

Adjoint-based sensitivity analysis in internal thermoviscous acoustic flows



Javier Lorente Macías

Department of Engineering
University of Cambridge

First Year Report

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 15,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Javier Lorente Macías
August 2022

Acknowledgements

I would like to express my gratitude to my supervisor, Prof. Matthew Juniper, for offering me this amazing opportunity. I really appreciate the independence that he gives me in my research, but also his ability to redirect me when there appears to be an impasse. I am sure that we will make a great job together in the following years.

I am very thankful to Prof. Patrick Farrell (University of Oxford) for the interesting discussions in functional analysis and preconditioning. His help has been very useful to make our method computationally more efficient. I would also like to thank Dr. Peter Boltryk (Xaar plc.) for sharing his knowledge in the inkjet printing industry.

Many thanks to the people that have also contributed to this report. I would like to thank Peter Benie (University of Cambridge) for his help with computer issues. I also acknowledge the great job of the FEniCSx developers for offering such a powerful open-source finite element solver. A especial thanks goes to Dr. Jørgen Døkken (University of Cambridge) for his help with this software.

I would like to thank my colleagues and friends from ISO-40: Matthew Yoko, Ekrem Ekici, Max Croci, Filip Gökstorp, Dr. Charles Sun, Alexandros Kontogiannis, Alessandro Giannotta and Ushnish Sengupta for the enjoyable moments and intellectual conversations that we have had in this year.

My warmest thanks to Fitzwilliam College for providing a really nice and enriching environment.

Last but not least, thank you to my family for their constant support in difficult times and for encouraging me to pursue my dreams.

Abstract

Drop-on-demand inkjet printing is one of the most widespread applications of microfluidics. Within the printhead, ink continuously flows through multiple microchannels. Each microchannel contains a piezoelectric actuator on the top face and a nozzle on the opposite face. When the actuator pulses, a droplet is forced through the nozzle. Acoustic oscillations then reverberate within the microchannel until they are damped by viscous and thermal dissipation. If a droplet is ejected before the reverberations from the previous droplet have been sufficiently damped, its size is affected by the reverberations, which spoils the image being printed.

In this study, we present a methodology to solve optimal control problems in thermoviscous acoustic flows. First, we derive the governing equations of the thermoviscous acoustic flow inside the printhead microchannels by linearizing the Navier-Stokes equations. We then formulate the adjoint method to obtain the gradient of a cost function with respect to the control parameters, as well as a tool to verify the resulting gradient. Moreover, we study different preconditioning techniques to make the numerical computations faster.

Finally, we apply this methodology to two different dynamical systems - a damped oscillator and a generic one-dimensional viscous acoustic flow. In both cases, the method is able to find the optimal control parameters that minimize the desired quantity.

Contents

List of Figures	xi
List of Tables	xiii
1 Introduction	1
1.1 Drop-on-demand inkjet printing	1
1.2 Adjoint-based sensitivity analysis in flow control	4
1.3 Research objectives	6
1.4 Report structure	7
2 Theoretical formulation	9
2.1 Thermoviscous acoustic equations for inkjet printheads	9
2.1.1 Strong formulation of the equations of motion	9
2.1.2 Weak formulation of the equations of motion	13
2.2 Adjoint-based optimization of dynamical systems	15
2.2.1 Taylor remainder convergence test	18
2.3 Numerical methods	19
3 Applications	33
3.1 Mass-spring-damper system	33
3.2 Viscous-acoustic system	39
3.2.1 One-dimensional geometry	45
3.2.2 Extension to two-dimensional and three-dimensional geometries	48
4 Conclusions and research plan	55
4.1 Key results	55
4.2 Research plan	56
Bibliography	59

List of Figures

1.1	Piezoelectric DOD inkjet printing cycle	2
1.2	Operating regime for stable drop-on-demand inkjet printing	3
2.1	Mapping property of the operator \mathcal{A}	20
2.2	Mapping property of the operator corresponding to the composition of \mathcal{A} and \mathcal{P}	21
2.3	Performance of different preconditioners in the Poisson equation	26
2.4	Performance of different preconditioners in the mixed Poisson equation	31
3.1	Optimal design space for $m = 1, b = 2, m = 4$ and the final time interval T	37
3.2	Evolution of the state vector (a) and the cost function (b) for the uncontrolled damped oscillator	38
3.3	Evolution of the state vector (a) and the cost function (b) for the controlled damped oscillator	38
3.4	Diagram of the duct that contains the acoustic system	40
3.5	Snapshots of the acoustic velocity field (a) and evolution of the cost function (b) for the uncontrolled one-dimensional system	46
3.6	Convergence of the cost function (a) and optimal control (b) for the one-dimensional geometry	47
3.7	Snapshots of the acoustic velocity field (a) and evolution of the cost function (b) for the controlled one-dimensional system	48
3.8	Second order Taylor remainder with quadratic fitting (a) and linear fitting (b)	49
3.9	Snapshots of the acoustic velocity field at $t = \frac{T}{2}$ (a) and $t = T$ (b) for the uncontrolled two-dimensional system	50
3.10	Convergence of the cost function (a) and optimal control (b) for the two-dimensional geometry	50
3.11	Snapshots of the acoustic velocity field at $t = \frac{T}{2}$ (a) and $t = T$ (b) for the controlled two-dimensional system	51

3.12	Snapshots of the acoustic velocity field at $t = \frac{T}{2}$ (a) and $t = T$ (b) for the uncontrolled three-dimensional system	52
3.13	Convergence of the cost function (a) and optimal control (b) for the three-dimensional geometry	52
3.14	Snapshots of the acoustic velocity field at $t = \frac{T}{2}$ (a) and $t = T$ (b) for the controlled three-dimensional system	53
3.15	Comparison of the optimal control (a) and the evolution of the cost function for the three geometries considered	53
4.1	Gantt chart	57

List of Tables

2.1	Time required to obtain the solution of the Poisson equation for different choices of preconditioners and for different number of grid points	26
2.2	Time required to obtain the solution of the mixed Poisson equation for different choices of preconditioners and for different number of grid points .	31

Chapter 1

Introduction

1.1 Drop-on-demand inkjet printing

Drop-on-demand (DOD) inkjet printing is a technique in which ink droplets are generated on demand and jetted toward a moving substrate. An inkjet printhead typically contains several microchannels and nozzles. Ink is flowing continuously through the channels at constant pressure difference (between one side and the other) and at constant pressure (relative to the atmosphere) just above the nozzle. An increase of the pressure must be generated inside the channel to expel droplets. This is usually achieved by either of the following two different mechanisms. In thermal DOD printing, a heater forms a vapour droplet which reduces the effective volume of ink inside the channel, so that a droplet is expelled to return to equilibrium. In piezoelectric DOD printing, an actuator placed on one or more sides of the channel moves a few tens of nanometers, which generates acoustic pressure waves. These oscillations eventually reach the nozzle and a droplet is expelled. The piezoelectric actuators are made of ceramic materials that exhibit the inverse piezoelectric effect (i.e., they deform when a voltage is applied).

Whereas thermal DOD printing is common in household printers, piezoelectric DOD is preferred in applications that require a high levels of precision. Some examples are printing microcircuits (Lee et al., 2011; Tomaszewski and Potencki, 2017), manufacturing of passive electronic components such as capacitors (Chiolerio et al., 2014), biomedicine (Lorber et al., 2013), or advanced manufacturing (Hoath, 2016).

Figure 1.1 illustrates the printing cycle of a piezoelectric DOD inkjet. At the beginning of an ideal cycle, ink is flowing continuously through the channel at a constant pressure difference. This avoids air bubbles becoming trapped in the channel. The piezoelectric actuator placed on one side of the channel bends and generates a wavefront of acoustic waves. These pressure waves reach the nozzle and push a droplet of ink out of the nozzle.

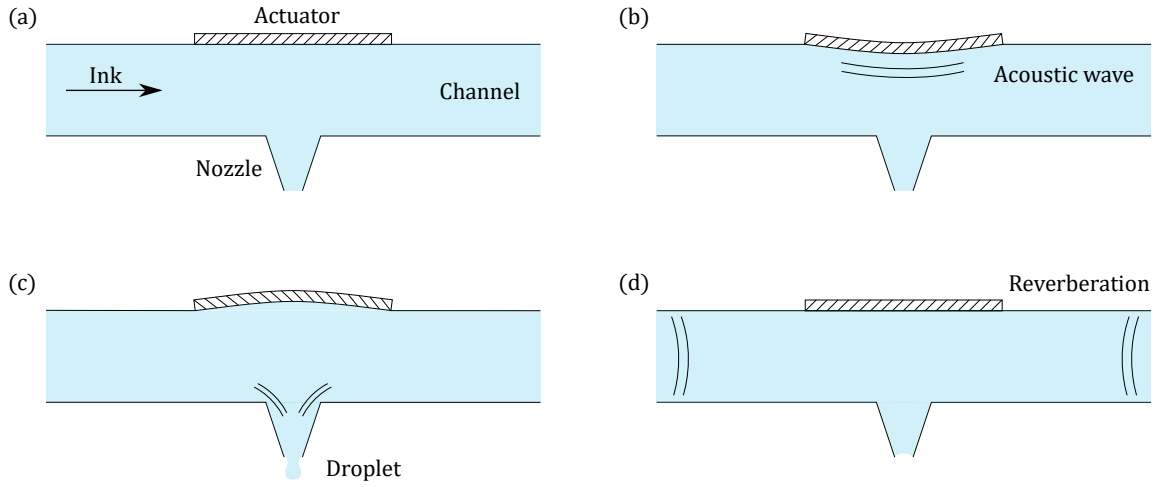


Figure 1.1 Piezoelectric DOD inkjet printing cycle. (a) Ink flows at constant pressure difference through the channel. (b) The actuator bends and generates pressure waves that travel towards the nozzle. (c) The waves reach the nozzle and a droplet is ejected. (d) Acoustic reverberations travel through the channel until they are damped by viscous dissipation.

The acoustic waves are reflected at the channel walls and continue to propagate within the printhead until the momentum is dissipated due to the flow viscosity and thermal diffusion (this latter effect is smaller in comparison with the former).

Manufacturers seek to improve mainly two characteristics of the printing process. The first one is to achieve a high reproducibility of the droplets, that is, to form droplets of same geometry and with uniform trajectories (Gan et al., 2009). This improves the resolution and the quality of the product. The second one is increasing the jetting frequency (Miers and Zhou, 2017), so that the printing time is reduced.

The characteristics of the droplet have a strong impact on the printing process. The most influential properties are ink viscosity ν , surface tension γ , droplet size R and droplet ejection velocity U (Lohse 2022). Two dimensionless numbers can be introduced that account for these properties. The first one is the Reynolds number $Re = \frac{UR}{\nu}$, which characterizes the ratio between inertial and viscous forces. The second one is the Ohnesorge number $Oh = \frac{\mu}{\sqrt{\rho\gamma R}}$, which characterizes the ratio of the viscous forces to inertial and surface tension forces.

The limiting conditions of the inkjet printing process are presented in Figure 1.2 (McKinley, 2011), which shows the operating conditions as a function of the parameter space defined by these dimensionless numbers. For too small Reynolds numbers, the actuation does not have energy enough to achieve a droplet pinch-off at the end of the nozzle. For too large Reynolds numbers, the droplet splashes when it hits the printing surface. For too small Ohnesorge numbers, the surface tension is high and satellite droplets are formed due to Plateau–Rayleigh instability (Plateau, 1873; Rayleigh, 1878). The reason is that, when

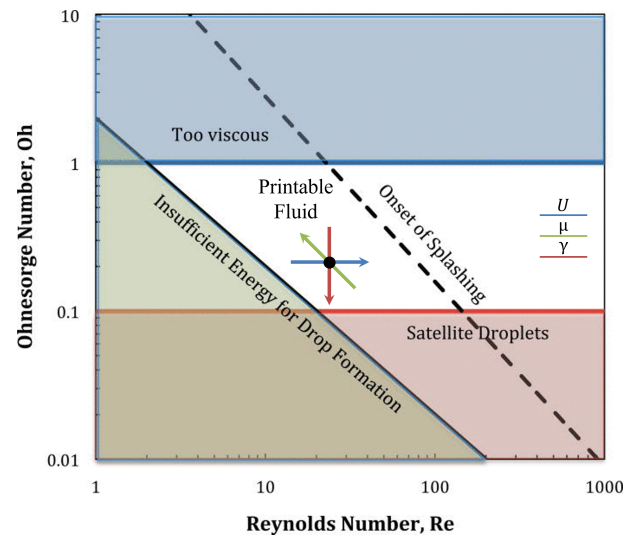


Figure 1.2 Operating regime for stable drop-on-demand inkjet printing. Re is the Reynolds number and Oh is the Ohnesorge number associated to the droplet. U is the droplet velocity, μ is the dynamic viscosity and γ is the surface tension. The black dot represents a particular ink. The arrows show the directions along which the operation point moves as the corresponding property is increased. Figure adapted from the results presented in McKinley (2011).

surface tension is high, the ejected droplets tend to minimize their surface area to reduce the surface tension forces, so that a given volume of ink splits into different droplets of smaller surface. For large Ohnesorge numbers, ink is too viscous to be jetted through pressure buildup. Thus, the inkjet printing process is only possible for intermediate values of the Reynolds and Ohnesorge numbers.

The actuator waveform also plays a fundamental role in the droplet formation and hence in the printing process. The amplitude of the pulse, its length and the time interval between pulses determine the droplet speed, size and frequency of the jetting. Thus, the droplet ejection frequency can be increased by shortening the time interval between pulses. However, if this time interval is very short, the flow is not able to dissipate the residual reverberations before the following droplet is ejected. This state of non-zero acoustic energy in the channel alters the geometry of the droplet, which affects the reproducibility.

In summary, we can observe that there exist two main approaches to improve the reproducibility and increase the jetting frequency. The first one is to optimize the properties of the ink, especially its viscosity or surface tension. A second alternative is by means of flow control. This includes passive control (i.e., without using external energy), or active control (i.e., using an external energy source). Extensive experimental campaigns have been performed to both to find the optimal ink properties and control the acoustic reverberations using the piezoelectric actuator.

1.2 Adjoint-based sensitivity analysis in flow control

The objectives mentioned previously can be formulated as mathematical optimization problems subjected to a certain number of constraints, such as governing equations or restrictions in the parameter space:

$$\begin{aligned}
 & \min_{\mathbf{c}} J(\mathbf{q}, \mathbf{c}) \\
 & \text{subject to:} \\
 & \mathbf{F}(\mathbf{q}, \mathbf{c}) = 0 \\
 & G_i(\mathbf{c}) = g_i, \quad \text{for } i = 1, \dots, m_G \\
 & H_j(\mathbf{c}) \leq h_j, \quad \text{for } i = 1, \dots, m_H
 \end{aligned} \tag{1.1}$$

where $\mathbf{c} \in \mathbb{R}^m$ is the vector of design parameters, J is the objective functional and \mathbf{F} is an operator that defines the governing equations that the state vector $\mathbf{q} \in \mathbb{R}^N$ satisfies.

For problems in which the design space has a high number of dimensions, global optimization methods become computationally expensive. In these cases, gradient-based optimization methods are usually the preferred choice.

Gradient-based optimization problems are reduced to finding the gradient of the cost function with respect to the parameters at a given point. The gradient defines a direction in the design space along which the cost function decreases. The design parameters are updated along that direction and the gradient is then recomputed at the new point. This process of computing the gradient and propagating the control parameters is repeated iteratively until convergence.

There exist essentially two methods to compute the gradient of the cost function: the tangent approach and adjoint approach. In the following analysis, we compare both methods and discuss the applications for which each of them is more suitable.

For simplicity, let us consider that the control constraints are embedded inside the cost function as penalty terms. The gradient of the cost function with respect to the control parameters is given by:

$$\frac{dJ}{d\mathbf{c}} = \frac{\partial J}{\partial \mathbf{q}} \frac{d\mathbf{q}}{d\mathbf{c}} + \frac{\partial J}{\partial \mathbf{c}} \tag{1.2}$$

The variables $\frac{\partial J}{\partial \mathbf{q}}$ and $\frac{\partial J}{\partial \mathbf{c}}$ are typically straightforward to compute. In contrast, the Jacobian of the solution, $\frac{d\mathbf{q}}{d\mathbf{c}}$, is more challenging. In the tangent approach, we differentiate the governing equations to obtain an expression for the Jacobian:

$$\frac{d\mathbf{F}}{d\mathbf{c}} = \frac{\partial \mathbf{F}}{\partial \mathbf{q}} \frac{d\mathbf{q}}{d\mathbf{c}} + \frac{\partial \mathbf{F}}{\partial \mathbf{c}} = 0 \quad (1.3)$$

which can be rearranged as:

$$\frac{\partial \mathbf{F}}{\partial \mathbf{q}} \frac{d\mathbf{q}}{d\mathbf{c}} = - \frac{\partial \mathbf{F}}{\partial \mathbf{c}} \quad (1.4)$$

The expression above is referred to as the tangent equation. The solution of this system gives the value of the Jacobian of the solution ¹. This variable is then introduced in Equation (1.2) to compute the gradient.

The adjoint approach avoids the direct computation of the Jacobian of the solution, for reasons that will be explained later. Assuming that $\frac{\partial \mathbf{F}}{\partial \mathbf{q}}$ is invertible, the Jacobian can be expressed as:

$$\frac{d\mathbf{q}}{d\mathbf{c}} = - \left(\frac{\partial \mathbf{F}}{\partial \mathbf{q}} \right)^{-1} \frac{\partial \mathbf{F}}{\partial \mathbf{c}} \quad (1.5)$$

Substituting this result into Equation (1.2) and taking the Hermitian transpose of the resulting expression yields:

$$\left(\frac{dJ}{d\mathbf{c}} \right)^* = - \left(\frac{\partial J}{\partial \mathbf{q}} \right)^* \left(\frac{\partial \mathbf{F}}{\partial \mathbf{q}} \right)^{-*} \left(\frac{\partial \mathbf{F}}{\partial \mathbf{c}} \right)^* + \left(\frac{\partial J}{\partial \mathbf{c}} \right)^* \quad (1.6)$$

We now define the adjoint variable, $\boldsymbol{\lambda}$, as follows:

$$\boldsymbol{\lambda} = \left(\frac{\partial \mathbf{F}}{\partial \mathbf{q}} \right)^{-*} \left(\frac{\partial \mathbf{F}}{\partial \mathbf{c}} \right)^* \quad (1.7)$$

which can be rearranged as:

$$\left(\frac{\partial \mathbf{F}}{\partial \mathbf{q}} \right)^* \boldsymbol{\lambda} = \left(\frac{\partial \mathbf{F}}{\partial \mathbf{c}} \right)^* \quad (1.8)$$

Equation (1.8) is the adjoint equation. Once this equation is solved, the gradient of the cost function can be computed as:

$$\frac{dJ}{d\mathbf{c}} = - \frac{\partial J}{\partial \mathbf{q}} \boldsymbol{\lambda}^* + \frac{\partial J}{\partial \mathbf{c}} \quad (1.9)$$

Thus, in both approaches, we first solve an auxiliary set of equations and then compute the gradient. However, the problem for which each method is efficient is radically different. When solving the tangent equation, the cost function J does not appear in the formulation.

¹In the literature, the Jacobian of the solution is also called as the tangent variable

This means that the value of the tangent variable is independent of the number of objective functions that one wants to minimize. In contrast, the tangent equation is actually a system of m equations. This makes the tangent approach attractive in problems with a few design parameters and many functionals to minimize. In the adjoint approach, the control vector \mathbf{c} does not appear in the adjoint equation, so that the adjoint vector is independent of the number of parameters considered. However, one needs to solve as many adjoint equations as the number of cost functions that need to be minimized. Therefore, the adjoint approach is useful for problems with many control parameters and a few cost functions to minimize.

In many engineering applications, we are interested in minimizing a single functional. Moreover, the parametrization of the control variable usually leads to a large number of parameters. Therefore, the adjoint method becomes more convenient in these cases and hence the preferred method.

The adjoint-based sensitivity analysis has long been considered as an efficient method to solve PDE-constrained optimization problems when the number of design parameters is large (Mitter and Lions, 1971). In fluid dynamics, Pironneau (1974) first used the adjoint method for design. Since then, one of the main applications of adjoint methods has been aerodynamic design. Jameson and other co-authors developed the adjoint approach for potential flow, the Euler equations and the Navier–Stokes equations (Jameson, 1988; Jameson, a; Jameson, b; Jameson and Reuther, 1994; Reuther et al., 1996; Jameson et al., 1998).

Adjoint methods have also been considered as design tools for thermoacoustic problems (Magri, 2019). Some examples are control of jet noise (Kim et al., 2014b), shape optimization to maximize the acoustic damping of a Helmholtz resonator (Caeiro et al., 2017), or shape optimization to stabilize a thermoacoustic annular configuration. Recently, the adjoint method has been also applied to find the optimal geometry of two-dimensional inkjet printing microchannels that maximizes the decay rate of the acoustic waves (Kungurtsev and Juniper, 2019).

1.3 Research objectives

The first objective of this project is to develop a three-dimensional generic model of the thermoviscous acoustic flow inside an inkjet printhead. We will then derive the corresponding adjoint model. As a first optimization problem, we will consider the piezoelectric actuator waveform as the control variable and we will find the optimal waveform that is able to minimize the acoustic energy in the minimum time interval.

For this process, it is essential to solve the numerical problem as efficiently as possible. The number of discretizing elements is large and the problem is unsteady. Moreover, gradient-

based optimization algorithms require evaluation of the cost function and its gradient multiple times. Thus, we will explore preconditioning and parallelization techniques. The results in this three-dimensional geometry will be compared with previous research in two-dimensional geometries. Once the results are verified, we will extend the model to a real inkjet printhead geometry of the industrial collaborator *Xaar*, and compare results with experiments. We may also consider different control parameters, such as the shape of the channel or nozzle.

As for every physics-based model, we will make several assumptions in the parameters of the flow inside the printhead. Therefore, as a final objective, we aim to make our model quantitatively more accurate by assimilating data from experiments.

1.4 Report structure

In Chapter 2, the governing equations of the oscillating flow inside an inkjet printhead is mathematically derived from the Navier-Stokes equations. The theory of adjoint-based sensitivity analysis applied to the optimal control of dynamical systems is also reviewed. Moreover, different numerical methods to speed up the computations are investigated in this chapter.

In Chapter 3, the mathematical framework presented in Chapter 2 is applied to two canonical problems in acoustics. The first one is a damped oscillator, for which the optimization problem has analytical solution. The second application is one-dimensional viscous-acoustic system. Chapter 4 closes this report by summarising the key results and presenting a research plan to achieve the project objectives.

Chapter 2

Theoretical formulation

This chapter contains the mathematical framework required to minimize the acoustic energy in the inkjet printhead. First, we derive the governing equations that model the propagation of acoustic waves inside the channels. The formulation starts from the compressible Navier-Stokes equations. These equations are then simplified into a set of linear equations using a small perturbations analysis. Second, we introduce the adjoint method to compute sensitivities in problems constrained by dynamical systems. The chapter concludes with a section focused on numerical methods to solve linear matrix systems efficiently. The poor behaviour of iterative solvers to solve matrix systems that arise from the discretization of PDEs and techniques to solve this issue are discussed. These considerations are applied to two canonical problems - the standard Poisson equation and the mixed formulation of the Poisson equation.

2.1 Thermoviscous acoustic equations for inkjet printheads

2.1.1 Strong formulation of the equations of motion

The motion of a compressible flow with viscosity and heat conductivity is governed by the Navier-Stokes equations, which in conservative form are:

$$\frac{\partial \mathbf{q}}{\partial t} + \nabla_j (\mathbf{f}_j^c(\mathbf{q}) + \mathbf{f}_j^v(\mathbf{q}, \nabla \mathbf{q})) = 0 \quad \text{in } \Omega \quad (2.1)$$

where \mathbf{q} is the state vector of conservative variables, and \mathbf{f}^c and \mathbf{f}^v are the convective and viscous components of the flux. These vectors are defined as

$$\mathbf{q} = \begin{bmatrix} \rho \\ \rho u_i \\ \rho E \end{bmatrix}, \quad \mathbf{f}^c = \begin{bmatrix} \rho u_k \\ \rho u_i u_j + p \delta_{ik} \\ u_j (\rho E + p) \end{bmatrix}, \quad \mathbf{f}^v = \begin{bmatrix} 0 \\ \tau_{ji} \\ u_i \tau_{ji} + \kappa_{th} \Delta_j T \end{bmatrix} \quad (2.2)$$

The flow variables ρ , \mathbf{u} , p and T are, respectively, flow density, velocity, pressure and temperature. E is the total energy and is the sum of the kinetic energy and the internal energy e :

$$\rho E = \rho e + \frac{1}{2} \rho \mathbf{u} \cdot \mathbf{u} \quad (2.3)$$

τ is the viscous stress tensor, which is expressed as:

$$\tau_{ij} = \mu_v \left(\nabla \mathbf{u} + (\nabla \mathbf{u})^T - \frac{2}{3} \delta_{ij} \nabla_k u_k \right) \quad (2.4)$$

where μ_v is the dynamic viscosity.

We also introduce an equation of state that relates the flow density, pressure and temperature, and an equation that relates the internal energy, pressure and temperature:

$$\rho = \rho(p, T) \quad (2.5a)$$

$$e = e(p, T) \quad (2.5b)$$

In inkjet printheads, the flow is governed to first order by two phenomena (Kungurtsev and Juniper, 2019). The first one is a steady flow with rigid boundaries whose characteristic speed is $\bar{U} \sim 1$ m/s. The second one is a periodic acoustic flow where the displacement amplitude at the boundary is $\Delta \sim 0.01 \mu\text{m}$ at high oscillation frequency $f \sim 100$ kHz. Thus, the characteristic speed of the acoustic flow is $\tilde{U} = \omega \Delta \sim 10^{-3}$ m/s. The sound speed in typical inks ranges from 800 to 1600 ms (Beltman, 1998; Dijksman, 2019; Kim et al., 2014a). In this study, we consider $c_0 = 1000$ m/s. Therefore, the Mach number associated with the steady flow, ϵ_s , and the acoustic flow, ϵ_{ac} , are:

$$\epsilon_s = \frac{\bar{U}}{c_0} \sim 10^{-3} \ll 1 \quad (2.6a)$$

$$\epsilon_{ac} = \frac{\tilde{U}}{c_0} \sim 10^{-6} \ll 1 \quad (2.6b)$$

As previously mentioned, these small Mach numbers allow us to consider a two-parameter low Mach number asymptotic expansion of the equations of motion. However, we can simplify the problem further by making the following observation. The characteristic time of the steady flow is $t_s \sim \frac{L}{U}$ whereas the characteristic time of the oscillating flow is $t_{ac} \sim \frac{L}{c_0}$. As a consequence we get the following ratio of characteristic times

$$\frac{t_s}{t_{ac}} \sim \frac{c_0}{U} \sim 10^3 \gg 1 \quad (2.7)$$

which implies that each phenomenon occurs at a completely different time scale. This allows us to decouple both flows and analyse them separately. In this study, we want to minimize the acoustic energy of the fluid, so we focus on the acoustic problem.

We then expand the flow variables in terms of the Mach number of the acoustic flow:

$$\mathbf{u}(\mathbf{x}, t) = \bar{\mathbf{u}}(\mathbf{x}) + \varepsilon_{ac} \mathbf{u}'(\mathbf{x}, t) + \mathcal{O}(\varepsilon_{ac}^2) \quad (2.8a)$$

$$p(\mathbf{x}, t) = \bar{p}(\mathbf{x}) + \varepsilon_{ac} p'(\mathbf{x}, t) + \mathcal{O}(\varepsilon_{ac}^2) \quad (2.8b)$$

$$T(\mathbf{x}, t) = \bar{T}(\mathbf{x}) + \varepsilon_{ac} T'(\mathbf{x}, t) + \mathcal{O}(\varepsilon_{ac}^2) \quad (2.8c)$$

where $(\bar{\cdot})$ denotes mean flow variable and (\cdot') denotes acoustic perturbation variable.

We also define the following non-dimensional acoustic variables:

$$\frac{u'_i}{c_0} \rightarrow \tilde{u}_i, \quad \frac{p'}{\bar{p}} \rightarrow \tilde{p}, \quad \frac{\alpha_p T'}{\bar{T}} \rightarrow \tilde{T} \quad (2.9a)$$

$$\frac{x}{L} \rightarrow \tilde{x}, \quad \frac{t}{\frac{L}{c_0}} \rightarrow \tilde{t} \quad (2.9b)$$

where α_p is the volumetric coefficient of thermal expansion.

We now introduce expressions (2.8) into Equation (2.1) and collect terms of order $\mathcal{O}(\varepsilon_{ac})$. After non-dimensionalizing, we obtain the set of thermoviscous acoustic equations:

$$\frac{\partial}{\partial \tilde{t}} (\gamma_{th} \tilde{p} - \tilde{T}) + \nabla \cdot \tilde{\mathbf{u}} = 0 \quad (2.10a)$$

$$\frac{\partial \tilde{\mathbf{u}}}{\partial \tilde{t}} = -\nabla \tilde{p} + \frac{1}{\tilde{Re}} \nabla \cdot \tilde{\boldsymbol{\tau}} \quad (2.10b)$$

$$\frac{\partial}{\partial \tilde{t}} \left(\frac{\tilde{T}}{\gamma_{th} - 1} - \tilde{p} \right) = \frac{1}{(\gamma_{th} - 1) \tilde{Pe}} \Delta \tilde{T} \quad (2.10c)$$

where $\tilde{Re} = \frac{\bar{\rho}c_0L}{\mu}$ and $\tilde{Pe} = \frac{\bar{\rho}c_0c_pL}{\kappa_{th}}$ are, respectively, the Reynolds number and the Péclet number based on the sound speed.

The system of equations (2.10) can be expressed in compact form as

$$A_c \frac{\partial \tilde{\mathbf{q}}}{\partial t_{ac}} + B_c \tilde{\mathbf{q}} = 0 \quad (2.11)$$

where

$$\tilde{\mathbf{q}} = \begin{bmatrix} \tilde{\mathbf{u}} \\ \tilde{p} \\ \tilde{T} \end{bmatrix}, \quad A_c = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \gamma_{th} & -1 \\ 0 & -1 & \frac{1}{\gamma_{th}-1} \end{bmatrix}, \quad B_c = \begin{bmatrix} -\frac{1}{Re} \nabla_j \tau_{ij} & \nabla_i & 0 \\ \nabla_i & 0 & 0 \\ 0 & 0 & -\frac{\Delta}{(\gamma_{th}-1)Pe} \end{bmatrix} \quad (2.12)$$

In order to complete the definition of the problem, we need to provide initial and boundary conditions. The initial state vector is given by:

$$\mathbf{q}(\mathbf{x}, 0) = \mathbf{q}_0(\mathbf{x}) \quad \text{in } \Omega \quad (2.13)$$

At the boundaries, either a velocity U or a force f is prescribed. We define the no-slip boundary Γ_{nsl} where $U = 0$, inlet boundary Γ_{in} where $U \neq 0$, stress-free boundary Γ_{free} where $\mathbf{f} = 0$, and forcing boundary Γ_{force} where $\mathbf{f} \neq 0$:

$$\mathbf{u} = U \quad \text{on } \Gamma_{in}, \quad \mathbf{u} = 0 \quad \text{on } \Gamma_{nsl} \quad (2.14a)$$

$$\sigma_{ij}n_j = f_i \quad \text{on } \Gamma_{force}, \quad \sigma_{ij}n_j = 0 \quad \text{on } \Gamma_{free} \quad (2.14b)$$

where σ is the stress tensor defined as:

$$\sigma_{ij} = -p\delta_{ij} + \tau_{ij} \quad (2.15)$$

Similarly, we define the boundary conditions for the energy equation. At the boundaries, either a temperature T or a heat flux $\frac{\partial T}{\partial \mathbf{n}}$ is prescribed. We define the isothermal boundary Γ_{iso} where $T = T_w$ and the adiabatic boundary Γ_a where $\frac{\partial T}{\partial \mathbf{n}} = 0$:

$$T = T_w \quad \text{on } \Gamma_{iso}, \quad \frac{\partial T}{\partial \mathbf{n}} = 0 \quad \text{on } \Gamma_a \quad (2.16)$$

It is possible for a boundary condition to represent a particular physical phenomenon, in which case the boundary condition is a function of the acoustic state:

$$\mathcal{B}(\tilde{\mathbf{q}}) = g(\tilde{\mathbf{q}}) \quad \text{on } \Gamma_{NL} \quad (2.17)$$

where \mathcal{B} is an operator over the acoustic variable and g is a function of the acoustic variable. Note that this makes the problem nonlinear.

2.1.2 Weak formulation of the equations of motion

We now introduce the weak form of the system of equations (2.10). Let us first define the volume inner product (\cdot, \cdot) and surface inner product $\{\cdot, \cdot\}$:

$$(u, v) = \int_{\Omega} uv d\Omega \quad (2.18a)$$

$$\{u, v\} = \int_{\partial\Omega} uv d\Omega \quad (2.18b)$$

We consider a scalar-valued test function $q \in \hat{\mathcal{Q}}$, a vector-valued test function $\mathbf{v} \in \hat{\mathcal{V}}$ and another scalar-valued test function $w \in \hat{\mathcal{W}}$. We define the vector of test functions as $\mathbf{p} = [q, \mathbf{v}, w]^T \in \hat{\mathcal{Y}}$, where $\hat{\mathcal{Y}}$ is the mixed function space $\hat{\mathcal{Y}} = \hat{\mathcal{Q}} \times \hat{\mathcal{V}} \times \hat{\mathcal{W}}$.

We multiply Equation (2.10a) by the test function q , Equation (2.10b) by the test function \mathbf{v} , Equation (2.10c) by the test function w , and integrate over the domain Ω :

$$\frac{d}{dt_{ac}} \int_{\Omega} (\gamma_{th} \tilde{p} - \tilde{T}) q d\Omega + \int_{\Omega} \nabla \cdot \tilde{\mathbf{u}} q d\Omega = 0 \quad (2.19a)$$

$$\frac{d}{dt_{ac}} \int_{\Omega} \tilde{\mathbf{u}} \cdot \mathbf{v} d\Omega = - \int_{\Omega} \nabla \tilde{p} \cdot \mathbf{v} d\Omega + \frac{1}{\tilde{Re}} \int_{\Omega} (\nabla \cdot \tilde{\boldsymbol{\tau}}) \cdot \mathbf{v} d\Omega \quad (2.19b)$$

$$\frac{d}{dt_{ac}} \int_{\Omega} \left(\frac{\tilde{T}}{\gamma_{th} - 1} - \tilde{p} \right) w d\Omega = \frac{1}{(\gamma_{th} - 1) \tilde{Pe}} \int_{\Omega} \Delta \tilde{T} w d\Omega \quad (2.19c)$$

We then integrate by parts the terms containing second order derivatives in space and set the test functions to zero in the boundaries where the corresponding trial function is known (or set the dot product of the test function and the normal vector to zero if the flux across the boundary is known). We also integrate the pressure term by parts for reasons that will become clear later. After grouping terms we can write the weak form as:

find $\mathbf{q} \in \mathcal{Y}$ such that:

$$a\left(\frac{\partial \mathbf{q}}{\partial t}, \mathbf{p}\right) + b(\mathbf{q}, \mathbf{p}) + b_\Gamma(\mathbf{q}, \mathbf{p}) = 0 \quad (2.20)$$

for all $\mathbf{p} \in \hat{\mathcal{Y}}$

where

$$a(\mathbf{q}, \mathbf{p}) = \langle \gamma_{th} \tilde{p} - \tilde{T}, q \rangle_\Omega + \langle \tilde{\mathbf{u}}, \mathbf{v} \rangle_\Omega + \left\langle \frac{\tilde{T}}{\gamma_{th} - 1} - \tilde{p}, w \right\rangle_\Omega \quad (2.21a)$$

$$b(\mathbf{q}, \mathbf{p}) = \langle \nabla_i \tilde{u}_i, q \rangle_\Omega - \langle \tilde{p}, \nabla_i v_i \rangle_\Omega + \left\langle \frac{1}{Re} \tilde{\tau}_{ij}, \nabla_j v_i \right\rangle_\Omega + \quad (2.21b)$$

$$+ \left\langle \frac{1}{(\gamma_{th} - 1) Pe} \nabla_j \tilde{T}, \nabla_j q \right\rangle_\Omega \quad (2.21c)$$

$$b_\Gamma(\mathbf{q}, \mathbf{p}) = - \{ \tilde{\sigma}_{ij} n_j, v_i \}_{\partial\Omega} - \left\{ \frac{1}{(\gamma_{th} - 1) Pe} \frac{\partial \tilde{T}}{\partial \mathbf{n}}, q \right\}_{\partial\Omega} \quad (2.21d)$$

By integrating by parts the pressure term, we can define the full stress tensor at the boundary as in Equation (2.21d). A suitable choice of the function spaces associated with the trial and test functions is:

$$\mathcal{Q} = \{L^2(\Omega)\} \quad (2.22a)$$

$$\hat{\mathcal{Q}} = \{L^2(\Omega)\} \quad (2.22b)$$

$$\mathcal{V} = \left\{ v = (H^1(\Omega))^d : v = u_D \text{ on } \Gamma_D \right\} \quad (2.22c)$$

$$\hat{\mathcal{V}} = \left\{ v = (H^1(\Omega))^d : v = 0 \text{ on } \Gamma_D \right\} \quad (2.22d)$$

$$\mathcal{W} = \{H^1(\Omega) : w = T_D \text{ on } \Gamma_D\} \quad (2.22e)$$

$$\hat{\mathcal{W}} = \{H^1(\Omega) : w = 0 \text{ on } \Gamma_D\} \quad (2.22f)$$

where the superscript d indicates topological dimension.

Finally, we can define the total thermoacoustic energy of the flow as follows (Chu 1965):

$$\mathcal{E}_{ac} = \frac{1}{2} ((\tilde{\mathbf{u}}, \tilde{\mathbf{u}}) + (\tilde{p}, \tilde{p}) + (\tilde{T}, \tilde{s})) \quad (2.23)$$

where the first term is the kinetic energy, the second term is the potential energy associated with compression and the third term is the potential energy associated with entropy fluctuations.

2.2 Adjoint-based optimization of dynamical systems

In this section, we lay out the theoretical foundations of the adjoint method to find the optimal control that minimizes a cost function in dynamical systems. We start from the formulation of a general constrained optimization problem subject to the governing equations of the dynamical system. We then transform the problem into an unconstrained optimization problem by defining its Lagrangian. Finally, we find the conditions that the Lagrange multipliers must satisfy so that the sensitivity of the Lagrangian with respect to the control parameters is equivalent to the sensitivity of the cost function that we want to minimize.

Let us consider that the governing equations of a dynamical system can be expressed in its general form as follows:

$$\frac{d\mathbf{q}}{dt} = \mathbf{f}(\mathbf{q}, \mathbf{c}) \quad (2.24a)$$

$$\mathbf{q}(0) = \mathbf{q}_0(\mathbf{c}) \quad (2.24b)$$

where \mathbf{q} is the state vector and \mathbf{c} is the vector of design parameters.

We assume that the solution of the Problem (2.24) is available either analytically or numerically. Thus, given the solution \mathbf{q} , the objective is to minimize a determined cost functional. In most engineering applications, a general cost function may be defined as follows:

$$J = J_F(\mathbf{q}(T)) + \int_0^T j(\mathbf{q}(t))dt + J_c(\mathbf{c}) \quad (2.25)$$

where the first term corresponds to a quantity defined at $t = T$; the second term is a time averaged function, and the third term is a function that depends on the control parameters. This last term is usually added to regularize the problem.

Thus, we can formulate the following constrained optimization problem:

$$\begin{aligned} \min_{\mathbf{c}} J &= J(\mathbf{q}(T), \mathbf{q}, \mathbf{c}) \\ \text{subject to:} \\ \frac{d\mathbf{q}}{dt} &= \mathbf{f}(\mathbf{q}, \mathbf{c}) \\ \mathbf{q}(0) &= \mathbf{q}_0(\mathbf{c}) \end{aligned} \quad (2.26)$$

We define the Lagrangian of Problem (2.26) as:

$$\mathcal{L} = J + \int_0^T \mathbf{q}^{\dagger T} \left(\frac{d\mathbf{q}}{dt} - \mathbf{f} \right) dt + \mathbf{q}_0^{\dagger T} (\mathbf{q}(0) - \mathbf{q}_0) \quad (2.27)$$

where \mathbf{q}^{\dagger} is the Lagrange multiplier associated with the state vector \mathbf{q} (i.e., the adjoint state vector), and \mathbf{q}_0^{\dagger} is the Lagrange multiplier associated with the initial condition $\mathbf{q}(0)$ (i.e., the adjoint initial condition). The time integral is taken because the integrand must be zero over all time.

We have defined an unconstrained optimization problem. The function to minimize now is the Lagrangian instead of the original cost function. Let us then define the sensitivity of the Lagrangian with respect to its variables:

$$\begin{aligned} \delta \mathcal{L} = & \frac{\partial J_F}{\partial \mathbf{q}(T)} \delta \mathbf{q}(T) + \int_0^T \frac{\partial j}{\partial \mathbf{q}} \delta \mathbf{q} dt + \frac{\partial J_c}{\partial \mathbf{c}} \delta \mathbf{c} + \\ & + \int_0^T \mathbf{q}^{\dagger T} \left(\frac{d\delta \mathbf{q}}{dt} - \frac{\partial \mathbf{f}}{\partial \mathbf{q}} \delta \mathbf{q} - \frac{\partial \mathbf{f}}{\partial \mathbf{c}} \delta \mathbf{c} \right) dt + \mathbf{q}_0^{\dagger T} \left(\delta \mathbf{q}(0) - \frac{\partial \mathbf{q}_0}{\partial \mathbf{c}} \delta \mathbf{c} \right) \end{aligned} \quad (2.28)$$

After integration by parts of the term containing a time derivative of $\delta \mathbf{q}$ we obtain:

$$\begin{aligned} \delta \mathcal{L} = & \frac{\partial J_F}{\partial \mathbf{q}(T)} \delta \mathbf{q}(T) + \int_0^T \frac{\partial j}{\partial \mathbf{q}} \delta \mathbf{q} dt + \frac{\partial J_c}{\partial \mathbf{c}} \delta \mathbf{c} + \mathbf{q}^{\dagger T} \delta \mathbf{q} \Big|_0^T - \\ & - \int_0^T \left(\delta \mathbf{q} \frac{d\mathbf{q}^{\dagger T}}{dt} + \mathbf{q}^{\dagger T} \left(\frac{\partial \mathbf{f}}{\partial \mathbf{q}} \delta \mathbf{q} + \frac{\partial \mathbf{f}}{\partial \mathbf{c}} \delta \mathbf{c} \right) \right) dt + \mathbf{q}_0^{\dagger T} \left(\delta \mathbf{q}(0) - \frac{\partial \mathbf{q}_0}{\partial \mathbf{c}} \delta \mathbf{c} \right) \end{aligned} \quad (2.29)$$

Equation (2.27) is satisfied for any value of the Lagrange multipliers and hence Equation (2.29) is as well. We set them such that the terms in $\delta \mathbf{q}(T)$, $\delta \mathbf{q}$ and $\delta \mathbf{q}(0)$ are zero. This way, the sensitivity of the Lagrangian with respect to the control parameters and the sensitivity of the cost function are the same. Thus, gathering terms in $\delta \mathbf{q}(T)$, the adjoint state vector at $t = T$ must satisfy:

$$\mathbf{q}^{\dagger}(T) = - \left(\frac{\partial J_F}{\partial \mathbf{q}(T)} \right)^T \quad (2.30)$$

and gathering terms in $\delta \mathbf{q}$, the time evolution of the adjoint state vector obeys the following equation:

$$\frac{d\mathbf{q}^{\dagger}}{dt} + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{q}} \right)^T \mathbf{q}^{\dagger} = \left(\frac{\partial j}{\partial \mathbf{q}} \right)^T \quad (2.31)$$

The problem defined by Equation (2.31) and (2.30) is often referred to as the adjoint problem. Note that, whereas the governing equations are defined as an initial value problem, the adjoint equation is a terminal value problem. Therefore, the adjoint problem is well-posed if we integrate it backwards in time.

Finally, gathering terms in $\delta q(0)$ gives the following expression for the adjoint initial condition:

$$\mathbf{q}_0^\dagger = \mathbf{q}^\dagger(0) \quad (2.32)$$

which is the value of the adjoint state vector at $t = 0$.

From this formulation, it can be observed that the adjoint variables have the same properties as its corresponding direct variable, that is, if the state vector has length N and depends on time, the adjoint state vector will have the same size and will be also time dependent.

For these values of the Lagrange multipliers, the sensitivity of the Lagrangian and the original cost function are the same, and is given by:

$$\delta \mathcal{L} = \left(\frac{\partial J_{\mathbf{c}}}{\partial \mathbf{c}} - \int_0^T \mathbf{q}^{\dagger T} \frac{\partial \mathbf{f}}{\partial \mathbf{c}} dt - \mathbf{q}_0^{\dagger T} \frac{\partial \mathbf{q}_0}{\partial \mathbf{c}} \right) \delta \mathbf{c} \quad (2.33)$$

The following lines enumerate the steps that we have to follow to find the optimal control vector \mathbf{c} that minimizes the functional defined in Equation (2.25):

1. Find the solution of the direct problem, \mathbf{q} , integrating (2.24) in time.
2. Find the terminal value of the multiplier \mathbf{q}^\dagger , using Equation (2.30).
3. Integrate Equation (2.31) backwards in time to get the time dependent multiplier $\mathbf{q}^\dagger(t)$.
4. Obtain the value of the multiplier \mathbf{q}_0^\dagger using Equation (2.32).
5. Compute the sensitivity of the cost function with respect to the control parameters using Equation (2.33), which contains a combination of the direct and adjoint variables.

In certain applications, it may be useful to change the definition of the time domain in the adjoint problem. One can define the adjoint time t^\dagger as $t^\dagger = T - t$. Thus, the adjoint equation can be rewritten as:

$$\frac{d\mathbf{q}^\dagger}{dt^\dagger} = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{q}} \right)^T \mathbf{q}^\dagger - \left(\frac{\partial j}{\partial \mathbf{q}} \right)^T \quad (2.34)$$

and the terminal condition becomes an initial condition:

$$\mathbf{q}^\dagger(0) = - \left(\frac{\partial J_F}{\partial \mathbf{q}(T)} \right)^T \quad (2.35)$$

This allows us to integrate the adjoint problem forwards in time in the same way as the direct problem, so that we can use the same time discretization algorithm.

2.2.1 Taylor remainder convergence test

A fundamental step of the design of an algorithm is the verification of its outputs. We have previously detailed a method for computing gradients. In this section, we explain a tool to check if the gradient is correct.

In the simple case, one can obtain the analytical expression of the gradient of the cost function. However, in the majority of fluid dynamics applications, an analytical expression is not available or it is not straightforward to obtain. The second-order Taylor remainder convergence test constitutes an efficient tool to check if the gradient is accurate up to first order.

Let us consider an objective function $J(\mathbf{c})$ where $\mathbf{c} \in \mathbb{R}^m$. We can approximate this function evaluated about \mathbf{c}_0 by means of Taylor series:

$$J(\mathbf{c}) = J(\mathbf{c}_0) + \nabla J|_{\mathbf{c}=\mathbf{c}_0} (\mathbf{c} - \mathbf{c}_0) + \frac{1}{2} (\mathbf{c} - \mathbf{c}_0)^T H|_{\mathbf{c}=\mathbf{c}_0} (\mathbf{c} - \mathbf{c}_0) + \mathcal{O} \left((\mathbf{c} - \mathbf{c}_0)^3 \right) \quad (2.36)$$

where $H = \nabla^2 J$ is the Hessian matrix.

Let $\delta \mathbf{c}$ be a perturbation to \mathbf{c} around \mathbf{c}_0 and let h be the size of this perturbation. Then, the previous expression can be rewritten as:

$$J(\mathbf{c}_0 + h\delta \mathbf{c}) = J(\mathbf{c}_0) + \nabla J|_{\mathbf{c}=\mathbf{c}_0} (h\delta \mathbf{c}) + \frac{1}{2} (h\delta \mathbf{c})^T H|_{\mathbf{c}=\mathbf{c}_0} (h\delta \mathbf{c}) + \mathcal{O} \left((h\delta \mathbf{c})^3 \right) \quad (2.37)$$

The Taylor test is based on the following observation:

$$\left| J(\mathbf{c}_0 + h\delta \mathbf{c}) - J(\mathbf{c}_0) - h \nabla J|_{\mathbf{c}=\mathbf{c}_0} \cdot \delta \mathbf{c} \right| \rightarrow 0 \quad \text{at } \mathcal{O}(h^2) \quad (2.38)$$

The quantity defined in Equation (2.38) is referred to as second order Taylor remainder. Therefore, we can compute this quantity for some choice of $\delta \mathbf{c}$ and h . We then halve h and check if the result decreases by a factor of 4 or higher. If this occurs, the gradient passes the test and we can ensure that it is accurate up to first order.

2.3 Numerical methods

The system of equations described in the previous sections are solved numerically. After the discretization of the problem, we obtain a linear matrix system of the form:

$$A\mathbf{x} = \mathbf{b} \quad (2.39)$$

where $\mathbf{x} \in \mathbb{R}^N$ is the solution vector, $\mathbf{b} \in \mathbb{R}^N$ is the right-hand side vector and $A \in \mathbb{R}^{N \times N}$ is the system matrix.

The solution of the linear problem defined by Equation (2.39) requires an inversion of the matrix A . Depending on the size and form of the matrix A , there exist certain strategies to perform this operation efficiently. Discretization of PDEs with the finite difference method or finite element method leads to a large sparse linear system of equations. Direct linear solvers such as LU factorization or Cholesky factorization decompose the matrix A into lower and upper triangular matrices. The resulting system is fast to solve using algorithms based on Gaussian elimination. For high dimensional problems, finding the factorization matrices can be expensive, not only in terms of computational time, but also in terms of memory requirements. In these cases, iterative methods are more efficient because matrices are not stored. Instead, matrix-vector products are directly computed.

One of the most famous class of iterative solvers are Krylov subspace methods. Krylov subspace methods are iterative solvers that find the solution in a subspace spanned by images of \mathbf{b} under the first k powers of A . These subspaces are mathematically defined as:

$$K_m(A, \mathbf{b}) = \text{span} \{ \mathbf{b}, A\mathbf{b}, \dots, A^{m-1}\mathbf{b} \} \quad (2.40)$$

For $m \geq N + 1$, the vectors $A^{m-1}\mathbf{b}$ are linear combinations of the previous vectors that define the Krylov subspace basis. Therefore, there exist coefficients α_k such that:

$$\alpha_N A^N \mathbf{b} + \alpha_{N-1} A^{N-1} \mathbf{b} + \dots + \alpha_0 \mathbf{b} = 0 \quad (2.41)$$

Let k be the smallest integer such that $\alpha_k \neq 0$. Then, if we multiply Equation (2.41) by A^{-k} , we obtain an expression for $\mathbf{x} = A^{-1}\mathbf{b}$ of the form:

$$A^{-1}\mathbf{b} = -\frac{1}{\alpha_k} \left(\alpha_{k+1}\mathbf{b} + \dots + \alpha_N A^{N-k-1}\mathbf{b} \right) \quad (2.42)$$

This is a direct result of the Cayley-Hamilton theorem (Hoffman and Kunze, 2004). Two of the most efficient Krylov subspace methods are the conjugate gradient method (CG) (Hestenes and Stiefel, 1952), which is applicable to positive-definite matrices, and the

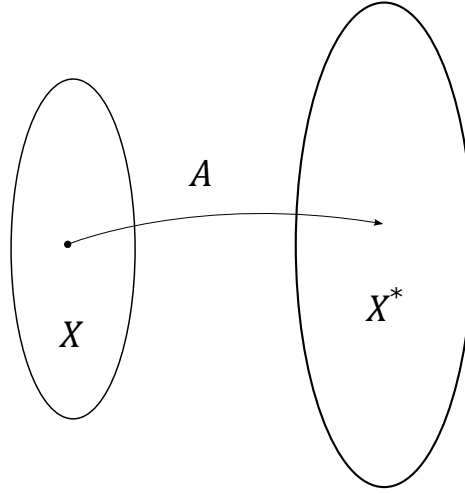


Figure 2.1 Mapping property of the operator \mathcal{A} .

generalized minimal residual method (GMRES) (Saad and Schultz, 1986), which can be applied to indefinite nonsymmetric matrices.

However, these methods fail when we have an operator with an unbounded spectrum (Mardal and Winther, 2011). An operator $\mathcal{A} : \mathcal{X} \rightarrow \mathcal{X}^*$ (see Figure 2.1) is said to be an unbounded weak operator if it is a bounded operator from a Hilbert space \mathcal{X} into its dual \mathcal{X}^* (Kirby, 2010). In these cases, as we refine the discretization, the number of iterations that the method needs to converge grows rapidly.

In order to understand the origin of this issue, let us consider the continuous version of the linear matrix system (2.39):

$$\mathcal{A}x = f \quad (2.43)$$

where $x \in \mathcal{X}$, $f \in \mathcal{X}^*$ and $\mathcal{A} : \mathcal{X} \rightarrow \mathcal{X}^*$.

The Krylov subspace of order m associated with the problem (2.43) is given by:

$$K_m(\mathcal{A}, f) = \text{span}\{f, \mathcal{A}f, \dots, \mathcal{A}^{m-1}f\} \quad (2.44)$$

The Krylov subspace (2.44) is not well defined because it involves the operations $\mathcal{A}^k f$, which are mathematically not valid due to the mapping properties of \mathcal{A} . In the discretized case we have that $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{b} \in \mathbb{R}^N$ and $A : \mathbb{R}^N \rightarrow \mathbb{R}^N$, but these subspaces are not the same². Although we can now perform the operations $A^k \mathbf{b}$, we may need many iterations to form a

¹The dual \mathcal{X}^* of a Hilbert space \mathcal{X} is the space of all bounded linear functionals on \mathcal{X} .

² \mathbf{b} can be an arbitrary vector in \mathbb{R}^N , but \mathbf{x} is a vector in \mathbb{R}^N that satisfies $A\mathbf{x} = \mathbf{b}$.

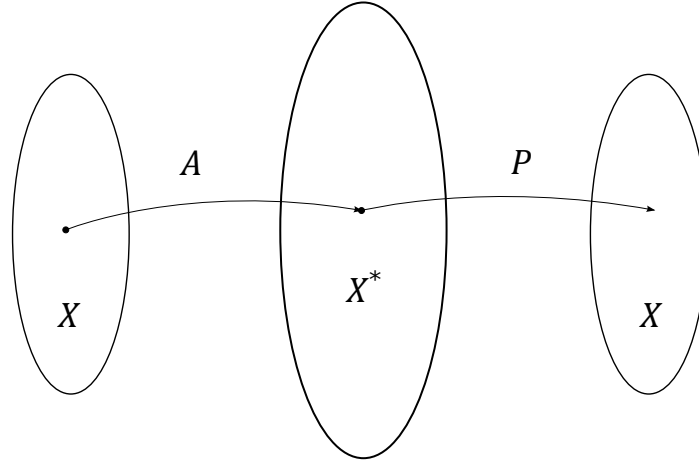


Figure 2.2 Mapping property of the operator corresponding to the composition of \mathcal{A} and \mathcal{P} .

basis whose linear combination contains a reasonable approximation of the solution vector, or even in the worst case, the iterative method will return a wrong value at $m = N + 1$.

This inconsistency is solved using the Riesz representation theorem, which states that a Hilbert space and its dual are isometrically isomorphic (Kirby 2010). Therefore, there exists a mapping $\mathcal{P} : \mathcal{X}^* \rightarrow \mathcal{X}$, which is referred to as the Riesz map. We can then left-multiply the operator \mathcal{A} with \mathcal{P} and obtain the following problem:

$$\mathcal{P}\mathcal{A}x = \mathcal{P}f \quad (2.45)$$

Now, the associated Krylov subspace is:

$$K_m(\mathcal{P}\mathcal{A}, \mathcal{P}f) = \text{span} \left\{ \mathcal{P}f, (\mathcal{P}\mathcal{A})(\mathcal{P}f), \dots, (\mathcal{P}\mathcal{A})^{m-1}(\mathcal{P}f) \right\} \quad (2.46)$$

The elements of the Krylov subspace (2.46) are inside the Hilbert space \mathcal{X} , which is the space where the solution x lies. Therefore, the Krylov subspace is now well-defined. A representation of the application of the Riesz map to the operator \mathcal{A} is illustrated in 2.2.

Returning to the discrete space, a preconditioner is just a matrix P that we left-multiply on both sides of the linear matrix system such that we obtain a problem that converges faster when we use a Krylov subspace method:

$$PAx = Pb \quad (2.47)$$

Thus, a trivial choice of a preconditioner is the discretized version of the Riesz map. This preconditioner ensures that the number of iterations required to converge is independent of the mesh size.

In summary, the idea of the Riesz map is to transform the original continuous problem into a new equation that, after discretization, is computationally less expensive, while preserving the information contained in the model.

Preconditioners based on the Riesz map are not the only choice. There exist a wide variety of operations that make the linear matrix system computationally cheaper to solve when using a Krylov subspace method.

The most simple preconditioners are just factorizations of the matrix A that are easier to invert, for example, the LU and Cholesky factorization. Other preconditioners are approximations of the inverse of A , so that $PA \simeq \mathbb{I}$. Some examples are the incomplete versions of LU and Cholesky factorization. These factorization matrices are faster to compute, but as we reduce the mesh size, the number of iterations to converge will grow because the approximations of the exact factorization matrices become worse.

Moreover, one does not need to have a preconditioner as a matrix. A computational procedure that applies a linear operation to a vector is what we need. Therefore, some iterative methods as Jacobi or Gauss-Seidel can be used as preconditioners. In particular, these preconditioners are very cheap to compute, but they require many iterations to converge. The slow convergence of iterative methods can be attributed to the difficulty of damping the residual along the direction of the slow eigenvectors (i.e., the eigenvectors with small decay rate).

Multigrid methods solve this issue. A classical multigrid solver considers a few iterations of Jacobi or Gauss-Seidel methods in the fine mesh. This process reduces the high frequency component of the error. The residual is then interpolated to a coarse mesh, where the coarse grid error vector can be easily computed. The solution is interpolated to the fine mesh and the process is repeated until convergence.

Now, we compare how different preconditioners behave in two different problems. Let us first consider the Poisson equation. The strong formulation of the Poisson problem is:

$$\Delta u = -f \quad \text{in } \Omega \quad (2.48a)$$

$$u = u_D \quad \text{on } \Gamma_D \quad (2.48b)$$

$$\nabla u \cdot n = g \quad \text{on } \Gamma_N \quad (2.48c)$$

where the Dirichlet boundary condition $u = u_D$ is a prescribed value for the unknown u on a subset Γ_D of the boundary and the Neumann boundary condition $\nabla u \cdot n = g$ is a prescribed value for the normal derivative of u on the remaining boundary Γ_N .

It is well-known that the discretization of the Poisson equation in space using the finite element method leads to ill-conditioned linear algebra because the Laplacian operator is unbounded (Lee and Min 2021). As the number of elements increases, the condition number of the matrix grows rapidly, making the Krylov subspace method computationally expensive. Therefore, it is important to consider a suitable preconditioner.

Let us multiply Equation (2.48) by a scalar-valued test function $v \in \hat{V}$, with \hat{V} to be determined:

$$\int_{\Omega} \Delta u v d\Omega = - \int_{\Omega} f v d\Omega \quad (2.49)$$

Integrating by parts the first term and taking into account that $\nabla u \cdot n = g$ on Γ_N and $v = 0$ on Γ_D ³:

$$\int_{\Omega} \nabla u \cdot \nabla v d\Omega = \int_{\Omega} f v d\Omega + \int_{\Gamma_N} g v d\partial\Omega \quad (2.50)$$

Thus, the abstract formulation of the Poisson problem is:

$$\begin{aligned} &\text{find } u \in V \text{ such that:} \\ &a(u, v) = F(v) \\ &\text{for all } v \in \hat{V} \end{aligned} \quad (2.51)$$

where

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v d\Omega \quad (2.52a)$$

$$F(v) = \int_{\Omega} f v d\Omega + \int_{\Gamma_N} g v d\partial\Omega \quad (2.52b)$$

We need one weak derivative on u and v . Moreover, we need to impose the Dirichlet boundary condition on Γ_D . Therefore, a suitable choice of the function spaces is:

$$V = \{v \in H^1(\Omega) : v = u_D \text{ on } \Gamma_D\} = H_{u_D}^1(\Omega) \quad (2.53a)$$

$$\hat{V} = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D\} = H_0^1(\Omega) \quad (2.53b)$$

³The test function is required to vanish at the boundaries where the associated trial function is known. For mathematical details see Langtangen and Mardal, 2019.

In this case, the Dirichlet boundary condition is called an essential boundary condition because it enters into the definition of the trial function space V , whereas the Neumann boundary condition is called a natural boundary condition because it enters into the variational problem (2.50).

In order to discretize the problem, we consider the Galerkin approximation of Equation (2.51). The idea is to seek a solution in a finite-dimensional subspace $V_h \in V$:

$$\begin{aligned} &\text{find } u_h \in V_h \text{ such that:} \\ &a(u_h, v_h) = F(v_h) \\ &\text{for all } v_h \in \hat{V}_h \end{aligned} \tag{2.54}$$

We then form a computational domain $\Omega_h \in \Omega$ and choose a basis for the subspace such that $V_h = \text{span}\{\phi_1, \dots, \phi_N\}$. Expanding u_h and v_h in terms of the basis gives:

$$u_h = \sum_{j=1}^N \hat{u}_j \phi_j \tag{2.55a}$$

$$v_h = \sum_{j=1}^N \phi_j \tag{2.55b}$$

After introducing these expressions into Equation (2.51), we obtain a linear matrix system:

$$A\hat{\mathbf{u}} = \mathbf{b} \tag{2.56}$$

where

$$A_{ji} = \int_{\Omega_h} \nabla \phi_i \cdot \nabla \phi_j d\Omega \tag{2.57a}$$

$$b_j = \int_{\Omega_h} f_j \phi_j d\Omega + \int_{\Omega_h} g_j \phi_j d\Omega \tag{2.57b}$$

Matrix A is usually referred to as the stiffness matrix. After choosing a basis for ϕ (for example, linear Lagrange elements of order 1), the linear system (2.56) is solved using a Krylov subspace method. However, we also need to find a suitable preconditioner. First,

we consider a preconditioner based on the Riesz map. The Poisson problem operator⁴ is a mapping $\mathcal{A} : H_0^1 \rightarrow H^{-1}$. Therefore, a trivial choice is of a Riesz map, \mathcal{P} , is:

$$\mathcal{P}^{-1} = \int_{\Omega} \nabla u \cdot \nabla v d\Omega \quad (2.58)$$

Thus, we have available a preconditioner P by discretizing \mathcal{P} , that does not depend on the refinement of the mesh. In this case, the preconditioner P is just the inverse of the stiffness matrix A . Therefore, factorizing A using exact Cholesky or LU factorization and taking their inverse is equivalent to find the discretization of \mathcal{P} .

Figure 2.3 illustrates the performance of different preconditioners. In the left figure, the preconditioners are compared for a fixed number of elements. We can observe that Cholesky factorization, LU factorization and Multigrid need a considerably smaller number of iterations to converge than the other preconditioners, especially Jacobi.

The right plot shows the number of iterations that each preconditioner needs to converge for different numbers of discretizing elements. We can observe that the number of iterations required to converge for Cholesky factorization, LU factorization and Multigrid are independent of the number of elements. This behaviour is expected for Cholesky and LU factorization as they are exact factorizations of the matrix A , but it may seem a surprise in the case of multigrid preconditioner. The reason of this is that multigrid solvers are spectrally equivalent to the inverse of the the Laplacian operator. As a consequence, multigrid behaves as a discretized version of a Riesz map. The mathematical explanation of this can be found in Bramble (2019); Hackbusch (1994); Trottenberg et al. (2000). In contrast, incomplete Cholesky, incomplete LU and Jacobi need more iterations as we reduce the mesh size. This is because incomplete Cholesky and LU factorization are approximations that become less accurate as we increase the number of elements of A .

Figure 2.3 gives algorithmic information about the preconditioners, but we still do not have explicit information about the computational cost. Table 2.1 collects the time required to converge for different preconditioners and number of grid points. For a large number of elements, the computational time of incomplete Cholesky and LU factorization are equivalent to exact LU, and smaller than exact Cholesky factorization, although the number of iterations required to converge are completely different. The reason is that computing the full Cholesky and LU factorization is very time consuming for a large number of elements. Multigrid preconditioner is by far the most efficient algorithm for this case, in terms of computational time, number of iterations and mesh independence.

In order to show that the choice of the preconditioner is strongly dependent of the form of the governing equations, we formulate the Poisson problem (2.48) in a different way. The

⁴The problem operator is defined such that $(\mathcal{A}u, v) = a(u, v)$.

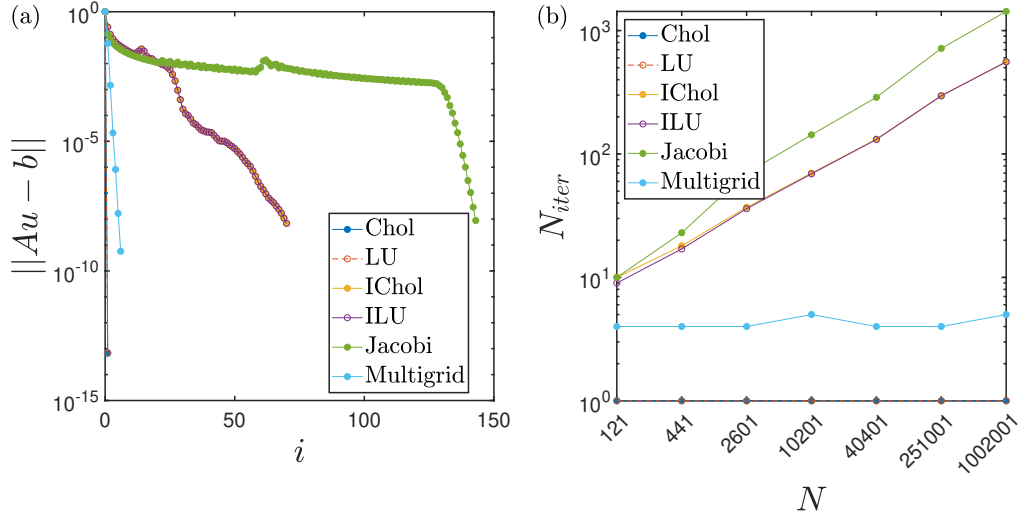


Figure 2.3 Performance of different preconditioners in the Poisson equation. (a) evolution of the residual norm with the number of iterations for $N = 10201$ and (b) number of iterations for different number of grid points. The relative tolerance is $1e - 8$, $r = \|Au - b\|$ is the residual, N_{iter} is the number of iterations necessary to reach the desired tolerance, i is the iteration and N is the number of grid points.

Table 2.1 Time required to obtain the solution of the Poisson equation for different choices of preconditioners and for different number of grid points.

Method	Number of grid points						
	121	441	2601	10201	40401	251001	1002001
Chol	0.0172	0.0221	0.0401	0.1426	0.6931	7.4959	49.6478
LU	0.0163	0.0187	0.0400	0.1269	0.5596	5.0664	37.0575
IChol	0.0163	0.0195	0.0362	0.1136	0.5751	6.0189	39.3289
ILU	0.0169	0.0194	0.0343	0.1515	0.7538	7.0049	39.1658
Jacobi	0.0174	0.0192	0.0338	0.1540	0.6861	7.7693	54.6831
Multigrid	0.0184	0.0192	0.0378	0.1248	0.4483	2.6590	10.2262

mixed formulation of the Poisson problem is formulated by adding a new variable $\sigma = \nabla u$. Thus, the strong form of the mixed Poisson equation is:

$$\sigma - \nabla u = 0 \quad \text{in } \Omega \quad (2.59a)$$

$$\nabla \cdot \sigma = -f \quad \text{in } \Omega \quad (2.59b)$$

$$u = u_D \quad \text{on } \Gamma_D \quad (2.59c)$$

$$\sigma \cdot n = g \quad \text{on } \Gamma_N \quad (2.59d)$$

Let us multiply Equation (2.59a) by a vector-valued test function $\tau \in \hat{\Sigma}$ and Equation (2.59b) by a scalar-valued test function $v \in \hat{V}$, with $\hat{\Sigma}$ and \hat{V} to be determined:

$$\int_{\Omega} \sigma \cdot \tau d\Omega - \int_{\Omega} \nabla u \cdot \tau d\Omega = 0 \quad (2.60a)$$

$$\int_{\Omega} (\nabla \cdot \sigma) v d\Omega = - \int_{\Omega} f v d\Omega \quad (2.60b)$$

We now integrate the second term in Equation (2.60a) by parts and set $u = u_D$ on Γ_D $\tau \cdot n = 0$ on Γ_N ⁵.

$$\int_{\Omega} \sigma \cdot \tau d\Omega + \int_{\Omega} u \nabla \cdot \tau d\Omega = \int_{\Gamma_D} u_D (\tau \cdot n) d\partial\Omega \quad (2.61a)$$

$$\int_{\Omega} (\nabla \cdot \sigma) v d\Omega = - \int_{\Omega} f v d\Omega \quad (2.61b)$$

The abstract formulation of Problem (2.61) is:

find $(u, \sigma) \in V \times \Sigma$ such that:

$$a(\sigma, \tau) + b(\tau, u) = G(\tau) \quad (2.62)$$

$$b(\sigma, v) = F(v)$$

for all $(v, \tau) \in \hat{V} \times \hat{\Sigma}$

where

⁵In this case, the flux normal to the boundary Γ_N is known, so we set $\tau \cdot n = 0$ at that boundary.

$$a(\boldsymbol{\sigma}, \boldsymbol{\tau}) = \int_{\Omega} \boldsymbol{\sigma} \cdot \boldsymbol{\tau} d\Omega \quad (2.63a)$$

$$b(\boldsymbol{\tau}, u) = \int_{\Omega} u \nabla \cdot \boldsymbol{\tau} d\Omega \quad (2.63b)$$

$$G(\boldsymbol{\tau}) = \int_{\Gamma_N} u_D (\boldsymbol{\tau} \cdot \boldsymbol{n}) d\partial\Omega \quad (2.63c)$$

$$F(v) = - \int_{\Omega} f v d\Omega \quad (2.63d)$$

A suitable choice of the subspaces is:

$$V = \{v \in L^2(\Omega)\} \quad (2.64a)$$

$$\hat{V} = \{v \in L^2(\Omega)\} \quad (2.64b)$$

$$\Sigma = \{\boldsymbol{\tau} \in H(\operatorname{div}, \Omega) : \boldsymbol{\tau} \cdot \boldsymbol{n} = g \text{ on } \Gamma_N\} = H_g(\operatorname{div}, \Omega) \quad (2.64c)$$

$$\hat{\Sigma} = \{\boldsymbol{\tau} \in H(\operatorname{div}, \Omega) : \boldsymbol{\tau} \cdot \boldsymbol{n} = 0 \text{ on } \Gamma_N\} = H_0(\operatorname{div}, \Omega) \quad (2.64d)$$

The variational problem (2.62) differs from the variational problem (2.51) in that the Dirichlet boundary condition enters now into the variational formulation, so it is a natural boundary condition. Similarly, the Neumann boundary condition enters now into the definition of the trial space V , so it is now an essential boundary condition.

The system of equations (2.62) is known as a saddle-point problem, and it has a block structured form.

We now consider the Galerkin approximation and find the solution in a subspace $V_h \times \Sigma_h \in V \times \Sigma$. We define the computational domain $\Omega_h \in \Omega$ and choose a basis for the subspaces $V_h = \operatorname{span}\{\phi_1, \dots, \phi_N\}$ and $\Sigma_h = \operatorname{span}\{\psi_1, \dots, \psi_N\}$. We expand the trial and test functions as follows:

$$u_h = \sum_{j=1}^N \hat{u}_j \phi_j \quad (2.65a)$$

$$v_h = \sum_{j=1}^N \phi_j \quad (2.65b)$$

$$\sigma_h = \sum_{j=1}^N \hat{\sigma}_j \psi_j \quad (2.65c)$$

$$\tau_h = \sum_{j=1}^N \psi_j \quad (2.65d)$$

After introducing these expressions into the Galerkin approximation of problem (2.62), we obtain:

$$\begin{bmatrix} A & B^* \\ B & 0 \end{bmatrix} \begin{bmatrix} \hat{\sigma} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} G \\ F \end{bmatrix} \quad (2.66)$$

where

$$A_{ji} = \int_{\Omega_h} \psi_i \cdot \psi_j d\Omega \quad (2.67a)$$

$$B_{ji}^* = \int_{\Omega_h} \phi_i \nabla \cdot \psi_j d\Omega \quad (2.67b)$$

$$B_{ji} = \int_{\Omega_h} \nabla \cdot \psi_i \phi_j d\Omega \quad (2.67c)$$

$$G_j = \int_{\Omega_h} u_{D_j} \psi_j \cdot n_j d\Omega \quad (2.67d)$$

$$F_j = - \int_{\Omega_h} f_j \phi_j d\Omega \quad (2.67e)$$

Establishing a stable discretization in mixed problems is more challenging as they must satisfy the inf-sup or Ladyzhenskaya–Babuška–Brezzi (LBB) conditions (Barth et al. 2005). For this particular problem, a stable choice of function spaces is the combination of order k Brezzi-Douglas-Marini (BDM) elements and order $k - 1$ discontinuous Galerkin elements (DG).

This mixed formulation of the problem loses the symmetry that the standard Poisson problem has. As a consequence, we are not allowed to use conjugate gradient as the Krylov subspace solver. We then use the GMRES solver instead. Again, we need to consider a

suitable preconditioner for the problem. In this case, we have more limitations because of the structure of the matrix A . For example, we cannot apply LU factorization directly to the system matrix because it has zeros in the diagonal. However, we can add a term $\varepsilon \int_{\Omega} uv d\Omega$ with $\varepsilon \ll 1$, so that all the diagonal elements are non-zero. Moreover, standard multigrid techniques fail when they are applied to this problem (Barth et al. 2005).

However, we can formulate a preconditioner based on a Riesz map associated with the problem (2.62). The mixed Poisson problem operator is a mapping $\mathcal{A} : L^2 \times H(\text{div}) \rightarrow L^2 \times H^*(\text{div})$. Therefore, a choice of the Riesz map, \mathcal{P} , is the following:

$$\mathcal{P}^{-1} = \int_{\Omega} \sigma \cdot \tau d\Omega + \int_{\Omega} (\nabla \cdot \sigma) (\nabla \cdot \tau) d\Omega + \int_{\Omega} uv d\Omega \quad (2.68)$$

The first two terms in Equation (2.68) is a mapping $H(\text{div}) \rightarrow H^*(\text{div})$, whereas the last term is a mapping $L^2 \rightarrow L^2$. Thus, the operator \mathcal{P} is a possible Riesz map for Problem (2.59). The corresponding preconditioner has a block-structured form:

$$P = \begin{bmatrix} S_1^{-1} & 0 \\ 0 & S_2^{-1} \end{bmatrix} \quad (2.69)$$

where S_1 is the discretized version of the first two terms in (2.68) and S_2 is the discretized version of the last term in (2.68).

Using this operator as a preconditioner ensures mesh independent convergence. We now need to find an efficient way of inverting this operator. The last term is just a mass matrix, so we can use sparse LU factorization to obtain S_2^{-1} . The first two terms are also sparse, but they are not just a mass matrix. Therefore, we can choose incomplete LU, algebraic multigrid or a few iterations of Jacobi solver to get an estimate of S_1^{-1} .

Figure 2.4 compares the performance of different preconditioners. The left plot shows the convergence of the residual for $N = 20200$ grid points. LU and the preconditioner based on the Riesz map converge considerably faster than ILU or Jacobi. Moreover, the number of iterations remains constant as we increase the number of elements when we use LU or Riesz map as preconditioners. This result confirms that (2.68) is a Riesz map of the mixed Poisson problem.

Table (2.2) collects the computational time required to solve the linear matrix system (2.66) for different preconditioners and number of grid points. ILU and Jacobi preconditioners become much more expensive in this case. In contrast, LU and the preconditioner based on the Riesz map show a similar behaviour.

In summary, the choice of the preconditioner strongly depends on the governing equations to solve. Ideally, one would choose the discretization of the Riesz map as the preconditioner because, as we have shown, the number of iterations will be independent of the number of

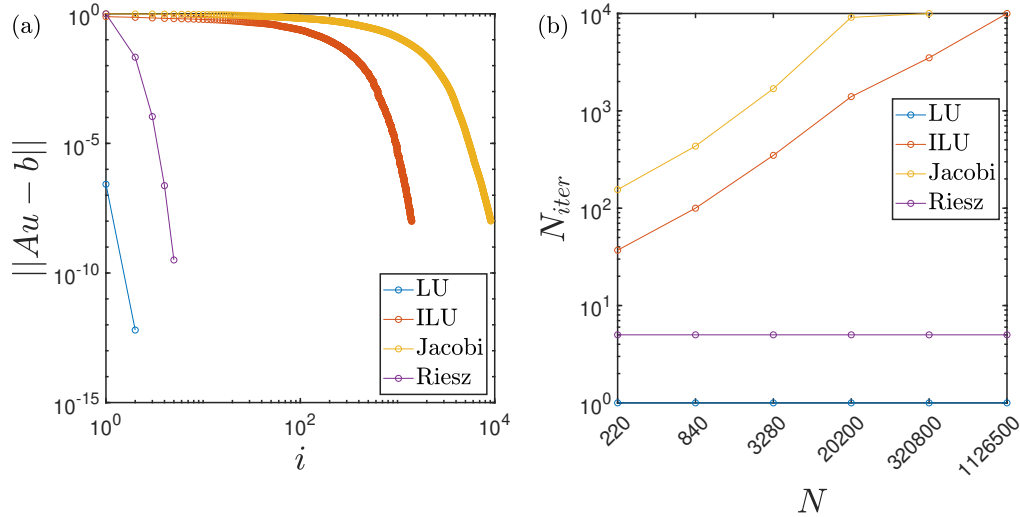


Figure 2.4 Performance of different preconditioners in the mixed Poisson equation. (a) evolution of the residual norm with the number of iterations for $N = 20200$ and (b) number of iterations for different number of grid points. The relative tolerance is $1e - 8$, $r = \|Au - b\|$ is the residual, N_{iter} is the number of iterations necessary to reach the desired tolerance, i is the iteration and N is the number of grid points.

Table 2.2 Time required to obtain the solution of the mixed Poisson equation for different choices of preconditioners and for different number of grid points. The hyphen symbol means that the algorithm has not converged in less than 10000 iterations.

Method	Number of grid points						
	220	840	3280	20200	80400	320800	1126500
LU	0.0006	0.0020	0.0082	0.1082	0.8504	5.3444	20.1714
ILU	0.0010	0.0045	0.0516	3.0025	31.3025	569.4607	-
Jacobi	0.0014	0.0091	0.1362	12.4096	-	-	-
Riesz map	0.0016	0.0029	0.0100	0.1263	0.9093	5.5262	21.2510

discretization elements. However, sometimes this choice of preconditioner leads to a long computation time because the inversion of the matrix can be expensive. In this case, it is preferable to use a less optimal preconditioner which will take a larger number of iterations to converge but the computational time will be shorter.

Chapter 3

Applications

In this chapter, we apply the adjoint method explained previously to different dynamical systems. The first application is a linear damped oscillator. We define the cost function as the kinetic energy at a given time and use the initial conditions as control parameters. The resulting dynamical system of two degrees of freedom can be solved analytically using the adjoint method. In this case, it is possible to obtain the gradient by exact differentiation, which allows us to verify the gradient computed through the adjoint approach. The second application is a one-dimensional viscous-acoustic problem governed by a system of PDEs. We show that this flow satisfies a diffusive wave equation and it is mathematically equivalent to a system of coupled damped oscillators. In this case, we solve the problem numerically by discretising in space using the finite element method and in time by using a finite difference scheme. We show the optimization process in a one-dimensional geometry, for which computational expense is not relevant. We verify the computed gradient by means of the Taylor test. We then consider a two-dimensional and a three-dimensional geometry and repeat the process. In this case, we also seek to compute the results with a reasonable computational cost by applying the concepts on iterative solvers and preconditioners explained previously.

3.1 Mass-spring-damper system

Understanding the behaviour of damped linear oscillators is important in acoustics because a generic thermoacoustic system behaves as a coupled system of oscillators Culick et al. (2006). Therefore, we first start by applying the adjoint method formulated in Section 2.2 to a mass-spring-damper model. The resulting optimization problem has an analytical solution. The displacement x around the equilibrium point of mass m attached to a spring of stiffness k and a damper of viscous damping b is given by the following differential equation:

$$m \frac{d^2 x}{dt^2} + b \frac{dx}{dt} + kx = 0 \quad (3.1)$$

To solve Equation (3.1), we need an initial condition for the displacement and an initial condition for its first time derivative:

$$x(0) = x_0 \quad (3.2a)$$

$$\left. \frac{dx}{dt} \right|_{t=0} = \left. \frac{dx}{dt} \right|_0 \quad (3.2b)$$

We can express this second order in time ODE as a system of first order in time ODEs by introducing the velocity variable, $y = \frac{dx}{dt}$:

$$\frac{d}{dt} \begin{Bmatrix} x \\ y \end{Bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{Bmatrix} x \\ y \end{Bmatrix} \quad (3.3)$$

which can be expressed in compact form as:

$$\frac{d\mathbf{q}}{dt} = L\mathbf{q} \quad (3.4a)$$

$$\mathbf{q}(0) = \mathbf{q}_0 \quad (3.4b)$$

where \mathbf{q} is the state vector that contains the displacement and the velocity.

We set the cost function to be the kinetic energy of the system at a time T .

$$J_F(\mathbf{q}(T)) = \frac{1}{2} (y(T))^2 \quad (3.5)$$

Thus, the formulation of the optimization problem reads:

$$\min_{\mathbf{c}} J = J_F(\mathbf{q}(T))$$

subject to:

$$\frac{d\mathbf{q}}{dt} = L(\mathbf{q}, \mathbf{c}) \quad (3.6)$$

$$\mathbf{q}(0) = \mathbf{q}_0$$

The first step is to obtain the time evolution of the state vector. Equation (3.4) is a homogeneous linear system of ODEs, so its solution can be expressed as a linear combination of pure exponential solutions:

$$\mathbf{q}(t) = a_1 \hat{\mathbf{q}}_1 e^{\lambda_1 t} + a_2 \hat{\mathbf{q}}_2 e^{\lambda_2 t} \quad (3.7)$$

where a_1 and a_2 are coefficients to be determined using the initial conditions.

The eigenvalues of the matrix L are:

$$\lambda_{1,2} = -\frac{b}{2m} \pm \frac{\sqrt{b^2 - 4mk}}{2m} \quad (3.8)$$

We assume that the system is underdamped, $b^2 - 4mk < 0$. In this case, the system has two complex conjugate eigenvalues:

$$\lambda_{1,2} = -\sigma \pm i\omega = -\frac{b}{2m} \pm i \frac{\sqrt{4mk - b^2}}{2m} \quad (3.9)$$

where σ is the decay rate and ω is the angular frequency.

After computing the eigenvectors and determining the coefficients a_k , the evolution of the state vector is given by the following expression:

$$\mathbf{q}(t) = \left\{ \begin{array}{l} e^{-\sigma t} (x_0 \cos(\omega t) + \frac{\sigma x_0 + y_0}{\omega} \sin(\omega t)) \\ e^{-\sigma t} [y_0 \cos(\omega t) - (\frac{\sigma}{\omega} (\sigma x_0 + y_0) + \omega x_0) \sin(\omega t)] \end{array} \right\} \quad (3.10)$$

We now define the adjoint problem. In this case, the control parameters are not parameters inside the governing equations. As a consequence, the adjoint equation is homogeneous:

$$\frac{d\mathbf{q}^\dagger}{dt} + L^T \mathbf{q}^\dagger = 0 \quad (3.11)$$

which needs to be integrated backwards in time using the terminal condition

$$\mathbf{q}^\dagger(T) = - \left(\frac{\partial J_F}{\partial \mathbf{q}(T)} \right)^T \quad (3.12)$$

where $\left(\frac{\partial J_F}{\partial \mathbf{q}(T)} \right)^T = \left[\begin{array}{cc} 0 & y(T) \end{array} \right]$.

However, by defining the adjoint time as $t^\dagger = T - t$, we can transform (3.11) into an initial value problem:

$$\frac{d\mathbf{q}^\dagger}{dt^\dagger} = L^T \mathbf{q}^\dagger \quad (3.13a)$$

$$\mathbf{q}^\dagger(0) = - \left(\frac{\partial J_F}{\partial \mathbf{q}(T)} \right)^T \quad (3.13b)$$

Now, we can solve the adjoint problem using the same method as in the direct problem. We obtain the time evolution of the adjoint state vector as a linear combination of pure exponential solutions. The eigenvalues of L^T are the same as the eigenvalues of L , but the eigenvectors are different. After computing the eigenvectors and determining the coefficients using the initial condition, the adjoint state vector is:

$$\mathbf{q}^\dagger(t^\dagger) = \left\{ \begin{array}{l} \frac{\sigma^2 + \omega^2}{\omega} y(T) e^{-\sigma t^\dagger} \sin(\omega t^\dagger) \\ y(T) e^{-\sigma t^\dagger} \left[-\cos(\omega t^\dagger) + \frac{\sigma}{\omega} \sin(\omega t^\dagger) \right] \end{array} \right\} \quad (3.14)$$

The next step is to compute the adjoint initial condition, which is the value of the adjoint state vector at the final time step due to the change in the time domain:

$$\mathbf{q}_0^\dagger = \left\{ \begin{array}{l} \frac{\sigma^2 + \omega^2}{\omega} y(T) e^{-\sigma T} \sin(\omega T) \\ y(T) e^{-\sigma T} \left[-\cos(\omega T) + \frac{\sigma}{\omega} \sin(\omega T) \right] \end{array} \right\} \quad (3.15)$$

Finally, because the initial conditions are the control parameters, $\frac{\partial \mathbf{q}_0}{\partial \mathbf{c}} = \mathbb{I}$, so the sensitivity of the cost function with respect to the parameters is given by the adjoint initial condition:

$$\delta J = -\mathbf{q}_0^\dagger \delta \mathbf{c} \quad (3.16)$$

From this result it turns out that the cost function is minimized if the initial conditions x_0 and y_0 are chosen such that adjoint initial condition is zero. Substituting the adjoint initial condition obtained in Equation (3.15) into Equation (3.16) yields:

$$\delta J = \left\{ \begin{array}{l} -\frac{\sigma^2 + \omega^2}{\omega} y(T) e^{-\sigma T} \sin(\omega T) \\ y(T) e^{-\sigma T} \left[\cos(\omega T) - \frac{\sigma}{\omega} \sin(\omega T) \right] \end{array} \right\} \delta \mathbf{c} \quad (3.17)$$

From Equation (3.17) we observe that the gradient becomes zero in two cases:

$$\sin(\omega T) = 0 \text{ and } \tan(\omega T) = \frac{\omega}{\sigma} \quad (3.18a)$$

$$y(T) = 0 \quad (3.18b)$$

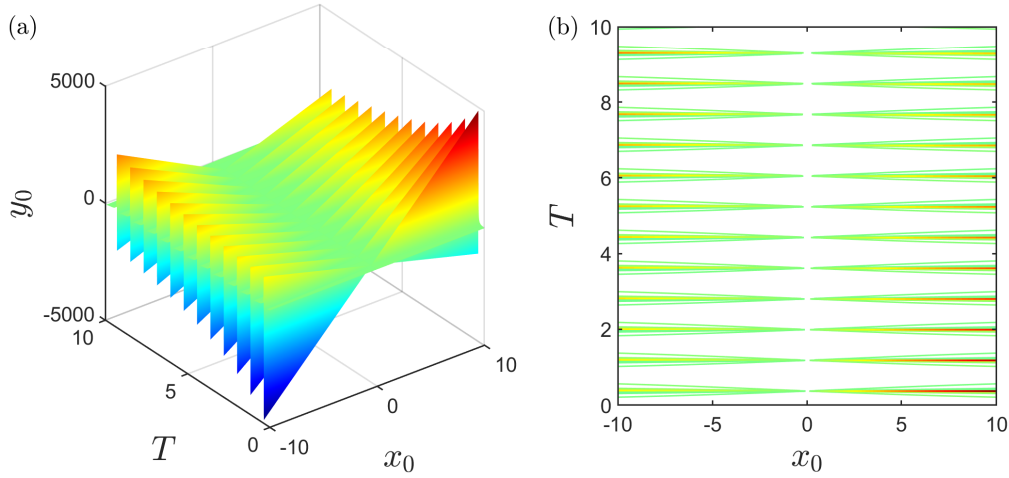


Figure 3.1 Optimal design space for different final times $T \in [0, 10]$. (a) is the surface representation and (b) is its contours. Both plots show that for some T , the required initial condition y_0 to minimize the kinetic energy is very high in magnitude.

The first condition is satisfied if $\omega = 0$. However, this condition is incompatible with the underdamped system assumption. Therefore, if $\omega = 0$, the previous analysis is not valid. Instead, the second expression contains a relation that the initial conditions must satisfy in order to minimize the cost function. This equation represents the optimal design space:

$$\frac{y_0}{x_0} = \frac{\left(\frac{\sigma^2}{\omega} + \omega\right) \tan(\omega T)}{1 - \frac{\sigma}{\omega} \tan(\omega T)} \quad (3.19)$$

From this result we observe that, if $\omega T = \arctan\left(\frac{\omega}{\sigma}\right)$, it is not possible to minimize the kinetic energy at time T because for any x_0 , an infinite value of y_0 would be required. Figure 3.1 illustrates the optimal design space defined by Equation (3.19) for $m = 1$, $b = 1$ and $k = 4$. For a given T , the design space is a straight line of slope given by Equation (3.19).

Figure 3.2 shows the evolution of the state vector and the cost function for the dynamical system for the uncontrolled case. The initial conditions considered are $x_0 = 1$ and $y_0 = 0$. The final time is $T = 2$. We can observe that the cost function at the final time is not zero. Figure 3.3 illustrates the same variables as in Figure 3.2, but for the controlled system. In this case, we maintain $x_0 = 1$, but y_0 is given by Equation (3.19). We can observe that the energy at $t = T$ has reached zero value.

In order to check that the gradient obtained by the adjoint method is correct, we compute the gradient by differentiating the cost function directly, since this low-order system has an

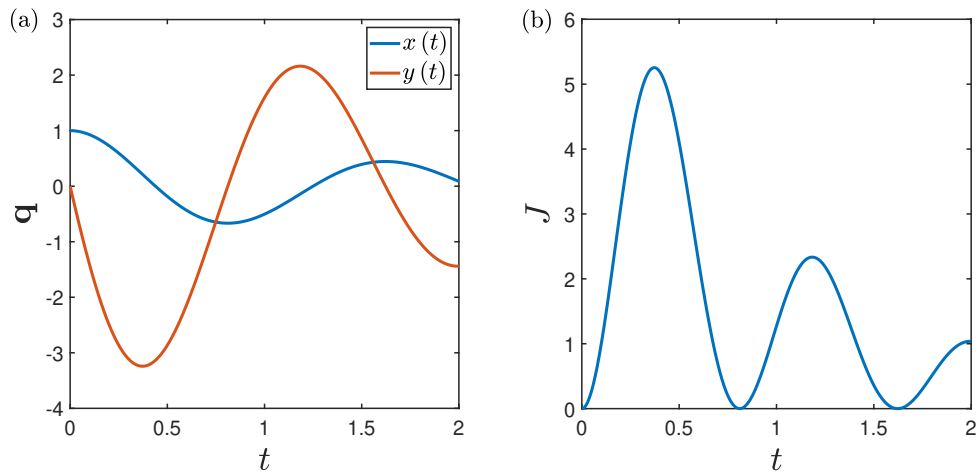


Figure 3.2 Evolution of the state vector (a) and the cost function (b) for the uncontrolled damped oscillator.

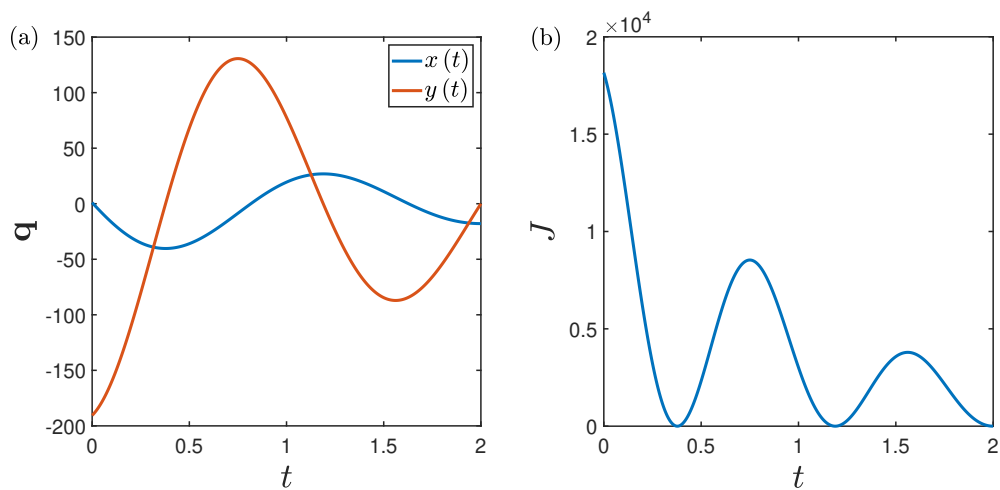


Figure 3.3 Evolution of the state vector (a) and the cost function (b) for the controlled damped oscillator.

analytical solution. By taking the derivative of the cost function with respect to the control parameters, we obtain:

$$\nabla J = \frac{\partial J_F}{\partial y(T)} \frac{\partial \mathbf{q}(T)}{\partial \mathbf{c}} \quad (3.20)$$

After substituting the terms into the expression above we get:

$$\nabla J = y(T) \left\{ \begin{array}{c} \frac{\partial y(T)}{\partial x_0} \\ \frac{\partial y(T)}{\partial y_0} \end{array} \right\} = \left\{ \begin{array}{c} -\frac{\sigma^2 + \omega^2}{\omega} y(T) e^{-\sigma T} \sin(\omega T) \\ y(T) e^{-\sigma T} [\cos(\omega T) - \frac{\sigma}{\omega} \sin(\omega T)] \end{array} \right\} \quad (3.21)$$

which is equivalent to the expression (3.17) obtained by means of the adjoint method.

Although both results are the same, we can observe a fundamental difference in the way that the gradients have been computed. When performing direct differentiation, we need to compute a partial derivative for each parameter to compute the Jacobian of the solution, $\frac{\partial \mathbf{q}(T)}{\partial \mathbf{c}}$. This operation is expensive when we have many parameters. On the contrary, the solution of the adjoint equation does not depend on the number of parameters.

3.2 Viscous-acoustic system

As a second model problem, we consider an acoustic flow in a duct. The fluid has a uniform density ρ_0 , a uniform pressure p_0 , and a kinematic viscosity ν . The thermal dissipation of the system is neglected. We set the mean flow variables to zero and study the acoustic oscillations for the velocity u and pressure p . We assume that the flow is one dimensional, so the flow variables depend on one spatial coordinate only. Under these assumptions, the conservation of momentum and energy equations are:

$$\frac{\partial u}{\partial t} + \frac{1}{\rho_0} \frac{\partial p}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2} \quad (3.22a)$$

$$\frac{\partial p}{\partial t} + \gamma p_0 \frac{\partial u}{\partial x} = 0 \quad (3.22b)$$

where γ is the ratio of specific heats, $\gamma = \frac{c_p}{c_v}$. Figure 3.4 shows a diagram of the domain Ω .

By taking the partial derivative of (3.22a) with respect to t and subtracting the divergence of (3.22b), we can eliminate the pressure to obtain a single expression for the velocity:

$$\frac{\partial^2 u}{\partial t^2} - c_0^2 \frac{\partial^2 u}{\partial x^2} = \nu \frac{\partial^3 u}{\partial t \partial x^2} \quad (3.23)$$

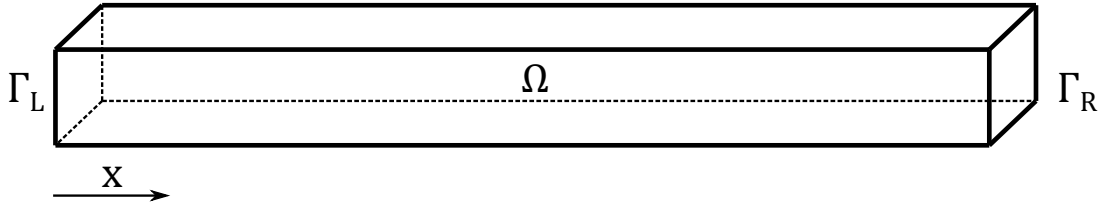


Figure 3.4 Diagram of the duct that contains the acoustic system. Ω is the domain, Γ_L is the left boundary and Γ_R is the right boundary.

where $c_0^2 = \left(\frac{\gamma p_0}{\rho_0}\right)^2$ is the squared sound speed.

Equation (3.23) represents the wave equation for the acoustic velocity with viscous dissipation. To solve this problem, we need to provide the initial conditions for the velocity and its first time derivative. Similarly, we need to impose certain boundary conditions. In this case, we consider Dirichlet boundary conditions for u at the left and right ends of the tube.

$$u(x, 0) = 0 \quad \frac{\partial u}{\partial t}(x, 0) = 0 \quad (3.24a)$$

$$u(x_{\Gamma_L}, t) = U_L(t) \quad u(x_{\Gamma_R}, t) = U_R(t) \quad (3.24b)$$

We now define the optimization problem. We prescribe the value of the time-dependent boundary condition at the left end, $U_L(t)$. As a result, waves generated at that boundary propagate within the domain. This will increase the acoustic kinetic energy of the flow, which will be dissipated at a rate given by the kinematic viscosity. Once this forcing ceases, the system will return to the quiescent state due to viscous dissipation. However, this process will take some time if no control is considered.

We then use the right end boundary condition $U_R(t)$ as an open loop control to minimize the acoustic kinetic energy at a given time T . Thus, the optimal control problem reads:

$$\begin{aligned} \min_c J &= J_F(T) \\ \text{subject to:} \\ \frac{\partial^2 u}{\partial t^2} - c_0^2 \frac{\partial^2 u}{\partial x^2} &= \nu \frac{\partial^3 u}{\partial t \partial x^2} \\ u(x, 0) &= 0, \quad \frac{\partial u}{\partial t}(x, 0) = 0 \\ u(x_{\Gamma_L}, t) &= U_L(t), \quad u(x_{\Gamma_R}, t) = U_R(t) \end{aligned} \quad (3.25)$$

where the cost function is defined as

$$J_F(T) = \frac{1}{2} (u(T))^2 \quad (3.26)$$

The first step to solve the optimization problem is to solve the direct problem (3.23). In this case, we solve the system numerically by discretising the equations in space using the finite element method and in time using the finite difference method.

We transform the second order in time PDE into a first order in time system of PDEs by defining a new variable $\eta = \frac{\partial u}{\partial t}$:

$$\frac{\partial u}{\partial t} = \eta \quad (3.27a)$$

$$\frac{\partial \eta}{\partial t} = c_0^2 \frac{\partial^2 u}{\partial x^2} + \nu \frac{\partial^2 \eta}{\partial x^2} \quad (3.27b)$$

We now derive the weak form of Equation (3.27). Let us multiply (3.27a) by a scalar-valued test function $v \in \hat{V}$ and (3.27b) by a scalar-valued test function $w \in \hat{W}$, with \hat{V} and \hat{W} to be determined. We integrate by parts the terms containing second order derivatives in space. The resulting surface integrals disappear if we choose $w = 0$. Thus, we obtain:

$$\frac{d}{dt} \int_{\Omega} uv d\Omega = \int_{\Omega} \eta v d\Omega \quad (3.28a)$$

$$\frac{d}{dt} \int_{\Omega} \eta w d\Omega = -c_0^2 \int_{\Omega} \frac{\partial u}{\partial x} \frac{\partial w}{\partial x} d\Omega - \nu \int_{\Omega} \frac{\partial \eta}{\partial x} \frac{\partial w}{\partial x} d\Omega \quad (3.28b)$$

By inspection of the system (3.28), we observe that we need one weak spatial derivative on u , and to impose Dirichlet boundary conditions on this variable. Then, the choice $\hat{V} = H_0^1(\Omega)$ is sufficient. For η and q , we also need one weak spatial derivative. Therefore, we can also choose $\hat{W} = H^1(\Omega)$.

To check that this choice is appropriate, let us assume that (u_s, η_s) is a solution of Equation (3.27). Then, the problem is well-defined if $(u_s + C_1(t), \eta_s + C_2(t))$ is not a solution (i.e., the problem has a unique solution). Substituting into Equation (3.27) yields:

$$\frac{\partial u_s}{\partial t} + \frac{dC_1}{dt} = \eta_s + C_2(t) \quad (3.29a)$$

$$\frac{\partial \eta_s}{\partial t} + \frac{dC_2}{dt} = c_0^2 \frac{\partial^2 u_s}{\partial x^2} + \nu \frac{\partial^2 \eta_s}{\partial x^2} \quad (3.29b)$$

Since (u_s, η_s) satisfies (3.27), this equation is simplified as:

$$\frac{dC_1}{dt} = C_2(t) \quad (3.30a)$$

$$\frac{dC_2}{dt} = 0 \quad (3.30b)$$

Expression (3.30a) implies that C_2 is constant in time. Taking into account the initial conditions, we have that $\eta(0) = \eta_s(0) + C_2(0) = 0$, so the value of the constant C_2 is zero. As a result, C_1 must also be constant in time and, by imposing the initial condition in u , we get the same result that $C_1 = 0$. This analysis proves the correct choice of the function spaces.

Therefore, the abstract variational formulation of the direct problem is:

find $(u, \eta) \in V \times W$ such that:

$$\frac{d}{dt} (a(u, v)) = a(\eta, v) \quad (3.31)$$

$$\frac{d}{dt} (a(\eta, w)) = -c_0^2 a\left(\frac{\partial u}{\partial x}, \frac{\partial w}{\partial x}\right) - va\left(\frac{\partial \eta}{\partial x}, \frac{\partial w}{\partial x}\right)$$

for all $(v, w) \in \hat{V} \times \hat{W}$

where

$$a(u, v) = \int_{\Omega} uv d\Omega \quad (3.32)$$

We now discretize the problem (3.31) using the Galerkin approximation. Let us consider the closed finite-dimensional subspace $V_h \in V$ and $W_h \in W$. We form a basis composed by piecewise polynomial basis functions ϕ_i such that $V_h = \text{span}(\phi_1, \phi_2, \dots, \phi_N)$ and $W_h = \text{span}(\phi_1, \phi_2, \dots, \phi_N)$. We can then express the problem variables as a linear combination of the basis elements:

$$u(x,t) = \sum_{i=1}^N \hat{u}_i(t) \phi_i(x) \quad (3.33a)$$

$$v(x) = \sum_{i=1}^N \phi_i(x) \quad (3.33b)$$

$$\eta(x,t) = \sum_{i=1}^N \hat{\eta}_i(t) \phi_i(x) \quad (3.33c)$$

$$w(x) = \sum_{i=1}^N \phi_i(x) \quad (3.33d)$$

We substitute these variables into the governing equations in weak form and obtain the following expressions:

$$\sum_{i=1}^N \frac{d\hat{u}_i}{dt} \int_{\Omega} \phi_i \phi_j d\Omega = \sum_{i=1}^N \hat{\eta}_i \int_{\Omega} \phi_i \phi_j d\Omega \quad (3.34a)$$

$$\sum_{i=1}^N \frac{d\hat{\eta}_i}{dt} \int_{\Omega} \phi_i \phi_j d\Omega = -c_0^2 \sum_{i=1}^N \hat{u}_i \int_{\Omega} \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} d\Omega - \nu \sum_{i=1}^N \hat{\eta}_i \int_{\Omega} \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} d\Omega \quad (3.34b)$$

which can be written in vector form as:

$$\frac{d}{dt} \begin{Bmatrix} \hat{\mathbf{u}} \\ \hat{\boldsymbol{\eta}} \end{Bmatrix} = \begin{bmatrix} 0 & \mathbb{I} \\ -c_0^2 M^{-1} K & -\nu M^{-1} K \end{bmatrix} \begin{Bmatrix} \hat{\mathbf{u}} \\ \hat{\boldsymbol{\eta}} \end{Bmatrix} \quad (3.35)$$

The matrices M and K are referred to as mass matrix and stiffness matrix, and they are given by:

$$M_{ji} = \int_{\Omega} \phi_i \phi_j d\Omega \quad (3.36a)$$

$$K_{ji} = \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j d\Omega \quad (3.36b)$$

are the mass matrix and the stiffness matrix, respectively.

Equation (3.35) represents a system of time-dependent ODEs, $\frac{d\mathbf{q}}{dt} = \mathbf{L}\mathbf{q}$. We can observe that the form of this system is equivalent to the equation that describes the damped oscillator introduced in Section 3.1. This shows that this acoustic system behaves as a coupled system of damped oscillators in which the damping is proportional to the stiffness.

For the time discretization, we consider the Crank-Nicolson scheme (Crank and Nicolson, 1947). Therefore, we split the continuous time domain \mathcal{T} into N_T intervals of the same size Δt . Given the state vector at time step n , \mathbf{q}^n , we obtain the state vector at the next time step, \mathbf{q}^{n+1} , as follows:

$$\frac{\mathbf{q}^{n+1} - \mathbf{q}^n}{\Delta t} = \frac{1}{2}L(\mathbf{q}^{n+1} + \mathbf{q}^n) \quad (3.37)$$

Equation (3.37) can be rearranged as a linear matrix system:

$$A\mathbf{q}^{n+1} = \mathbf{b}(\mathbf{q}^n) \quad (3.38)$$

where $A = \mathbb{I} - \frac{\Delta t}{2}L$ and $\mathbf{b}(\mathbf{q}^n) = [\mathbb{I} + \frac{\Delta t}{2}L]\mathbf{q}^n$.

Equation (3.38) has to be solved using a matrix system solver, for example, a Krylov subspace method. Note that the matrix A is constant, but the right hand side needs to be updated at each time step.

Once the direct problem is solved, we can formulate the adjoint problem. In this case, we use a discrete-adjoint approach, so we start from the system of ODEs that arises from the spatial discretization, that is, Equation (3.35). In order to reuse the same finite differences algorithm for the time integration, we define the adjoint time as $t^\dagger = T - t$. Thus, we rewrite the adjoint equation as the following initial value problem:

$$\frac{d}{dt^\dagger} \begin{Bmatrix} \hat{\mathbf{u}}^\dagger \\ \hat{\boldsymbol{\eta}}^\dagger \end{Bmatrix} = \begin{bmatrix} 0 & \mathbb{I} \\ -c_0^2 M^{-1} K & -v M^{-1} K \end{bmatrix}^T \begin{Bmatrix} \hat{\mathbf{u}}^\dagger \\ \hat{\boldsymbol{\eta}}^\dagger \end{Bmatrix} \quad (3.39a)$$

$$\mathbf{q}^\dagger(0) = - \begin{Bmatrix} \hat{\mathbf{u}}(T) \\ 0 \end{Bmatrix} \quad (3.39b)$$

After applying the Crank-Nicolson scheme, we obtain the linear matrix system

$$A^\dagger \mathbf{q}^{\dagger n+1} = \mathbf{b}^\dagger(\mathbf{q}^{\dagger n}) \quad (3.40)$$

where $A^\dagger = \mathbb{I} - \frac{\Delta t}{2}L^T$ and $\mathbf{b}^\dagger(\mathbf{q}^{\dagger n}) = [\mathbb{I} + \frac{\Delta t}{2}L^T]\mathbf{q}^{\dagger n}$.

Finally, we obtain the sensitivity of the cost function with respect to the problem parameters as follows:

$$\delta J_F = - \left(\int_0^T \mathbf{q}^{\dagger T} \frac{\partial (L\mathbf{q})}{\partial \mathbf{c}} dt \right) \delta \mathbf{c} \quad (3.41)$$

The last variable that we need to compute to obtain the sensitivity is $\frac{\partial f}{\partial \mathbf{c}} = \frac{\partial(L\mathbf{q})}{\partial \mathbf{c}}$. By inspecting $L\mathbf{q}$, we can determine that it is a matrix of size $N \times N_T$ with zeros in all the entries except the row corresponding to the control boundary. Therefore, the gradient will be proportional to a vector containing the values of the adjoint time derivative of the acoustic velocity at the right boundary at each time step:

$$\delta J_F \propto - \left[\hat{\eta}_{x_{TR}}^\dagger(t_0) \quad \cdots \quad \hat{\eta}_{x_{TR}}^\dagger(t_{N_T}) \right]^T \delta \mathbf{c} \quad (3.42)$$

From Equation (3.42) we can establish a minimization condition. The cost function is minimized if the control vector is such that the adjoint time derivative of the acoustic velocity at the control boundary is zero at all time steps.

Thus, in order to minimize the cost function, one can compute the full gradient from Equation (3.41) or the proportional quantity defined in Equation (3.42). In both cases, the gradient vector points to the same direction in the design space. However, it is important to remark that if we use the Taylor test for verification, the latter approach will fail, although both gradients will give us the same optimal solution.

The solution steps for this application are summarized in Algorithm 1.

Algorithm 1: Finding the optimal control vector that minimizes a cost function defined at the final time step using the adjoint method

Input: Initial guess of the control vector, \mathbf{c}_0 .

Result: Optimal control vector, \mathbf{c}_{opt} .

while *Stopping criterion is not satisfied* **do**

 Compute the acoustic state vector \mathbf{q} at each time step by solving Equation (3.38);

 Solve the adjoint problem (3.40) to get the adjoint state vector \mathbf{q}^\dagger ;

 Compute the gradient of the cost function using Equation (3.41);

 Update the control parameters \mathbf{c} using a line search algorithm;

 Check for convergence;

end

3.2.1 One-dimensional geometry

We first consider a one-dimensional domain of length L , $x \in [0, L]$. The final time is set to the time that acoustic perturbations take to travel from one end to the other, that is, $T = \frac{L}{c_0}$. At the left boundary, we prescribe the following unsteady forcing:

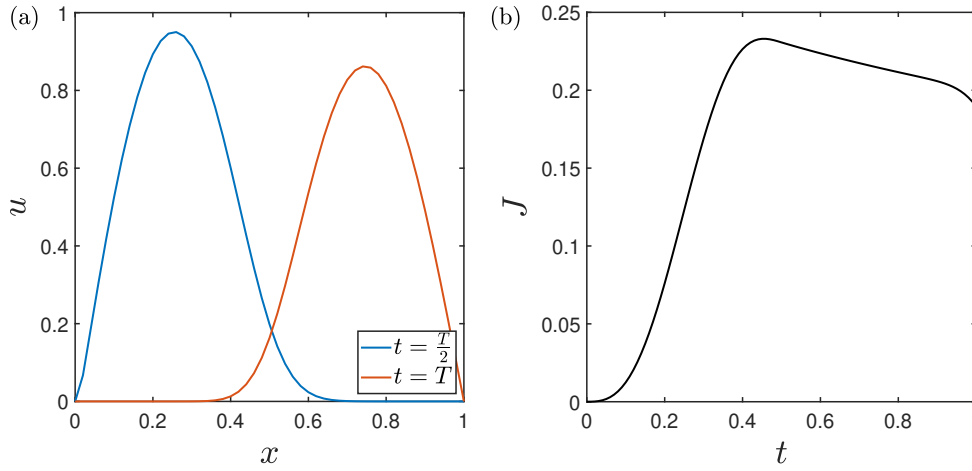


Figure 3.5 Snapshots of the acoustic velocity field (a) and evolution of the cost function (b) for the uncontrolled one-dimensional system.

$$U_L(t) = \begin{cases} A \sin(\omega t) & \text{for } 0 \leq t \leq \frac{T}{2} \\ 0 & \text{for } t > \frac{T}{2} \end{cases} \quad (3.43)$$

where ω is the angular frequency and A is the amplitude.

We set the system parameters to $c_0 = 1$ and $\nu = 0.01$. The parameters of the prescribed boundary condition are $A = 1$ and $\omega = 2\pi$. The domain length is $L = 1$. For these values, the time interval is $T = 1$.

We then discretize the domain using P_1 linear Lagrange elements of size $\Delta x = 0.02$. As a result, we need to solve a sparse linear matrix system of size $\mathcal{O}(10^2)$. Since the size of the system is not very large, we can consider a sparse direct solver. In particular, we choose UMFPACK. We discretize the time domain in intervals of size $\Delta t = 0.01$, so we have 100 control parameters.

First, we solve the uncontrolled problem. Figure 3.5 shows two snapshots of the acoustic velocity field at $t = \frac{T}{2}$ and $t = T$, along with the time evolution of the cost function. The wave generated at $x = x_{\Gamma_L}$ reaches half of the duct at $t = \frac{T}{2}$, which is consistent with the value of the sound speed considered. At $t = T$, the acoustic perturbations reach the right end of the domain, but the amplitude has decreased due to viscous dissipation. We can also observe that the acoustic energy grows while the forcing is acting. However, from $t = \frac{T}{2}$, the acoustic energy decreases due to viscous dissipation. Since the kinematic viscosity is considerably low, there is a remaining amount of energy at $t = T$.

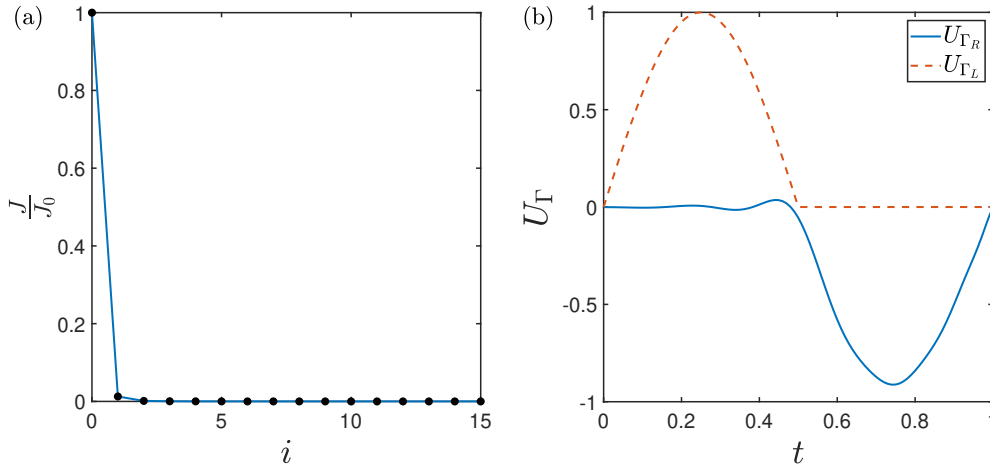


Figure 3.6 Convergence of the cost function (a) and optimal control (b) for the one-dimensional geometry.

We then perform the optimization process in order to damp the residual acoustic energy at $t = T$. The convergence results are collected in Figure 3.6. The left plot illustrates how the cost function decreases with the number of iterations of the optimization algorithm. The algorithm converges very quickly. In fact, it is able to minimize the cost function by 99% with respect to its initial value in the first iteration. The right plot shows the unsteady forcing along with the optimal control obtained as a result of the optimization. The control has a negative value, which means that the control boundary is generating acoustic waves of the opposite sign as the ones generated by the forcing.

Figure 3.7 shows the flow field at $t = \frac{T}{2}$ and $t = T$ when the optimal control is considered. At $t = \frac{T}{2}$, the acoustic velocity field is practically the same as in the uncontrolled case. The only difference is that the control has just started to act, so we can observe that the acoustic velocity is non-zero close to the right boundary. At $t = T$, the acoustic velocity is zero everywhere, leading to a zero acoustic kinetic energy. From $t = 0$ to $t = \frac{T}{2}$, the cost function evolves as in the uncontrolled case. However, from $t = \frac{T}{2}$, the evolution is completely different. At first, the cost function increases because the control is generating new acoustic waves. When these waves of negative velocity meet the waves generated by the forcing, they cancel each other and the kinetic energy goes to zero rapidly.

Finally, we verify the computed gradients by performing a Taylor test as explained in Section 2.2.1. Figure 3.8. shows the value of the second-order Taylor remainder for different values of the perturbation size h . When we plot the Taylor remainder against h^2 , we can

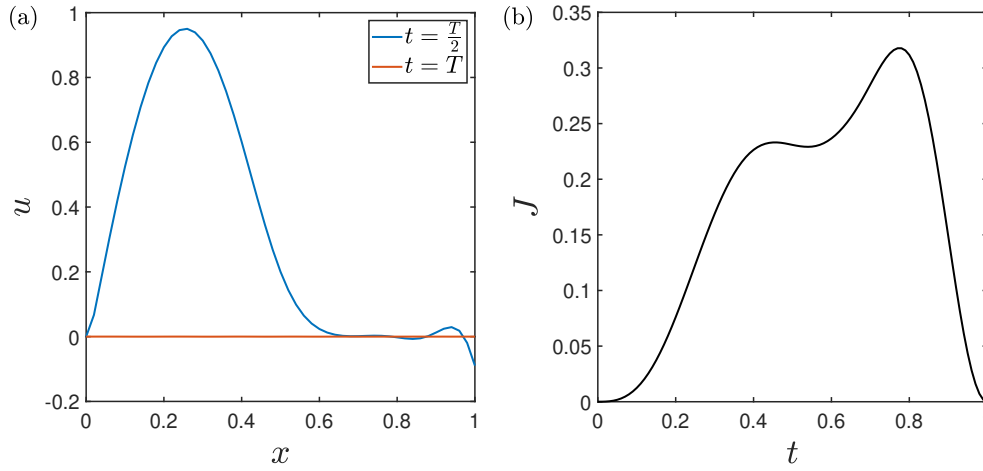


Figure 3.7 Snapshots of the acoustic velocity field (a) and evolution of the cost function (b) for the controlled one-dimensional system.

observe a straight line that passes through the origin. As a result, we can state that the gradient that we obtain using the adjoint method is accurate up to first order.

3.2.2 Extension to two-dimensional and three-dimensional geometries

The next step is to extend the algorithm to two-dimensional and three-dimensional geometries. It is important to remark that the problem is still one-dimensional in the sense that the flow variables depend on one spatial coordinate only. This observation will be useful to check that the solution is the same as the one obtained in the previous section.

The purpose of the extension to higher dimensions is to apply the adjoint method for gradient computation with a reasonable computational cost. For this objective, we will apply the concepts explained in Section 2.3 on preconditioning of Krylov subspace methods and parallel computing.

Let us begin with a two-dimensional geometry. We consider a square domain of side length L , $\mathbf{x} \in [0, L] \times [0, L]$. The time domain is divided into intervals of size $\Delta t = 0.01$. We discretize the domain using P_1 triangular elements and use 50 cells for the x and y directions. For these values, the linear matrix system has size $N \sim \mathcal{O}(10^4)$. Since the size of the matrix becomes large, we now consider iterative methods. The matrix is positive-definite, so we use the conjugate gradient method. If we apply this method to the system naively, the algorithm will take many iterations to converge because the condition number of the matrix is very large.

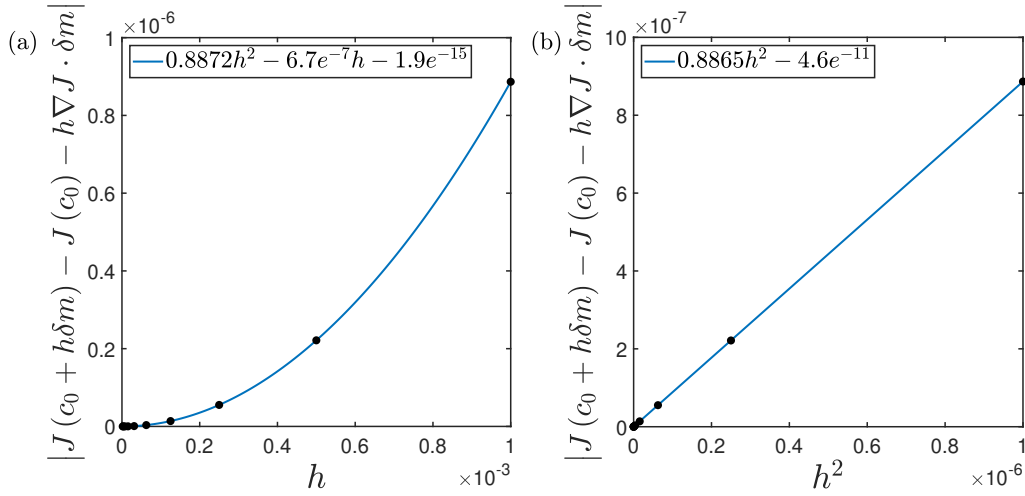


Figure 3.8 Second order Taylor remainder with quadratic fitting (a) and linear fitting (b).

Therefore, a suitable preconditioner must be considered. In this case, we use incomplete Cholesky factorization.

Figure 3.9 shows two snapshots of the acoustic velocity at $t = \frac{T}{2}$ and $t = T$. The forcing generates a sinusoidal wave that travels along the domain and reaches the right end at $T = \frac{L}{c_0}$. Although some momentum has been diffused due to the fluid viscosity, the fluid still contains a considerable amount of kinetic energy at the final time.

We then perform the optimization process using the adjoint method. Figure 3.10 illustrates the value of the cost function as the number of iterations of the optimization algorithm increases. The convergence is slightly slower than in the one-dimensional case. We can also observe that the optimal boundary condition is practically identical to the one in the one-dimensional case, as expected.

Figure 3.10 illustrates the flow field when the optimal control is applied to the right boundary. As in the one-dimensional case, the acoustic velocity field at $t = \frac{T}{2}$ is almost the same as in the uncontrolled case. However, the optimal boundary condition is able to drive the system to the quiescent state at $t = T$.

We finally consider a cubic domain of side length L , $\mathbf{x} \in [0, L] \times [0, L] \times [0, L]$. We discretize the domain using 50 cells per direction, so that the resulting matrix has size $N \sim \mathcal{O}(10^5)$. In this case, the size of the linear system is even larger than in the two-dimensional case. Therefore, the computation is distributed over 10 processor cores based on a message passing interface (MPI) parallelization. We again use conjugate gradient for solving the linear system, but in this case we use a geometric multigrid V-cycle preconditioner.

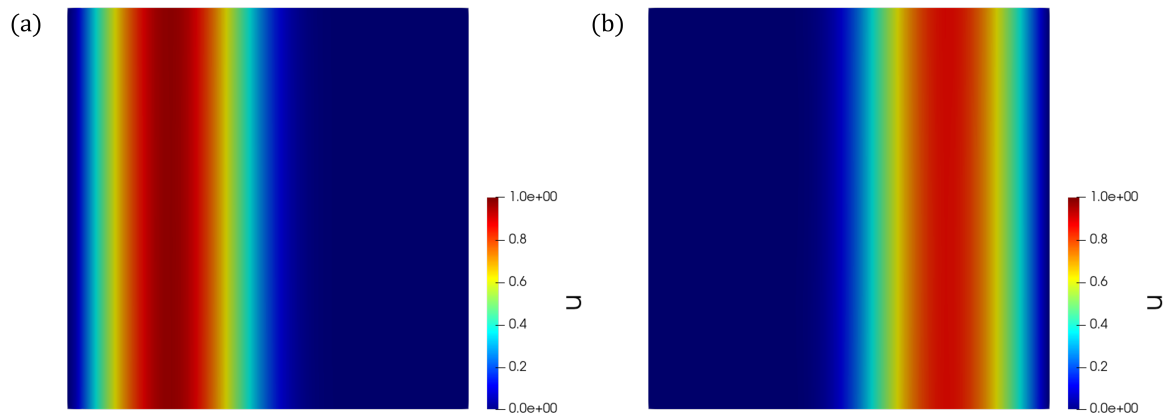


Figure 3.9 Snapshots of the acoustic velocity field at $t = \frac{T}{2}$ (a) and $t = T$ (b) for the uncontrolled two-dimensional system.

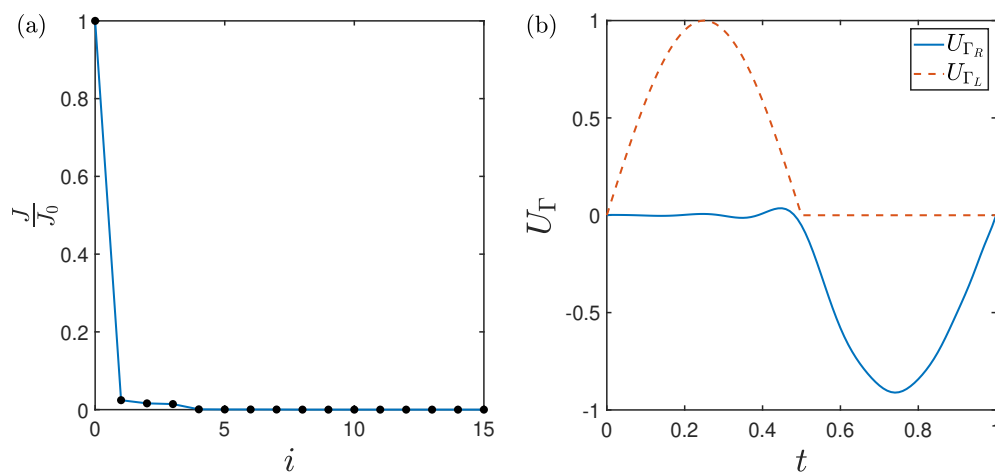


Figure 3.10 Convergence of the cost function (a) and optimal control (b) for the two-dimensional geometry.

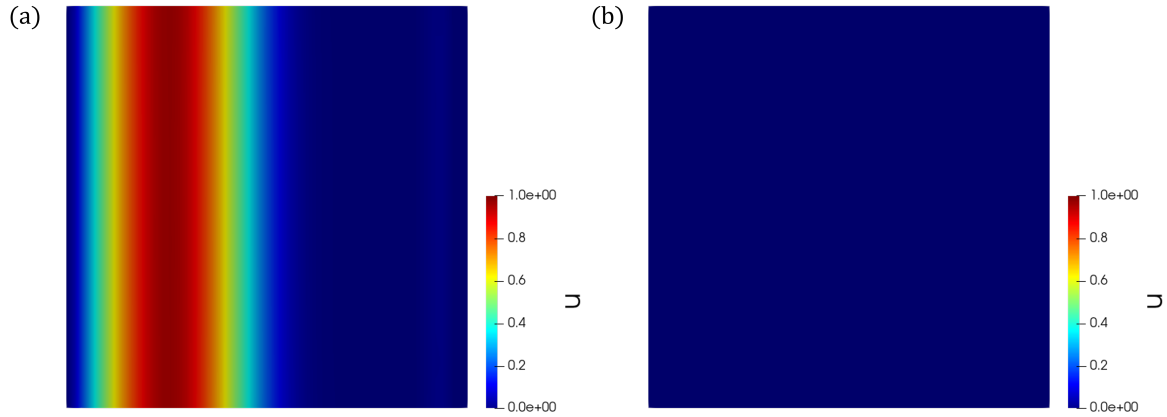


Figure 3.11 Snapshots of the acoustic velocity field at $t = \frac{T}{2}$ (a) and $t = T$ (b) for the controlled two-dimensional system.

Figure 3.12 illustrates two snapshots of the acoustic velocity at the plane $z = \frac{L}{2}$ at times $t = \frac{T}{2}$ and $t = T$ for the uncontrolled case. Figure 3.13 shows the convergence of the optimization process. Figure 3.14 shows the flow evolution of the controlled case. The results are equivalent to the ones obtained for the one-dimensional and two-dimensional cases.

Finally, we collect the results for the different cases in Figure 3.15. In summary, the three optimal solutions are very similar to each other, which is consistent to the one-dimensional physics of the problem. The optimal control remains still until $t = \frac{T}{2}$ and from this time, it generates a negative acoustic velocity wave that cancels the acoustic wave generated by the prescribed forcing. In all the cases, the control is able to minimize the acoustic kinetic energy at the final time. It is also worth mentioning the fast convergence of the algorithm in all the cases.

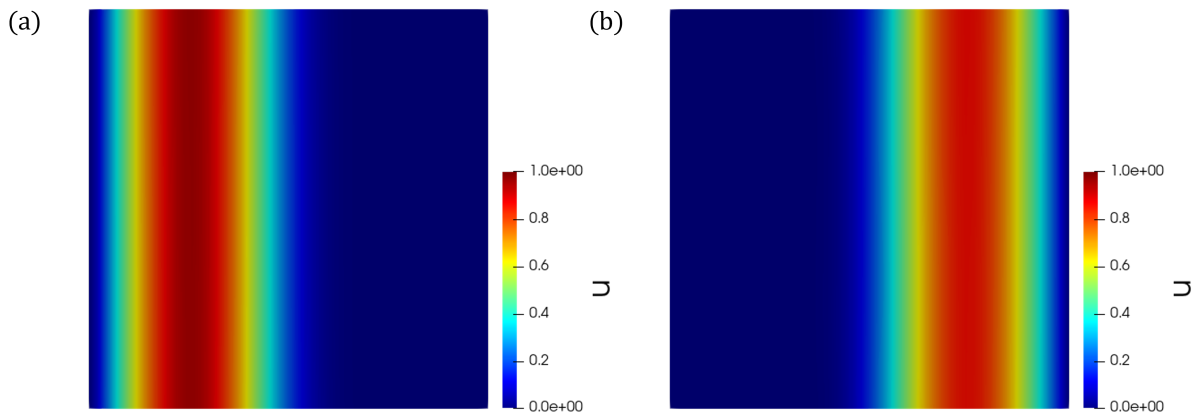


Figure 3.12 Snapshots of the acoustic velocity field at $t = \frac{T}{2}$ (a) and $t = T$ (b) for the uncontrolled three-dimensional system.

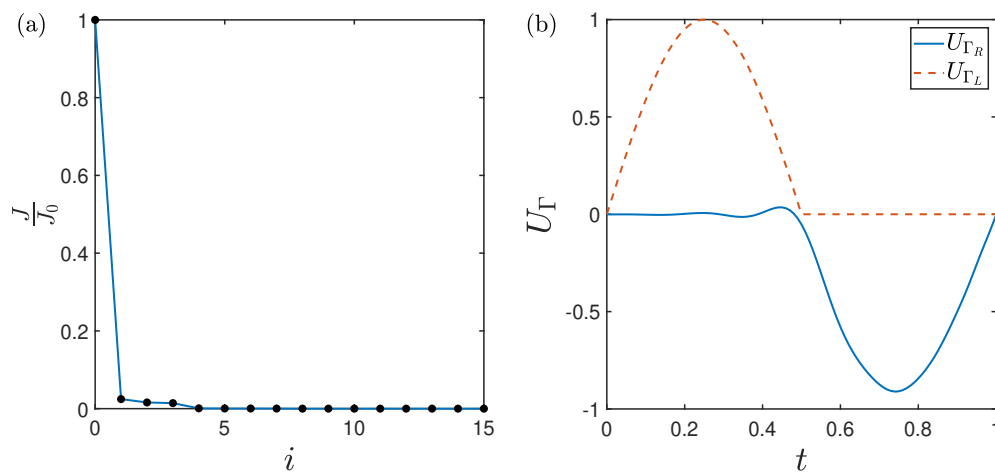


Figure 3.13 Convergence of the cost function (a) and optimal control (b) for the three-dimensional geometry.

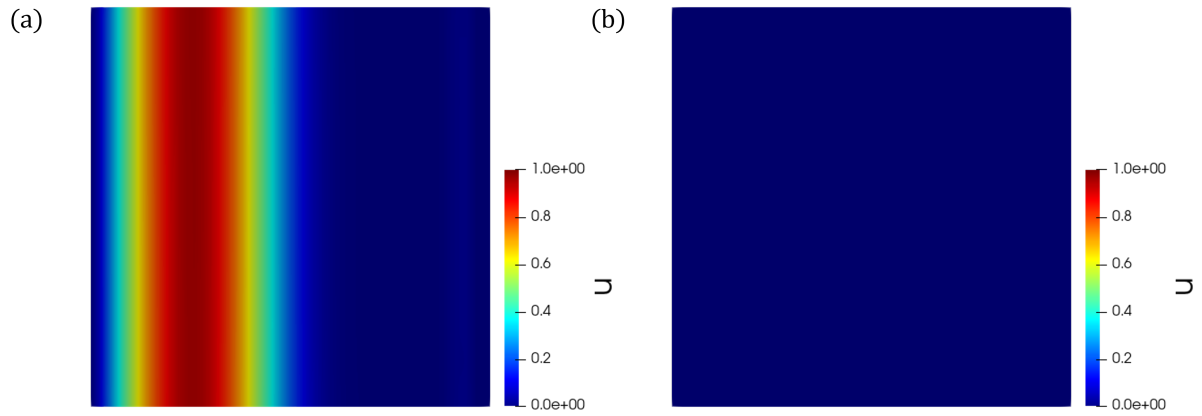


Figure 3.14 Snapshots of the acoustic velocity field at $t = \frac{T}{2}$ (a) and $t = T$ (b) for the controlled three-dimensional system.

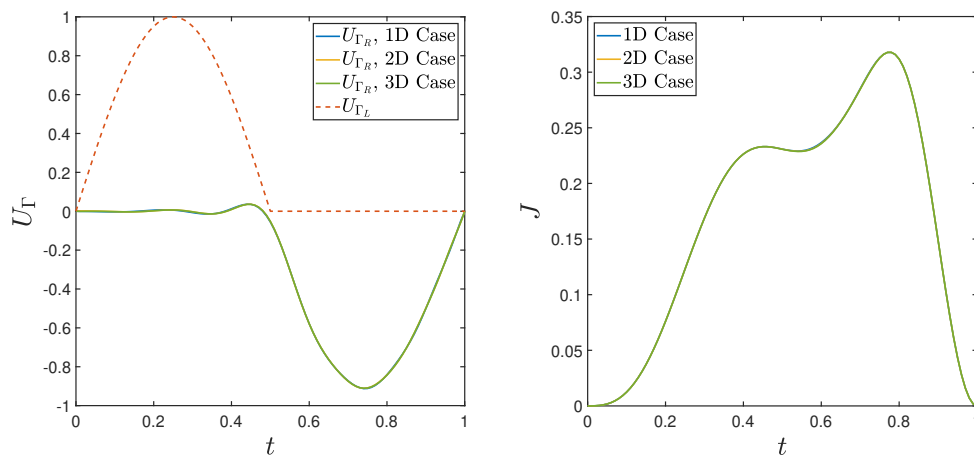


Figure 3.15 Comparison of the optimal control (a) and the evolution of the cost function for the three geometries considered.

Chapter 4

Conclusions and research plan

4.1 Key results

In this report, we have established a methodology to solve optimal control problems in inkjet printing applications using the adjoint method.

First, we have derived the thermoviscous acoustic equations from the Navier-Stokes equations. These equations govern the evolution of acoustic perturbations inside the microchannels. Second, we have formulated a procedure to compute sensitivities using the adjoint method.

We have also studied numerical methods to solve the problem efficiently. First, we have shown why iterative methods need many iterations to converge (or do not converge) if they are applied naively to the linear matrix system that arises from the discretization of the PDE. Then, we have introduced an operator, the Riesz map, that is able to make the problem well-defined. The discretized version of this map is the so-called preconditioner. We have tested these ideas in two problems - the standard Poisson equation and the mixed Poisson equation. The results revealed that preconditioners based on the Riesz map are independent of the mesh size. Moreover, the choice of a preconditioner depends on the equations to solve.

Finally, the optimization method has been applied to simplified versions of the thermoviscous acoustic equations. The first application has been a mass-spring-damper system. The sensitivity has been computed analytically using the adjoint method and the result has been checked by computing the gradient directly. As a second application, we have considered a generic one-dimensional acoustic problem. This system differs from the governing equations of the inkjet printhead in that the simplified problem neglects the thermal dissipation of the fluid, and the flow variables are one-dimensional. The sensitivity has been computed numerically and verified using the Taylor test. The test ensured first-order accuracy of the computed gradient. We have also been able to minimize the acoustic energy at a given time.

4.2 Research plan

The following lines summarizes the future work in this project ¹:

Three-dimensional generic inkjet printhead

As a first optimization problem, we will find the optimal actuator waveform that minimizes the acoustic energy at a given time in an inkjet printhead geometry. This system has two main differences with respect to the viscous acoustic flow presented in Chapter 3. The first one is that the governing equations are different. The flow is not one-dimensional and we also need to include the thermal dissipation. The second difference is that we also need to consider the flow inside the nozzle. This is typically modelled as a boundary condition in the boundary defined by the interface between the nozzle and the channel. The results will be compared against numerical simulations and experiments from the industrial partner *Xaar*.

The following task will be to consider a realistic initial state. In the previous work, the initial condition is assumed to be a quiescent flow. We plan to perform experiments to find the flow field inside the inkjet printhead after a droplet ejection. Moreover, the speed of propagation of information is assumed to be constant. In real inkjet printheads, this parameter is spatial dependent.

We may also consider different optimization problems, such as finding the optimal shape of the channel or nozzle that damps the acoustic waves passively.

Part of this work will be presented in the 75th Annual Meeting of the APS Division of Fluid Dynamics.

Data-assimilation

A future objective is to make our physics-based model quantitatively more accurate. Our physics-based model contains assumptions in the parameters. The most important ones are some fluid parameters (for example, the sound speed) and the flow inside the nozzle.

Therefore, we aim to assimilate experimental data with Bayesian inference to obtain a better approximation of the model parameters and their uncertainties.

¹This research plan will be reassessed periodically as make advances in the project.

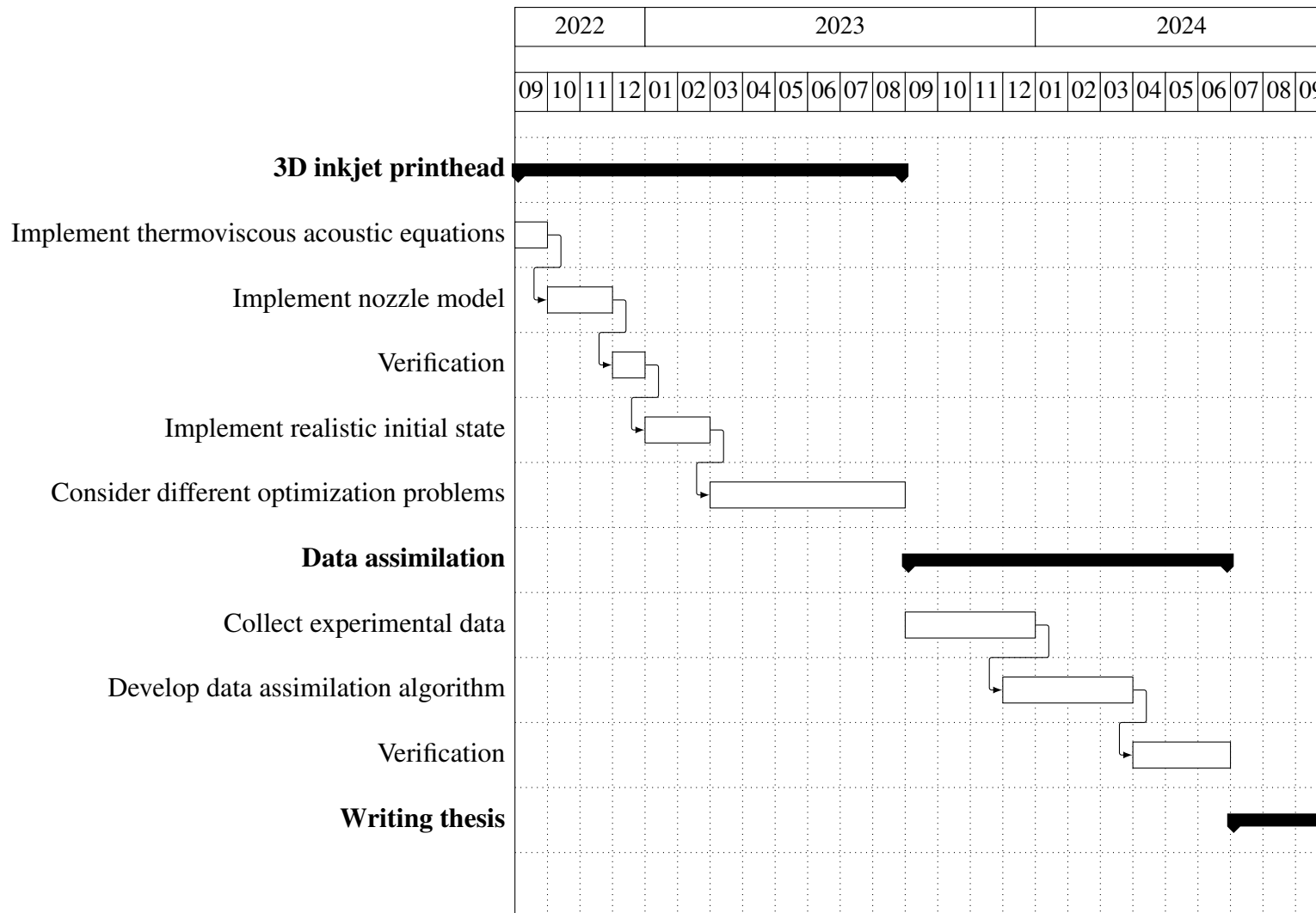


Figure 4.1 Gantt chart

Bibliography

- Barth, T. J., Griebel, M., Keyes, D. E., Nieminen, R. M., Roose, D., and Schlick, T. (2005). *Lecture Notes in Computational Science and Engineering*, volume 49. Springer.
- Beltman, W. M. (1998). Viscothermal wave propagation including acousto-elastic interaction. *PhD Thesis*.
- Bramble, J. H. (2019). *Multigrid methods*. Chapman and Hall/CRC.
- Caeiro, F., Sovardi, C., Förner, K., and Polifke, W. (2017). Shape optimization of a helmholtz resonator using an adjoint method. *International Journal of Spray and Combustion Dynamics*, 9(4):394–408.
- Chiolerio, A., Bocchini, S., and Porro, S. (2014). Inkjet printed negative supercapacitors: synthesis of polyaniline-based inks, doping agent effect, and advanced electronic devices applications. *Advanced Functional Materials*, 24(22):3375–3383.
- Chu, B.-T. (1965). On the energy transfer to small disturbances in fluid flow (part i). *Acta Mechanica*, 1(3):215–234.
- Crank, J. and Nicolson, P. (1947). A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Mathematical Proceedings of the Cambridge Philosophical Society*, 43(1):50–67.
- Culick, F., Kuentzmann, P., RESEARCH, N., and (France), T. O. N.-S.-S. (2006). *Unsteady Motions in Combustion Chambers for Propulsion Systems*. AD-a466 461. nato research and technology organization neuilly-sur-seine (france).
- Dijksman, J. F. (2019). *Design of Piezo Inkjet Print Heads: From Acoustics to Applications*. John Wiley & Sons.
- Gan, H. Y., Shan, X., Eriksson, T., Lok, B. K., and Lam, Y. C. (2009). Reduction of droplet volume by controlling actuating waveforms in inkjet printing for micro-pattern formation. *Journal of Micromechanics and Microengineering*, 19(5):055010.
- Hackbusch, W. (1994). *Iterative solution of large sparse systems of equations*, volume 95. Springer.
- Hestenes, M. R. and Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *Journal of research of the National Bureau of Standards*, 49:409–436.
- Hoath, S. D. (2016). *Fundamentals of inkjet printing: the science of inkjet and droplets*. John Wiley & Sons.

- Hoffman, K. and Kunze, R. A. (2004). *Linear Algebra*. PHI Learning, second edition.
- Jameson, A. *Optimum aerodynamic design using CFD and control theory*.
- Jameson, A. Optimum aerodynamic design using control theory, computational fluid dynamics review 1995, edited by m. hafez and k. oshima (1995).
- Jameson, A. (1988). Aerodynamic design via control theory. *Journal of scientific computing*, 3(3):233–260.
- Jameson, A., Martinelli, L., and Pierce, N. A. (1998). Optimum aerodynamic design using the navier–stokes equations. *Theoretical and computational fluid dynamics*, 10(1):213–237.
- Jameson, A. and Reuther, J. (1994). Control theory based airfoil design using the euler equations. In *5th Symposium on Multidisciplinary Analysis and Optimization*, page 4272.
- Kim, B.-H., Lee, H.-S., Kim, S.-W., Kang, P., and Park, Y.-S. (2014a). Hydrodynamic responses of a piezoelectric driven mems inkjet print-head. *Sensors and Actuators A: Physical*, 210:131–140.
- Kim, J., Bodony, D. J., and Freund, J. B. (2014b). Adjoint-based control of loud events in a turbulent jet. *Journal of Fluid Mechanics*, 741:28–59.
- Kirby, R. C. (2010). From functional analysis to iterative methods. *SIAM Review*, 52(2):269–293.
- Kungurtsev, P. V. and Juniper, M. P. (2019). Adjoint-based shape optimization of the microchannels in an inkjet printhead. *Journal of Fluid Mechanics*, 871:113–138.
- Langtangen, H. P. and Mardal, K.-A. (2019). *Introduction to numerical methods for variational problems*, volume 21. Springer Nature.
- Lee, B. and Min, C. (2021). Optimal preconditioners on solving the poisson equation with neumann boundary conditions. *Journal of Computational Physics*, 433:110189.
- Lee, C.-L., Chang, K.-C., and Syu, C.-M. (2011). Silver nanoplates as inkjet ink particles for metallization at a low baking temperature of 100°C. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 381(1):85–91.
- Lohse, D. (2022). Fundamental fluid dynamics challenges in inkjet printing. *Annual Review of Fluid Mechanics*, 54(1):349–382.
- Lorber, B., Hsiao, W.-K., Hutchings, I. M., and Martin, K. R. (2013). Adult rat retinal ganglion cells and glia can be printed by piezoelectric inkjet printing. *Biofabrication*, 6(1):015001.
- Magri, L. (2019). Adjoint methods as design tools in thermoacoustics. *Applied mechanics reviews*, 71(2).
- Mardal, K.-A. and Winther, R. (2011). Preconditioning discretizations of systems of partial differential equations. *Numerical Linear Algebra with Applications*, 18(1):1–40.

- McKinley, Gareth H. and Renardy, M. (2011). Wolfgang von ohnesorge. *Physics of Fluids*, 23(12):127101.
- Miers, J. C. and Zhou, W. (2017). Droplet formation at megahertz frequency. *AIChE Journal*, 63(6):2367–2377.
- Mitter, S. and Lions, J. (1971). *Optimal Control of Systems Governed by Partial Differential Equations*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg.
- Pironneau, O. (1974). On optimum design in fluid mechanics. *Journal of Fluid Mechanics*, 64(1):97–110.
- Plateau, J. (1873). Experimental and theoretical statics of liquids subject to molecular forces only. *Gauthier-Villars, Paris*, 4.
- Rayleigh, L. (1878). On the instability of jets. *Proceedings of the London mathematical society*, 1(1):4–13.
- Reuther, J., Jameson, A., Farmer, J., Martinelli, L., and Saunders, D. (1996). Aerodynamic shape optimization of complex aircraft configurations via an adjoint formulation. In *34th aerospace sciences meeting and exhibit*, page 94.
- Saad, Y. and Schultz, M. H. (1986). Gmres: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7(3):856–869.
- Tomaszewski, G. and Potencki, J. (2017). Drops forming in inkjet printing of flexible electronic circuits. *Circuit World*.
- Trottenberg, U., Oosterlee, C. W., and Schuller, A. (2000). *Multigrid*. Elsevier.

