Numerical Optimization of Partial Differential Equations

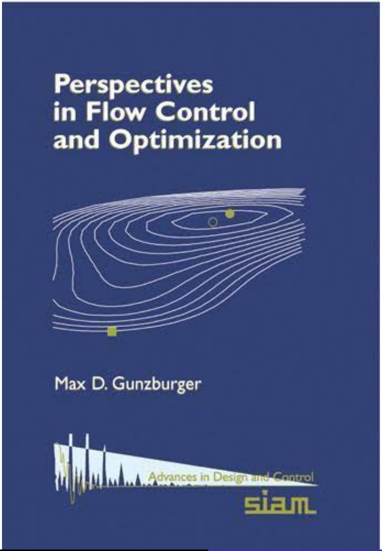
Part II: optimization with PDE constraints

Pritpal 'Pip' Matharu and Bartosz Protas

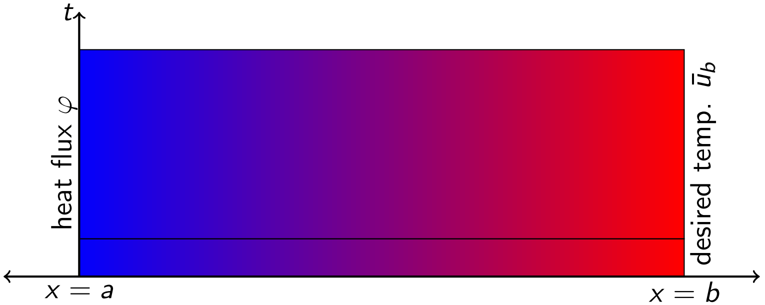
Department of Mathematics & Statistics
McMaster University, Hamilton, Ontario, Canada
URL: http://www.math.mcmaster.ca/bprotas

Rencontres Normandes sur les aspects théoriques et numériques des EDP
5–9 November 2018, Rouen

A good reference for standard approaches



- Consider heat conduction in a bar. How do we choose the heat flux φ applied at the left endpoint $(x = a)$, so that the temperature at the right endpoint $(x = b)$ has a desired time-history $\bar{u}_b = \bar{u}_b(t)$?



- Using the heat flux φ as the control variable, we formulate this problem as minimization of a (reduced) least-squares cost functional

$$\mathcal{J}(\varphi) = \frac{1}{2}\int_0^T [u(\varphi)|_b - \bar{u}_b]^2 \, dt$$

- Since $u = u(\varphi)$, we thus have the following optimization problem

$$\min_{\varphi} \mathcal{J}(\varphi) \quad \text{subject to} \quad \begin{cases} \frac{\partial u}{\partial t} - \Delta u = 0, & (t,x) \in [0,T] \times [a,b] \\ \frac{\partial u}{\partial x}\big|_{x=a} = \varphi(t), & t \in [0,T] \\ \frac{\partial u}{\partial x}\big|_{x=b} = 0, & t \in [0,T] \\ u(x, t=0) = u_0(x), & x \in [a,b] \end{cases}$$

where:

- $a, b, T \in \mathbb{R}$ are given parameters
- u_0 is an appropriate initial condition

- We wish to find the optimal boundary data (heat flux) $\widetilde{\varphi}$ such that

$$\widetilde{\varphi} = \operatorname{argmin}_{\varphi \in \mathcal{U}} \mathcal{J}(\varphi)$$

  where $\mathcal{U}$ is a suitable Hilbert space of functions $\varphi : [0, T] \to \mathbb{R}$

- The optimal control $\widetilde{\varphi}$ can be computed using a gradient descent algorithm as $\widetilde{\varphi} = \lim_{n\to\infty} \varphi^{(n)}$, where

$$\begin{cases} \varphi^{(n+1)} & = \varphi^{(n)} - \tau^{(n)} \nabla_{\varphi} \mathcal{J}(\varphi^{(n)}), \qquad n = 1, 2, \ldots \\ \varphi^{(1)} & = \varphi_0 \end{cases}$$

  - $\nabla_{\varphi} \mathcal{J}(\varphi)$ is the gradient (sensitivity) of the cost functional with respect to the control variable
  - $\tau^{(n)}$ is step length along the descent direction at the n-th iteration
  - φ_0 is the initial guess for the heat flux

Gâteaux Differential

- To determine the gradient $\nabla_\varphi \mathcal{J}(\varphi)$, we must compute the Gâteaux (directional) differential of the cost functional

$$\mathcal{J}'(\varphi;\varphi') = \lim_{\epsilon\to 0} \frac{\mathcal{J}(\varphi+\epsilon\varphi') - \mathcal{J}(\varphi)}{\epsilon} = \frac{d}{d\epsilon}\mathcal{J}\left(\varphi+\epsilon\varphi'\right)\Big|_{\epsilon=0}$$
$$= \int_0^T \left[u(\varphi)|_b - \bar{u}_b\right] u'(x,t;\varphi,\varphi')\, dt$$

  where:
  - $u'(x,t;\varphi,\varphi')$ is the perturbation variable that satisfies the *linearization* of the governing system
  - $\varphi'(t)$ is an arbitrary perturbation of the control variable $\varphi(t)$

- A (local) minimizer of the functional $\mathcal{J}(\varphi)$ is characterized by the condition

$$\forall\, \varphi' \in \mathcal{U} \qquad \mathcal{J}'(\widetilde{\varphi};\varphi') = 0$$

Perturbation System

- The perturbation system for $u'(x,t;\varphi,\varphi')$ is obtained by linearizing the governing system system around the state $u(\varphi)$

$$\begin{cases} \frac{\partial u'}{\partial t} - \Delta u' = 0 \\ \frac{\partial u'}{\partial x}\Big|_{x=a} = \varphi'(t) \\ \frac{\partial u'}{\partial x}\Big|_{x=b} = 0 \\ u'(x,t=0) = 0 \end{cases}$$

- In the present problem the governing system is linear, hence the perturbation system has
  - an identical operator (equation),
  - different data (boundary and initial conditions)
- In general, the governing and perturbations systems are defined in terms of different operators (nonlinear vs. linear)

- The following fundamental result from functional analysis will allow to extract the gradient $\nabla_\varphi \mathcal{J}(\varphi)$ from the Gâteaux differential $\mathcal{J}'(\varphi;\varphi')$

Theorem (Riesz Representation Theorem)

Let $\mathcal{X}$ be a Hilbert space. Then any bounded linear functional $h(x)$ defined on $\mathcal{X}$ ($x \in \mathcal{X}$) can be uniquely written as $h(x) = \langle x, y \rangle_{\mathcal{X}}$ for some $y \in \mathcal{X}$ (the element y is referred to as the "Riesz representer").

- Since $\forall \varphi \in \mathcal{U}$ the Gâteaux differential

$$\mathcal{J}'(\varphi;\cdot) \; : \; \mathcal{U} \to \mathbb{R}$$

is a bounded linear functional, we have the Riesz representation

$$\mathcal{J}'(\varphi;\varphi') = \left\langle \nabla_\varphi \mathcal{J}, \varphi' \right\rangle_{\mathcal{U}}$$

The gradient $\nabla_\varphi \mathcal{J}$ is thus the Riesz representer!

- However, the Gâteaux differential

$$\mathcal{J}'(\varphi;\varphi') = \int_0^T [u(\varphi)|_b - \bar{u}_b]\, u'(x,t;\varphi,\varphi')\, dt$$

  is not yet consistent with the Riesz representation, because the perturbation variable φ' does not appear explicitly in it, but is hidden in the boundary condition of the perturbation system

- To convert the Gâteaux differential $\mathcal{J}'(\varphi;\varphi')$ we will use the *adjoint calculus*
  - let $u^* \ : \ [a,b]\times[0,T] \to \mathbb{R}$ be the “adjoint state”

- Let the adjoint variable u^* satisfy the following judiciously chosen *adjoint system*

$$
\begin{cases}
-\dfrac{\partial u^*}{\partial t} - \Delta u^* = 0 \\
\dfrac{\partial u^*}{\partial x}\Big|_{x=a} = 0 \\
\dfrac{\partial u^*}{\partial x}\Big|_{x=b} = u(\varphi)|_b - \bar{u}_b \qquad \Longleftarrow \\
u^*(x, t = T) = 0
\end{cases}
$$

- The "forcing term" in the boundary condition at $x = b$ is related to the Gâteaux differential
- Note that this is a *terminal-value* problem, so we must solve this system backwards in time!
  - however, the term with the time derivative has a negative sign, so the problem is well posed
- Now we will now demonstrate that the adjoint system defined in this particular way will allow us to determine the gradient $\nabla_{\varphi}\mathcal{J}$

- Start by integrating the perturbation system against the adjoint field u^* over space and time
  Then integrate by parts with respect to space (x) and time (t)

$$
\begin{aligned}
0 &= \int_0^T \int_a^b \left(\frac{\partial u'}{\partial t} - \Delta u' \right) u^* \, dx \, dt \\
&= \int_0^T \int_a^b \underbrace{\left(-\frac{\partial u^*}{\partial t} - \Delta u^* \right)}_{=0} u' \, dx \, dt + \int_a^b \left[u^* \, u' \right] \Big|_{t=0}^{T} dx \\
&\quad - \int_0^T \left[u^* \frac{\partial u'}{\partial x} \right] \Big|_{x=a}^{b} dt + \int_0^T \left[\frac{\partial u^*}{\partial x} \, u' \right] \Big|_{x=a}^{b} dt = 0
\end{aligned}
$$

- We will now analyze the boundary terms resulting from the integration by parts

$$0 = \int_a^b \left[u^* \, u'\right] \Big|_{t=0}^{T} dx - \int_0^T \left[u^* \, \frac{\partial u'}{\partial x}\right] \Bigg|_{x=a}^{b} - \left[\frac{\partial u^*}{\partial x} \, u'\right] \Bigg|_{x=a}^{b} dt$$

$$0 = \int_a^b \underbrace{u^*}_{=0} \, u'\Big|_{t=T} - u^* \, \underbrace{u'}_{=0}\Big|_{t=0} dx$$

$$- \int_0^T u^* \, \underbrace{\frac{\partial u'}{\partial x}}_{=0}\Bigg|_{x=b} - u^* \, \underbrace{\frac{\partial u'}{\partial x}}_{=\varphi'}\Bigg|_{x=a} dt$$

$$+ \int_0^T \underbrace{\frac{\partial u^*}{\partial x}}_{=u(\varphi)|_b - \bar{u}_b} \quad u'\Bigg|_{x=b} - \underbrace{\frac{\partial u^*}{\partial x}}_{=0} \, u'\Bigg|_{x=a} dt$$

$$\Longrightarrow \quad \underbrace{\int_0^T \left[u(\varphi)|_b - \bar{u}_b\right] \, u'\Bigg|_{x=b} dt}_{\mathcal{J}'(\varphi;\varphi')} = \int_0^T -u^*\Bigg|_{x=a} \varphi' \, dt$$

- Thus, choosing $\mathcal{U} = L^2(0, T)$, we obtain an expression for the L^2 gradient of the cost functional

$$\mathcal{J}'(\varphi;\varphi') = \int_0^T -u^*\Big|_{x=a} \varphi' \, dt$$

$$= \left\langle \nabla_\varphi^{L^2} \mathcal{J}, \varphi' \right\rangle_{L^2} = \int_0^T \nabla_\varphi^{L^2} \mathcal{J} \, \varphi' \, dt$$

$$\Longrightarrow \quad \nabla_\varphi^{L^2} \mathcal{J} = -u^*\Big|_{x=a} \quad \text{on } [0, T]$$

- Determination of the gradient $\nabla_\varphi^{L^2} \mathcal{J}$ requires:
  - solution of the governing system forward in time
  - solution of the adjoint system backwards in time
- When properly defined, the adjoint system conveys information about the *sensitivity* of the solutions of the governing system to perturbations of the data (here, the Neumann boundary condition)

- We will now consider an alternative formulation involving the *Lagrange multiplier* $\lambda : [a, b] \times [0, T]$ (instead of the *reduced* objective functional)

$$\mathcal{L}(\varphi, u, \lambda) = \widetilde{J}(\varphi, u) - \left\langle \frac{\partial u}{\partial t} - \Delta u, \lambda \right\rangle_{L^2(0,T;L^2(a,b))}$$
$$= \frac{1}{2}\int_0^T [u(\varphi)|_b - \bar{u}_b]^2 \, dt - \int_0^T \int_a^b \left(\frac{\partial u}{\partial t} - \Delta u \right) \lambda \, dx \, dt$$

- Solution of the problem $\sup_{\lambda\in\mathcal{X}} \inf_{(x,\varphi)\in\mathcal{X}\times\mathcal{U}} \mathcal{L}(x,\varphi,\lambda)$ requires:

$$\nabla_\lambda \mathcal{L}(\varphi,u,\lambda)=0 \quad \Longrightarrow \quad \begin{cases} \frac{\partial u}{\partial t} - \Delta u = 0, \\ \frac{\partial u}{\partial x}\Big|_{x=a} = \varphi(t), \quad \frac{\partial u}{\partial x}\Big|_{x=b} = 0, \\ u(x,t=0) = u_0(x) \end{cases}$$

$$\nabla_u \mathcal{L}(\varphi,u,\lambda)=0 \quad \Longrightarrow \quad \begin{cases} -\frac{\partial \lambda}{\partial t} - \Delta \lambda = 0 \\ \frac{\partial \lambda}{\partial x}\Big|_{x=a} = 0, \quad \frac{\partial \lambda}{\partial x}\Big|_{x=b} = u(\varphi)|_b - \bar{u}_b \\ \lambda(x,t=T) = 0 \end{cases}$$

$$\nabla_\varphi \mathcal{L}(\varphi,u,\lambda)=0 \quad \Longrightarrow \quad -\lambda\Big|_{x=a} = 0$$

- Thus, the three conditions form a *two-point boundary-value problem in time* for u, λ and φ

- At the optimum, the adjoint variable u^* *coincides* with the Lagrange multiplier λ
- Away from the optimum the adjoint variable u^* can be interpreted in terms of the *sensitivity* of the solutions to the governing system with respect to perturbations of the data
  - the operator defining the adjoint system is determined by the governing equation (it is the Hilbert space adjoint of its linearization)
  - there is some freedom in choosing the data for the adjoint system (terminal & boundary conditions, source term)
  - The Riesz theorem guarantees that this freedom can always be exploited to obtain the required sensitivity
- The action of the Hessian of the objective functional $\mathcal{H}\mathcal{J}(\varphi;\varphi')$ on some perturbation φ' can be determined in a similar way
  - the second-order adjoint is needed

- The L^2 gradients $\nabla_{\varphi}^{L^2}\mathcal{J}$ may not by regular (smooth) enough (they are only square-integrable!)

- We should extract the gradient in the space of smoother functions: the Sobolev space $H^1(0,T)$ endowed with the inner product

$$\begin{aligned}\forall_{p_1,p_2\in H^1(0,T)} \quad \langle p_1, p_2\rangle_{H^1} &= \langle p_1, p_2\rangle_{L^2} + \ell^2 \left\langle \frac{dp_1}{dt}, \frac{dp_2}{dt}\right\rangle_{L^2} \\ &= \int_0^T p_1\, p_2\, dt + \ell^2 \int_0^T \frac{dp_1}{dt}\,\frac{dp_2}{dt}\, dt\end{aligned}$$

- $\ell \in \mathbb{R}$ is a "length-scale" parameter
  - the H^1 inner produce are *equivalent* for $0 < \ell < \infty$

- More precisely, we will assume that $\nabla_{\varphi}^{H^1}\mathcal{J}, \varphi' \in H_0^1(0,T)$ such that

$$\nabla_{\varphi}^{H^1}\mathcal{J}(t) = \varphi'(t) = 0 \quad \text{at } t = 0, T$$

- Invoking again the Riesz representation theorem, we obtain an expression for the Gâteaux differential in terms of the H^1 inner product

$$
\begin{aligned}
\mathcal{J}'(\varphi;\varphi') &= \left\langle \nabla_{\varphi}^{L^2}\mathcal{J}, \varphi' \right\rangle_{L^2} \\
&= \left\langle \nabla_{\varphi}^{H^1}\mathcal{J}, \varphi' \right\rangle_{H^1} \\
&= \int_0^T \nabla_{\varphi}^{H^1}\mathcal{J}\,\varphi'\,dt + \ell^2 \int_0^T \frac{d(\nabla_{\varphi}^{H^1}\mathcal{J})}{dt}\,\frac{d\varphi'}{dt}\,dt
\end{aligned}
$$

- We shall use integration by parts to transform the second term

$$
\begin{aligned}
\left\langle \nabla_{\varphi}^{H^1} \mathcal{J}, \varphi' \right\rangle_{H^1} &= \int_0^T \nabla_{\varphi}^{H^1} \mathcal{J}\, \varphi'\, dt + \ell^2 \int_0^T \frac{d(\nabla_{\varphi}^{H^1} \mathcal{J})}{dt}\, \frac{d\varphi'}{dt}\, dt \\
&= \int_0^T \nabla_{\varphi}^{H^1} \mathcal{J}\, \varphi'\, dt - \ell^2 \int_0^T \frac{d^2(\nabla_{\varphi}^{H^1} \mathcal{J})}{dt^2}\, \varphi'\, dt + \ell^2 \underbrace{\left[\frac{d(\nabla_{\varphi}^{H^1} \mathcal{J})}{dt}\, \varphi' \right]\Bigg|_{t=0}^{T}}_{=0} \\
&= \int_0^T \left[\nabla_{\varphi}^{H^1} \mathcal{J} - \ell^2 \frac{d^2(\nabla_{\varphi}^{H^1} \mathcal{J})}{dt^2} \right] \varphi'\, dt = \int_0^T -u^*\Big|_{x=a} \varphi'\, dt
\end{aligned}
$$

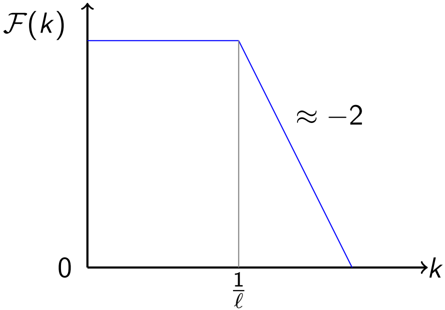
- Since the last relation must hold for any $\varphi' \in H_0^1(0,T) \subset L^2(0,T)$, we obtain

$$
\begin{cases}
\left[\mathrm{Id} - \ell^2 \dfrac{d^2}{dt^2} \right] \nabla_{\varphi}^{H^1} \mathcal{J} = \nabla_{\varphi}^{L^2} \mathcal{J} & \text{on } (0,T) \\
\nabla_{\varphi}^{H^1} \mathcal{J}(0) = \nabla_{\varphi}^{H^1} \mathcal{J}(T) = 0 &
\end{cases}
$$

- The Sobolev gradient $\nabla_{\varphi}^{H^1} \mathcal{J}$ is obtained from the L^2 gradient $\nabla_{\varphi}^{L^2} \mathcal{J}$ by solving an *elliptic boundary-value problem*

- Consider the equation determining the Sobolev gradient $\nabla_{\varphi}^{H^1}\mathcal{J}$ in the Fourier space (for $k = 1, 2, \dots$)

$$\left[1+\ell^2\,k^2\right]\,\left[\widehat{\nabla_{\varphi}^{H^1}\mathcal{J}}\right]_k = \left[\widehat{\nabla_{\varphi}^{L^2}\mathcal{J}}\right]_k \quad \Longrightarrow \quad \left[\widehat{\nabla_{\varphi}^{H^1}\mathcal{J}}\right]_k = \underbrace{\frac{1}{1+\ell^2\,k^2}}_{\mathcal{F}(k)}\,\left[\widehat{\nabla_{\varphi}^{L^2}\mathcal{J}}\right]_k$$



- Extraction of gradients is Sobolev spaces is equivalent to *low-pass filtering* in the frequency space
  - $1/\ell$ is the cut-off frequency

- How to choose an optimal value of ℓ to produce fastest convergence?
  $\Longrightarrow$ open research problem!

- Some results:

  A. Novruzi and B. Protas, "A gradient method in a Hilbert space with an optimized inner product: achieving a Newton-like convergence", (see `arXiv:1803.02414`), 2018.

Conjugate Gradients

- When using the nonlinear conjugate gradients, we need to evaluate the "momentum" term (the Polak-Ribière version)

$$\beta = \frac{\left\langle \nabla_{\varphi}^{H^1} \mathcal{J}(\varphi^{(n)}), \left(\nabla_{\varphi}^{H^1} \mathcal{J}(\varphi^{(n)}) - \nabla_{\varphi}^{H^1} \mathcal{J}(\varphi^{(n-1)}) \right) \right\rangle_{\mathcal{U}}}{\left\langle \nabla_{\varphi}^{H^1} \mathcal{J}(\varphi^{(n-1)}), \nabla_{\varphi}^{H^1} \mathcal{J}(\varphi^{(n-1)}) \right\rangle_{\mathcal{U}}}$$

- Since $H_0^1(0,T) \subset L^2(0,T)$, we have a choice between using
  - the L^2 inner product $\langle \cdot, \cdot \rangle_{L^2}$, or
  - the Sobolev H^1 inner product $\langle \cdot, \cdot \rangle_{H^1}$

- Suppose we wish to impose the a linear constraint on the control variable, e.g., fix its mean value

$$\int_0^T \varphi \, dt = m, \qquad m \in \mathbb{R}$$

- If we impose this condition on the initial guess, i.e., $\int_0^T \varphi_0 \, dt = m$, then we need to ensure that the gradients have *zero mean*

$$\int_0^T \nabla_\varphi \mathcal{J} \, dt = 0$$

- This property defines a *linear subspace*

$$\mathcal{S} = \left\{ f \in L^2(0,T) \ : \ \int_0^T f(t)\, dt = 0 \right\}$$

- Since the gradient need not satisfy the constraint, it must be *projected* on the subspace defined by this constraint

- The projection operator $P_{\mathcal{S}} : L^2 \to \mathcal{S}$

$$P_{\mathcal{S}} \nabla_\varphi \mathcal{J} = \nabla_\varphi^{H^1} \mathcal{J} - \alpha, \qquad \text{where} \quad \alpha = \int_0^T \nabla_\varphi \mathcal{J} \, dt$$

(the projection is realized by subtracting the mean)

- The Sobolev gradient then must be found in $\mathcal{S} \cap H_0^1(0,T)$ using the Riesz theorem with the representer in $\mathcal{S}$

$$\mathcal{J}'(\varphi;\varphi') = \left\langle P_{\mathcal{S}} \nabla_\varphi^{H^1} \mathcal{J}, \varphi' \right\rangle_{H^1} = \left\langle \nabla_\varphi^{H^1} \mathcal{J} - \alpha, \varphi' \right\rangle_{H^1}$$
$$= \left\langle \nabla_\varphi^{L^2} \mathcal{J}, \varphi' \right\rangle_{L^2}$$

- Proceeding as before, we obtain the *projected Sobolev gradient* $P_{\mathcal{S}} \nabla_{\varphi}^{H^1} \mathcal{J}$ as solution of an elliptic boundary-value problem with a global constraint

$$
\begin{cases}
\left[\mathrm{Id} - \ell^2 \dfrac{d^2}{dt^2} \right] \nabla_{\varphi}^{H^1} \mathcal{J} - \alpha = \nabla_{\varphi}^{L^2} \mathcal{J} \quad \text{on } (0, T) \\
\nabla_{\varphi}^{H^1} \mathcal{J}(0) = \nabla_{\varphi}^{H^1} \mathcal{J}(T) = 0 \\
\displaystyle\int_0^T \nabla_{\varphi}^{H^1} \mathcal{J} \, dt = 0
\end{cases}
$$

- The parameter α acts like a "Lagrange multiplier" necessary to accommodate an additional constraint

Algorithm 1 Projected Steepest Descent Line-Search (PSDLS) for PDEs

$\varphi^{(0)} \leftarrow \varphi_0$ (initial guess)
$n \leftarrow 0$
repeat
  solve the governing system with data $\varphi^{(n)}$ forward in time
  solve the corresponding adjoint problem backwards in time
  determine the L^2 gradient $\nabla_\varphi^{L^2}\mathcal{J}$
  determine the projector $P_{\mathcal{S}_{\varphi^{(n)}}}$
  determine the projected Sobolev gradient gradient $P_{\mathcal{S}_{\varphi^{(n)}}}\nabla_\varphi^{H^1}\mathcal{J}$
  determine optimal step size $\tau_n = \operatorname{argmin}_{\tau>0}\mathcal{J}(\varphi^{(n)} - \tau\, P_{\mathcal{S}_{\varphi^{(n)}}}\nabla_\varphi^{H^1}\mathcal{J})$
  update $\varphi^{(n+1)} = \varphi^{(n)} - \tau_n\, P_{\mathcal{S}_{\varphi^{(n)}}}\boldsymbol{\nabla}\mathcal{J}(\varphi^{(n)})$
  $n \leftarrow n+1$
until $\frac{|\mathcal{J}(\varphi^{(n)})-\mathcal{J}(\varphi^{(n-1)})|}{|\mathcal{J}(\varphi^{(n-1)})|} < \varepsilon_f$

Input:

φ_0 — initial guess, ε_τ — tolerance in line search
ε_f — tolerance in the termination condition

Output: an approximation of the minimizer $\widetilde{\varphi}$

- For the purpose of numerical solution, the heat equation is discretized
  - using second-order central/forward finite differences in space
  - using second-order Crank-Nicolson scheme in time

- At each time step we need to solve the following linear system

$$\begin{bmatrix} -3 & 4 & -1 & & \cdots & 0 \\ -\frac{1}{2}h & 1+h & -\frac{1}{2}h & & \cdots & 0 \\ 0 & -\frac{1}{2}h & 1+h & -\frac{1}{2}h & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & -\frac{1}{2}h & 1+h & -\frac{1}{2}h \\ 0 & 0 & 0 & -3 & 4 & -1 \end{bmatrix} \begin{bmatrix} u_{1,n+1} \\ u_{2,n+1} \\ u_{3,n+1} \\ \vdots \\ u_{M-1,n+1} \\ u_{M,n+1} \end{bmatrix} = \begin{bmatrix} 3\,u_{1,n} - u_{2,n} + u_{3,n} + 2\Delta x(\phi_n + \phi_{n+1}) \\ \frac{1}{2}h\,u_{1,n} + (1-h)u_{2,n} + \frac{1}{2}h\,u_{3,n} \\ \frac{1}{2}h\,u_{2,n} + (1-h)u_{3,n} + \frac{1}{2}h\,u_{4,n} \\ \vdots \\ \frac{1}{2}h\,u_{M-2,n} + (1-h)u_{M-1,n} + \frac{1}{2}h\,u_{M,n} \\ 3\,u_{M,n} - u_{M-1,n} + u_{M-2,n} \end{bmatrix}$$

where

- $\{x_1 = a, x_2 = a + \Delta x, \ldots, x_M = b\}$
- $\{t_1 = 0, t_2 = \Delta t, \ldots, t_N = T\}$
- $u_{j,n} = u(x_j, t_n)$
- $h = \frac{\Delta t}{\Delta x^2}$

- The same approach is also used to solve the adjoint problem

- How can we validate the derivation and computation of gradients $\nabla_{\varphi}\mathcal{J}$?
- Compare the Gâteaux differential $\mathcal{J}'(\varphi;\varphi')$
  - approximated using finite differences, and
  - evaluated using the Riesz representation and the gradient $\nabla_{\varphi}\mathcal{J}$
  - $$\kappa(\epsilon) = \frac{\epsilon^{-1}\left[\mathcal{J}(\varphi+\epsilon\varphi') - \mathcal{J}(\varphi)\right]}{\left\langle \nabla_{\varphi}^{L^2}\mathcal{J}, \varphi' \right\rangle_{L^2}}, \quad \forall \varphi, \varphi'$$
- Properties of the quantity $\kappa(\epsilon)$:
  - for intermediate ϵ, $\kappa(\epsilon) \approx 1$ (in fact, $\kappa(\epsilon) \to 1$ as $\Delta x, \Delta t \to 0$)
  - $|\kappa(\epsilon)| \to \infty$ as $\epsilon \to 0$, due to round-off errors
  - $|\kappa(\epsilon)| \to \infty$ as $\epsilon \to \infty$, due to truncation errors in the finite-difference approximation of $\mathcal{J}'(\varphi;\varphi')$