,v^n) - d_v(x,t,y,v)p^n)\ v\ dx\ dt \\ &+ \int \int_\Sigma (\psi_u(x,t,y^n,u^n) - b_u(x,t,y,u)p^n)\ u\ ds\ dt \end{aligned}$$

This can also be found in [Tröltzsch, p. 243f], for the special case of $d(x,t,y,v) = v$, $b(x,t,y,u) = u$.

Solution algorithm:

1. Solve forward system for given $(u^n, v^n) \rightsquigarrow y_n$

2. Solve adjoint system $\rightsquigarrow p_n$

3. Descent directions

$$\begin{aligned} h_n &:= -(\varphi_v(\cdot, y_n, v_n) - d_v(\cdot, y_n, v_n)p_n) \\ r_n &:= -(\psi_u(\cdot, y_n|_\Sigma, u_n) - b_u(\cdot, y_n, u_n)p_n|_\Sigma) \end{aligned}$$

4. Compute step size $\rightsquigarrow s_n$ (e.g. use $\min_{s>0} f(\mathbb{P}_V(v_n + sh_n), \mathbb{P}_U(u_n + sr_n))$.

5. New iterates:

$$(v_{n+1}, u_{n+1}) := (\mathbb{P}_V(v_n + s_n h_n), \mathbb{P}_U(u_n + s_n r_n)) \tag{11}$$

8 Solver Options/Preconditioning

Reference