SHAPE DESIGN BY OPTIMAL FLOW CONTROL AND REDUCED BASIS TECHNIQUES:
APPLICATIONS TO BYPASS CONFIGURATIONS IN HAEMODYNAMICS

THÈSE N° 3400 (2005)

PRÉSENTÉE À LA FACULTÉ SCIENCES DE BASE
Institut d'analyse et de calcul scientifique
SECTION DE MATHÉMATIQUES

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Gianluigi ROZZA

Laurea in Ingegneria Aerospaziale, Politecnico di Milano, Italie
et de nationalité italienne

acceptée sur proposition du jury:

Prof. A. Quarteroni, directeur de thèse
Prof. V. Agoshkov, rapporteur
Prof. A. Patera, rapporteur
Prof. J. Rappaz, rapporteur

Lausanne, EPFL
2005
Ai miei Nonni,
Rosa e Luigi, Giuseppina ed Ernesto
Abstract

The purpose of this thesis is to develop numerical methods for optimization, control and shape design in computational fluid dynamics, more precisely in haemodynamics. The application studied is related with the shape optimization of an aorto-coronaric bypass. The optimization process has to keep into account aspects which are very different and sometimes conflicting, for this reason the process has been organized in more levels dealing with a geometrical scale. Moreover we have chosen to use simplified low fidelity models during the application of the complex optimization tools and to verify in feed-back with higher fidelity models the configurations previously obtained. In our case we deal with fluid models based on Stokes and Navier-Stokes equations, for lower and higher fidelity approach respectively, also in the unsteady formulation.

At an outer level of the optimization process, efficient numerical methods based on parametrized partial differential equations have been developed to get real-time and accurate information concerning the preliminary configurations, and to get a sensitivity analysis on geometrical quantities of interest and on functionals, related with fluid mechanics quantities. This approach is carried out by reduced basis methods which let us rebuild approximate solutions for parametrized equations by other solutions already computed and stored, allowing huge computational savings.

At an inner level we have developed local shape optimization methods by optimal flow control theory based on adjoint approach. Two different approaches have been developed: the former is based on the local displacement of each node on the boundary, the latter is based on small perturbation theory into a reference domain. This approach is more complex but let us avoid mesh reconstruction at each iteration and study the problem into a deeper context from a theoretical point of view and do a generalization dealing with unsteady flows.
Version abrégée

Le but de cette thèse est le développement de méthodes numériques pour l’optimisation, le contrôle et le project de forme dans le cadre de la dynamique des fluides, notamment en hémodynamique. On étudie en particulier une application de ces méthodes à l’optimisation d’un bypass aortocoronaire.

Le processus d’optimisation est basé sur des aspects différents et des exigences parfois contradictoires; pour cette raison il a été structuré sur plusieurs niveaux correspondants à différentes échelles géométriques. En outre, on a choisi d’utiliser des modèles simplifiés pour l’application des complexes outils d’optimisation et de vérifier les états optimaux à posteriori par des modèles plus précis. Dans notre cas, il s’agit de modèles portant respectivement sur les équations de Stokes et Navier-Stokes, éventuellement non stationnaires.

Dans un niveau extérieur (par rapport au processus d’optimisation) on a développé des méthodes numériques efficaces basées sur des équations paramétrisées qui donnent des informations précises “en temps réel” sur une configuration indicative à considérer à travers l’analyse de sensibilité pour les quantités d’intérêt et les fonctionnelles typiques de la dynamique des fluides. Cette approche utilise les méthodes dites “à bases réduites”, qui permettent de reconstruire des solutions approchées d’équations paramétrisées en fonction d’autres solutions déjà calculées et disponibles, ce qui réduit énormément les ressources de calcul nécessaires.

Dans un niveau interne on a développé des méthodes d’optimisation de forme locale par la théorie du contrôle optimal des fluxes en utilisant la formulation basée sur l’état adjoint. On a étudié deux approches: une première portant sur le déplacement des noeuds du bord, et une deuxième qui utilise une technique de perturbations sur un domaine de référence. Cette dernière approche, même si plus complexe, a l’avantage d’éviter la reconstruction du maillage et permet d’étudier le problème d’une façon plus profonde, ainsi que de le généraliser.
Estratto

Acknowledgements

First of all, I would like to thank Prof. Alfio Quarteroni for being more than a thesis supervisor during these years at EPFL with his human qualities and kindness. He gave me unique opportunities to study and learn new branches of applied mathematics and numerical analysis since my master degree thesis, leading enthusiastically my research activity and always providing new ideas, suggestions and insights.

I am very grateful to Prof. Anthony T. Patera, who has given me the opportunity to visit his research team at Massachusetts Institute of Technology and opened the doors of the reduced basis world. I am particularly grateful for his availability, kindness, patience during the several video conferences between EPFL and MIT and in revising my work.

Special thanks to Prof. Valery Agoshkov of Russian Academy of Sciences for his strong contributions and guidelines on optimal control and adjoint approach during his visiting periods at EPFL.

I thank all the members of the jury, especially Prof. Jacques Rappaz, director of IACS, Institute of Analysis and Scientific Computing, and Prof. Charles Stuart for presiding the jury.

Several people contributed to the achievement of this work. Special recognition goes to all my colleagues of the Chair of Modelling and Scientific Computing (CMCS) of EPFL, past and present members, for sharing with me a part of their precious time, for their friendship and their availability and for creating a lively and nice place and a great team. Special thanks to our information technology team and my office mates.

Thanks to M.me J. Mosetti and M.me C. Craman, our secretaries, for their work and their kindness, and to all other people with whom I interacted during my period at EPFL, especially all the members and colleagues of IACS.

I mention also other three professional environments in which I have been involved: members of MOX -Modelling and Scientific Computing Center– at Politecnico di Milano, especially Prof. L. Formaggia, Prof. F. Saleri, Prof. A. Veneziani and Dr. S. Perotto, with whom I shared many parts of my professional life and I had always suggestions and encouragement; the Haemodel project group to be not only a network but also a big family after several
meetings and conferences; and all the A.T. Patera's group at Fluids Laboratory at MIT, especially Dr. Karen Veroy and Dr. Martin Grepl for creating a very exciting place to work and collaborate.

This work has also benefited from the collaboration with Luca Dedè and Annalisa Quaini who carried out their master thesis under my supervision.

Acknowledgements to Bernoulli Center of EPFL for the special semester organized on the mathematical modelling of the cardiovascular system in 2003.

I acknowledge the European Community’s Human Potential Programme (contract HPRN-CT-2002-00270), the European Union Research Training Network –Haemodel Project– and the OFES –Swiss Federal Office for Education and Schools– which supported my activity during these three years at EPFL.

Personally, I would like to remember all people with whom I interacted during my PhD –the list would be too long and incomplete– for understanding, support, interest and affection, especially all my friends.

Last but not least, a very special thanks to my parents, for their support, their love and patience. This work is dedicated to my grandparents to remember everything they made for our family.

Lausanne, November 2005
“In motu igitur sanguinis explicando easdem offendimus insuperabiles difficultates, quae nos impediant omnia plane opera Creatoris accuratius prescrutari; ubi perpetuo multo magis sum-man sapientiam cum omnipotentia coniunctam admirari ac venerari debemus, cum ne sumum quidem ingenium humanum vel levissimae vibrillae veram structuram perciere atque explicare valeat.”
Leonhard Euler (1707-1783).

“Mathematics seems to endow one with something like a new sense.”
Charles Darwin (1809-1882).

“È nelle difficoltà che si prova la fedeltà e la costanza
– Omnia possum in Eo qui me confortat (St.Paul, Philipp. 4,13) –.”
Francesca Cabrini (1850-1917).
# Contents

1 Introduction and Motivation 1
   1.1 Bypass clinical aspects: surgery, materials, state of the art ................. 1
   1.2 Framework: mathematical modelling in haemodynamics ........................ 4
       1.2.1 On control approaches ........................................ 5
       1.2.2 Reduced basis methods for pre-process and real time optimization .... 6
       1.2.3 Thesis outline ............................................... 8

I Optimal Flow Control and Shape Optimization 13

2 Optimal Control and Shape Optimization by Local Boundary Variations 15
   2.1 A control approach: looking for cost functional optimization ................ 15
   2.2 Design control quantities ........................................ 16
   2.3 Mathematical modelling of the problem. .................................... 18
       2.3.1 The geometrical model ...................................... 18
       2.3.2 The state problem: Navier-Stokes and Stokes equations ............. 19
       2.3.3 The observation on the system ................................ 23
       2.3.4 The adjoint problem ....................................... 25
   2.4 Optimization and control algorithm ...................................... 26
       2.4.1 Gradient method ........................................... 26
       2.4.2 Shape variation and control ................................ 27
       2.4.3 The iterative process ...................................... 28
       2.4.4 Mesh strategy ............................................. 29
   2.5 An equivalent interpretation of the problem .................................. 31
   2.6 Some numerical results ............................................. 33
   2.7 Feedback procedures and results interpretation ............................. 35

3 Shape Design by Optimal Control using Small Perturbation Theory 39
   3.1 An alternative approach to local boundary variations ....................... 39
   3.2 Problem statement ................................................ 40
   3.3 The problem for the perturbation functions ................................ 42
   3.4 The shape optimization problem ..................................... 45
### 3. Variational Equations of the Optimal Control Problem
- 3.5 The variational equations of the optimal control problem
- 3.6 Existence and uniqueness results
- 3.7 Iterative processes
- 3.8 Test problem and numerical results
- 3.9 A comparison between the two methods
- 3.10 Further developments

### 4. Shape Design by Optimal Control with Unsteady Flows and Perturbation Method
- 4.1 Introduction
- 4.2 The non-stationary problem
- 4.3 The unsteady problem for the perturbed functions
- 4.4 The shape optimization problem: unsteady cost functionals
- 4.5 The optimal control formulation for unsteady flows
- 4.6 Existence and uniqueness results for unsteady case
- 4.7 The iterative process in the non-stationary case
- 4.8 Numerical results for unsteady generalized Stokes problem
- 4.9 Conclusions about Part I

### II. Reduced Basis Techniques for Optimization

#### 5. Reduced Basis Methods for Stokes Equations
- 5.1 Reduced basis for viscous flows
- 5.2 The parametrized Stokes problem
- 5.3 The reduced basis formulation of the Stokes equations
  - 5.3.1 Outputs of interest
  - 5.3.2 Reduced basis on-line complexity
- 5.4 Off-line optimized basis assembling: adaptive procedure
  - 5.4.1 $L^2$ pressure error projection
  - 5.4.2 $H^1$ velocity error projection
- 5.5 The computation of the constant $\beta$ of the inf-sup condition
- 5.6 Some preliminary numerical results
  - 5.6.1 First Test: homogeneous Dirichlet boundary conditions, forced flow, 3(1) varying parameter(s)
  - 5.6.2 Second Test: mixed Dirichlet/Neuman boundary conditions, 5(2) varying parameters
  - 5.6.3 Third Test: homogeneous Dirichlet boundary conditions, forced flow, 6(3) varying parameters
  - 5.6.4 On computational costs
- 5.7 On algebraic and approximation stability
  - 5.7.1 Orthonormalization: Gram-Schmidt (GS) algorithm
  - 5.7.2 Approximation stability: other supremizer options
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.8 Some preliminary results on bypass configurations</td>
<td>117</td>
</tr>
<tr>
<td>5.9 Non-compliant outputs: dual residual approach</td>
<td>119</td>
</tr>
<tr>
<td>5.10 Reduced basis approximation for adjoint problem</td>
<td>120</td>
</tr>
<tr>
<td>5.10.1 Some elements for output error estimation</td>
<td>122</td>
</tr>
<tr>
<td>5.10.2 Results using non-compliant outputs</td>
<td>123</td>
</tr>
<tr>
<td>5.11 Conclusions</td>
<td>126</td>
</tr>
<tr>
<td>6 Reduced Basis Methods for Stokes Equations in Domains with Non-Affine</td>
<td>127</td>
</tr>
<tr>
<td>Parametric Dependence</td>
<td></td>
</tr>
<tr>
<td>6.1 An extension on the use of reduced basis</td>
<td>127</td>
</tr>
<tr>
<td>6.2 Empirical interpolation for (coefficient) functions approximation</td>
<td>128</td>
</tr>
<tr>
<td>6.3 The Stokes problem with non-affine parametric dependence</td>
<td>129</td>
</tr>
<tr>
<td>6.4 The Stokes reduced basis formulation with non-affine parameters</td>
<td>132</td>
</tr>
<tr>
<td>6.5 Numerical results</td>
<td>134</td>
</tr>
<tr>
<td>6.5.1 First test: curved upper wall</td>
<td>134</td>
</tr>
<tr>
<td>6.5.2 Second test: bypass with curved incoming branch</td>
<td>139</td>
</tr>
<tr>
<td>6.6 Conclusions</td>
<td>146</td>
</tr>
<tr>
<td>7 Reduced Basis for Navier-Stokes Equations</td>
<td>147</td>
</tr>
<tr>
<td>7.1 Introduction</td>
<td>147</td>
</tr>
<tr>
<td>7.2 The state of the art</td>
<td>148</td>
</tr>
<tr>
<td>7.3 Steady Navier-Stokes equations</td>
<td>149</td>
</tr>
<tr>
<td>7.4 Parametrized formulation: affine parametric dependence</td>
<td>150</td>
</tr>
<tr>
<td>7.5 Parametrized formulation: non-affine parametric dependence</td>
<td>154</td>
</tr>
<tr>
<td>7.6 Parametrized formulation: affine and non-affine</td>
<td></td>
</tr>
<tr>
<td>parametric dependence</td>
<td>156</td>
</tr>
<tr>
<td>7.7 Some preliminary results</td>
<td>162</td>
</tr>
<tr>
<td>7.8 Furrowed channel test</td>
<td>162</td>
</tr>
<tr>
<td>7.8.1 Results for low Reynolds number</td>
<td>166</td>
</tr>
<tr>
<td>7.8.2 Results for higher Reynolds number</td>
<td>166</td>
</tr>
<tr>
<td>7.8.3 Reduced basis computational costs</td>
<td>167</td>
</tr>
<tr>
<td>7.9 The bypass problem</td>
<td>168</td>
</tr>
<tr>
<td>7.9.1 Outputs sensitivities</td>
<td>170</td>
</tr>
<tr>
<td>7.10 Conclusions</td>
<td>172</td>
</tr>
<tr>
<td>8 Perspectives in Flow Control by Reduced Basis Methods</td>
<td>173</td>
</tr>
<tr>
<td>8.1 Introduction</td>
<td>173</td>
</tr>
<tr>
<td>8.2 Optimal control problem for advection-diffusion equations</td>
<td>174</td>
</tr>
<tr>
<td>8.2.1 A formulation for optimal control problem based on Lagrangian</td>
<td>174</td>
</tr>
<tr>
<td>8.2.2 Optimal control problem governed by advection-diffusion equations</td>
<td>175</td>
</tr>
<tr>
<td>8.2.3 Numerical discretization and stabilization</td>
<td>177</td>
</tr>
<tr>
<td>8.3 Reduced basis method for optimal control</td>
<td>178</td>
</tr>
</tbody>
</table>
CONTENTS

8.3.1 Computational procedure: off-line/on-line decomposition ........................................ 180
8.3.2 Error on control and error on cost functional .......................................................... 180
8.4 An application to flow control ...................................................................................... 181
8.5 Control input: variable emission rates ........................................................................... 183
  8.5.1 Parametrized state equation ...................................................................................... 183
  8.5.2 Parametrized adjoint equation .................................................................................. 184
  8.5.3 Some results ............................................................................................................. 185
  8.5.4 Example .................................................................................................................. 186
8.6 Physical input: variable emission rates and velocity field ............................................ 187
  8.6.1 Parametrized state equation ...................................................................................... 187
  8.6.2 Parametrized adjoint equation .................................................................................. 188
  8.6.3 Some results ............................................................................................................. 189
  8.6.4 Example .................................................................................................................. 190
8.7 Geometrical input: parametrized domains ..................................................................... 191
  8.7.1 Parametrized state equation ...................................................................................... 191
  8.7.2 Parametrized adjoint equation .................................................................................. 194
  8.7.3 Geometrical sensitivity analysis ................................................................................ 195
  8.7.4 Sensitivity analysis applied to control problem ........................................................ 196
8.8 Some concluding remarks .............................................................................................. 197
8.9 Appendix: reduced basis for time-dependent problem .................................................. 198
  8.9.1 Reduced basis formulation ....................................................................................... 200

A An Introduction to Reduced Basis Methods: Elliptic Equations in Parametrized Domains 205
A.1 Overview on reduced basis .............................................................................................. 205
A.2 Problem formulation ...................................................................................................... 206
  A.2.1 Galerkin approximation ............................................................................................ 207
  A.2.2 Reduced basis method .............................................................................................. 209
A.3 Reduced basis output bounds ........................................................................................ 211
  A.3.1 Bound conditioner .................................................................................................... 212
  A.3.2 Error and output bounds .......................................................................................... 213
  A.3.3 Minimum coefficient bound conditioner .................................................................. 213
  A.3.4 Bounding properties ................................................................................................. 214
  A.3.5 Computational procedure: off-line/on-line decomposition .................................... 215
A.4 A priori convergence theory ........................................................................................... 216
  A.4.1 Best approximation ................................................................................................... 217
A.5 Adaptive procedure for basis construction ..................................................................... 218
A.6 A multi-parameter application ........................................................................................ 218
  A.6.1 Geometrical model .................................................................................................... 218
  A.6.2 Formulation .............................................................................................................. 219
  A.6.3 Affine mapping ......................................................................................................... 221
  A.6.4 Model problem ......................................................................................................... 222
A.7  Some numerical results ........................................ 224
A.8  Developments on reduced basis ................................. 231
  A.8.1  Reduced basis for the study of a complete geometrical configuration . 231
  A.8.2  Reduced basis for pre-process optimization ............... 231
  A.8.3  Reduced basis for a shape optimization problem ....... 232
  A.8.4  Reduced basis for optimal flow control .................. 233
A.9  The model-order reduction ..................................... 233
  A.9.1  Earlier Work ............................................... 234
  A.9.2  Remarks on {	extit{a posteriori}} error estimation .......... 237
List of Tables

5.1 $\beta$ and $\beta_N$ in one parameter dependent configuration ($D$ for example). Other parameters are frozen. ................................................................. 99
5.2 Table of $H^1$ and $L^2$ relative errors on velocity and pressure, respectively, 50 test configurations, 3(1) parameters, $N < 7$. ............................................. 101
5.3 Table of $H^1$ and $L^2$ relative errors on velocity and pressure, respectively, 50 test configurations, 3(1) parameters, $N > 7$. ............................................ 102
5.4 $\beta_N$ equiv. LBB inf-sup constant. $\beta = 5.0164$ ($\tau = 0.76$). ....................... 102
5.5 Output error: $\Delta s = s - s_N$ for $\tau = 0.356$. ................................................. 103
5.6 Table of $H^1$ relative errors on velocity and $L^2$ relative errors on pressure, 50 test configurations, 5(2) parameters, $N < 20$. ........................................... 106
5.7 $\Delta s = s - s_N$: max and mean error over output $s$ for 50 configurations. ....... 106
5.8 Table of velocity relative errors $H^1$ and pressure $L^2$, 90 test configurations, 6(3) parameters, $N \leq 25$. ......................................................... 108
5.9 Max and mean for 90 configurations Functional output difference $\Delta s = (s - s_N)$: Max and mean values for 90 configurations $s = \int_{\Omega} f \cdot u$. ......................... 108
5.10 Computational costs (CPU time) varying $N$ (and $H^1$ velocity error) and comparison with computational cost of FEM solution (%). ................................. 109

6.1 Mean $H^1$ velocity relative errors for different $\varepsilon_{\text{max}}$. .......................... 137
6.2 Mean $H^1$ velocity relative errors for different $\varepsilon_{\text{max}}$ and comparison with “true” $g^1$. ................................................................................................. 138
6.3 $H^1$ velocity relative errors for different number of varying parameters and $\varepsilon_{\text{max}} = 10^{-5}$: (a) only $v$ (non-affine) parameter varying, (b) $v, L, S$ (affine and non-affine) parameters (in different subdomains) varying. ...................... 142
6.4 $H^1$ velocity relative errors for different number of varying parameters and $\varepsilon_{\text{max}} = 10^{-5}$: (c) $v, L, S, t$ parameters varying, (d) $v, t, L, S, \theta$ parameters varying.143
6.5 $L^2$ pressure relative errors for different number of varying parameters and $\varepsilon_{\text{max}} = 10^{-5}$: (a) only $v$ (non-affine) parameter varying, (b) $v, L, S$ (affine and non-affine) parameters (in different subdomains) varying. ...................... 144
6.6 $L^2$ pressure relative errors for different number of varying parameters and $\varepsilon_{\text{max}} = 10^{-5}$: (c) $v, L, S, t$ parameters varying, (d) $v, t, L, S, \theta$ parameters varying.145


8.1 Control input: number of basis functions, mean errors on cost functional and control function at convergence and time saving.  

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1</td>
<td>Control input: number of basis functions, mean errors on cost functional and control function at convergence and time saving.</td>
<td>186</td>
</tr>
</tbody>
</table>

8.2 Control input for state equation: mean $H^1$ error and time saving for control input.  

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.2</td>
<td>Control input for state equation: mean $H^1$ error and time saving for control input.</td>
<td>186</td>
</tr>
</tbody>
</table>

8.3 Physical input: number of basis functions, mean errors on cost functional and control function, at convergence, and time saving (including also error calculation).  

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.3</td>
<td>Physical input: number of basis functions, mean errors on cost functional and control function, at convergence, and time saving (including also error calculation).</td>
<td>189</td>
</tr>
</tbody>
</table>

8.4 Physical input for state equation during optimal control (velocity field is fixed at first iteration): mean $H^1$ error and time saving. Without fixing the online velocity field at first iteration computational saving are $\sim 65 - 70\%$.  

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.4</td>
<td>Physical input for state equation during optimal control (velocity field is fixed at first iteration): mean $H^1$ error and time saving. Without fixing the online velocity field at first iteration computational saving are $\sim 65 - 70%$.</td>
<td>190</td>
</tr>
</tbody>
</table>

A.1 $\Delta_N$ and $N$ using two-parameters configuration ($L$ and $D$) and adaptive procedure (Section A.5). Without adaptive procedure we would need $N = 50$ to get a $\Delta_N < 10^{-3}$, we reach a consistent computational load reduction.  

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1</td>
<td>$\Delta_N$ and $N$ using two-parameters configuration ($L$ and $D$) and adaptive procedure (Section A.5). Without adaptive procedure we would need $N = 50$ to get a $\Delta_N &lt; 10^{-3}$, we reach a consistent computational load reduction.</td>
<td>225</td>
</tr>
</tbody>
</table>

A.2 $\Delta_{N_{\text{max}}}$ and $N$ using five-parameters configuration and adaptivity procedure to get a $\Delta_{N_{\text{max}}} < 10^{-4} = \varepsilon_d^{\text{prior}}$ in the worst case. Note values of the effectivity $\eta_N$.  

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.2</td>
<td>$\Delta_{N_{\text{max}}}$ and $N$ using five-parameters configuration and adaptivity procedure to get a $\Delta_{N_{\text{max}}} &lt; 10^{-4} = \varepsilon_d^{\text{prior}}$ in the worst case. Note values of the effectivity $\eta_N$.</td>
<td>226</td>
</tr>
</tbody>
</table>

A.3 Third step future developments options.  

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.3</td>
<td>Third step future developments options.</td>
<td>233</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Simplified bypass model. ......................................................... 2
1.2 Heart, coronary arteries and bypass. ......................................... 2
1.3 Different anastomoses configurations: end-to-side (top) and cuffed on the bed (bottom). ................................................................. 3
1.4 Miller cuffed bypass model. ....................................................... 3
1.5 Bypass schemes: coarse geometrical configuration (left), parametrized by $H$, $L$, $S$, $t$, $D$, $\theta$ and local configuration (right), where $\Gamma_{C,0}$ represents the initial configuration, $\Gamma_C$ the one optimized by small perturbation techniques and $\Gamma_{C+w}$ the one optimized by local boundary variations. ......................................................... 4
1.6 A double level of interest: local shape optimization and complete configuration. ................................................................. 10
1.7 Thesis organization: theoretical and conceptual links. .................. 11
1.8 Thesis organization: application to bypass optimization. ................. 12

2.1 3-D Anastomoses fibreglass model (Aarhus University Hospital, Denmark). ................................................................. 19
2.2 Anastomosis schematic model and symbolic notation used for domain and boundaries, with down-field observation zone $\Omega_{wd}$. ......................... 20
2.3 Vorticity in the down-field zone of a bypass (from a numerical simulation carried out by Fischer, Lee, Loth; University of Illinois and Argonne National Laboratories by permission). ................................................................. 24
2.4 Vorticity in the down-field zone of a bypass: 2D details at the incoming branch (from a numerical simulation carried out by Fischer, Lee, Loth; University of Illinois and Argonne National Laboratories by permission). ................................................................. 24
2.5 Optimal shape design by flow control: scheme of the algorithm. ........ 30
2.6 Complete bypass configuration (bridge) before optimal shape design process: iso-velocity $[ms^{-1} \cdot 10^{-2}]$. ................................................................. 34
2.7 Complete bypass configuration at the end of shape optimization: iso-velocity, same colorbar of the previous picture. ................................................................. 34
2.8 Bypass configuration near the incoming branch before shape optimization (left) and with 11% vorticity reduction (right). ................................................................. 35
2.9 Bypass configuration near the incoming branch with 22% (left) and with 32% vorticity reduction (right). The colorbar is the same of pictures in the previous page. ................................................................. 35
2.10 Bypass configuration near the incoming branch with 38% (left) and with 45% vorticity reduction (right). ........................................... 35
2.11 Fluid dynamics unsteady phenomena into the incoming branch of bypass at $t = 0.3s$, velocity $[ms^{-1} \cdot 10^{-2}]$. ........................................... 36
2.12 Fluid dynamics unsteady phenomena into the incoming branch of bypass at $t = 0.5s$. .......................................................... 36
2.13 Variations in time of the Wall Shear Stress $\Sigma(t)$ $[Nm^{-1}s^{-1}]$ (2.2.7) along $\Gamma_c$ (pulsatory inflow condition) in the two different bypass configurations, at the beginning (continuous line) and at the end (dashed line) of optimal shape design process. ................................. 37
2.14 Bypass Scheme and Green’s Theorem interpretation. .......................... 38

3.1 Idealized, 2-D bypass bridge configuration (left) and detail of the sensible part for the optimization process (right), where the dotted curve represents a possible shape variation. ............................................ 40
3.2 $\Omega = \overline{\Omega}_1 \cup \overline{\Omega}_2, \Gamma_w = \Gamma_{w_1} \cup \Gamma_{w_2} \cup \Gamma_{w_3}, \Gamma_0 = \partial \Omega_1 \cap \partial \Omega_2$. .................................................. 41
3.3 The original “unperturbed domain” $\Omega_0$. ........................................ 42
3.4 The “simple” domain $\bar{\Omega}$. ....................................................... 42
3.5 Subdomain $\Omega_{\text{obs}}$. ................................................................. 47
3.6 Transformation of the simple domain into the unperturbed domain $\Omega_0$ .... 48
3.7 Domain $\Omega$ with $N$ shape functions: (a) $N = 1, f = \beta_1 + a\varphi_0(x), \varphi_0 = x(x_2 - x)$; (b) $N = 3, f = \beta_1 + \sum_{i=1}^{3} a_i \varphi_i$. .................................................. 56
3.8 Example of configuration proposed and studied by Lee, Loth and Fisher [77]. 57
3.9 Idealized 2-D bypass configuration before optimal shape design process: iso-velocity $[ms^{-1} \cdot 10^{-2}]$. .................................................. 57
3.10 Bypass configuration at the end of shape optimization using first corrections: iso-velocity $[ms^{-1} \cdot 10^{-2}]$. ........................................... 58
3.11 First correction $\frac{\partial y}{\partial \Gamma_y}$, quantity which is related with the adjoint solution $(q, \sigma)$ and represents the shape variation in bypass configuration (reported in the reference domain) $[m \cdot 10^{-2}]$. ........................................... 58
3.12 Bypass configuration (velocity $[ms^{-1} \cdot 10^{-2}]$, same colorbar of Fig. 3.10) near the incoming branch after shape optimization by boundary variations (30 iterations, 40% vorticity reduction). ........................................... 59
3.13 Bypass configuration (velocity $[ms^{-1} \cdot 10^{-2}]$, same colorbar of Fig. 3.10, 25 iterations and 30% vorticity reduction) after design by small perturbations. .... 60
3.14 Bypass configuration and its shape sensitivity in reference domain with respect to the cost functional (quantity related with adjoint solution) to underline the most sensible zone related with observation $[m \cdot 10^{-2}]$. ........................................... 60
3.15 Distributed vorticity $[m^2 s^{-1}]$ reduction achieved during the two optimization processes: shape design by local boundary variations and small perturbations. 61
4.1 Idealized, 2-D bypass bridge configuration (left) and detail of the sensible part for the optimization process (right). The dotted curve represents the portion of the boundary that is subjected to change. ................................. 64
4.2 Main notations: $\Omega = \Omega_1 \cup \Omega_2$, $\Omega_w = \Gamma_{w1} \cup \Gamma_{w2} \cup \Gamma_{w3}$, $\Gamma_0 = \partial \Omega_1 \cap \partial \Omega_2$ .... 64
4.3 The original “unperturbed domain” $\Omega_0$. .................................. 65
4.4 The “simple” domain $\Omega$. .............................................. 65
4.5 Transformation of the simple domain into the unperturbed domain $\Omega_0$. ... 72
4.6 Flow-rate waveform at the inlet of the ascending Aorta from [140]. ......... 80
4.7 Flow-rate waveform at the left coronary from [140]. .......................... 80
4.8 Unsteady (pulsatile) Stokes velocity profiles at the inflow $[m s^{-1}]$. ....... 81
4.9 Horizontal velocity $[m s^{-1} \cdot 10^{-2}]$ at $t = 0.1s$ for initial test configuration (left) and after 5 iterations (right). ............................................ 82
4.10 Horizontal velocity $[m s^{-1} \cdot 10^{-2}]$ after $t = 0.1s$ at 11 (left) and 17 (right) iterations. ................................................................. 82
4.11 Horizontal velocity $[m s^{-1} \cdot 10^{-2}]$ at $t = 0.4s$ for initial test configuration (left) and after 5 iterations (right). ................................. 82
4.12 Horizontal velocity $[m s^{-1} \cdot 10^{-2}]$ after $t = 0.4s$ at 11 (left) and 17 (right) iterations. ................................................................. 83
4.13 Horizontal velocity $[m s^{-1} \cdot 10^{-2}]$ at $t = 0.7s$ for initial test configuration (left) and after 5 iterations (right). ................................. 83
4.14 Horizontal velocity $[m s^{-1} \cdot 10^{-2}]$ after $t = 0.7s$ at 11 (left) and 17 (right) iterations. ................................................................. 83
4.15 Distributed vorticity $[s^{-1} \cdot 10^{-2}]$ in original configuration (left) and at the end of the optimization process (right). Vorticity in the upper corner and in the bed of the artery is diminished. ............................. 84
4.16 Variation of corrections $(x, \delta f^n(x) y/\beta)$ at the first iteration (left) and after 14 iterations of the process (right) $[m \cdot 10^{-2}]$. ................... 84
4.17 Total vorticity $[m^2 s^{-1}]$ reduction at different time steps during shape optimization. ................................................................. 84
4.18 Total vorticity $[m^2 s^{-1}]$ in time during shape optimization for different configurations. ................................................................. 85

5.1 True physical domain: sub-domains and parameters. ....................... 91
5.2 Reference domain $\Omega$. .................................................. 92
5.3 Velocity and pressure solution for $\tau = 0.1$. ................................ 100
5.4 Velocity and pressure solution for $\tau = 1.45$. .............................. 101
5.5 Relative max projected $H^1$ error (see Section 5.4.2) on velocity at each iteration during adaptive basis assembling (left) and parameters distribution during offline reduced basis assembling procedure (right). ............. 101
5.6 Left: relative true $H^1$ velocity errors, max and mean error over a large test sampling (50 configurations). Right: relative true $L^2$ pressure errors, max and mean error over a large test sampling (50 configurations); errors in log-log scale.102
5.7 Left: Functional output difference $\Delta s = (s - s_N)$; $s = \int_{\Omega} f \cdot u$. Right: convergence of $s_N$ to $s$ ($\tau = 0.356$). .......................... 103
5.8 Example of flow solution (velocity $[ms^{-1}.10^{-2}]$) in the “T” shape parametrized domain. .......................... 104
5.9 Basis assembling: parameters distribution during off-line reduced basis optimized assembling procedure (left). Functional output difference $\Delta s = (s - s_N)$; $s = \int_{\Omega} f \cdot u$: max and mean error over large test sample configurations (right). 105
5.10 Left: relative true $H^1$ velocity errors: max and mean error over a large test sampling (50 configurations). Right: relative true $L^2$ pressure errors: max and mean error over a large test sampling (50 configurations); errors in log-log scale.105
5.11 Basis assembling: total $H^1$ velocity (projected) error reduction during basis assembling (left) and parameters distribution during off-line reduced basis assembling (right). .......................... 107
5.12 Left: relative true $H^1$ velocity errors: max and mean error over a large test sampling (90 configurations.). Right: relative true $L^2$ pressure errors: Max and mean error over a large test sampling (90 configurations); errors in log-log scale. .......................... 107
5.13 Basis assembling: total $H^1$ velocity and $L^2$ pressure (projected) errors reduction during basis assembling (left) and corresponding parameters distribution during off-line reduced basis assembling (right). ......................... 108
5.14 Left: relative true $H^1$ velocity errors: max and mean error over a large test sampling (90 configurations). Right: relative true $L^2$ pressure errors: Max and mean error over a large test sampling (90 configurations); errors in log-log scale.109
5.15 Orthonormalization: condition number of reduced basis Stokes linear system matrix with complete orthonormalized basis (left) and with a partial orthonormalization only on velocity and pressure (right), but not on supremizer. .... 111
5.16 New supremizer first option, $H^1$ relative error (for velocity) and $L^2$ relative error (for pressure) using 50 configurations. ...................... 114
5.17 New supremizer second option, $H^1$ relative error (velocity) and $L^2$ relative error (pressure) over 50 configurations. ...................... 116
5.18 Condition number (max and mean) for the Stokes reduced basis system, with the new supremizer second option. On the left results with orthonormalization, on the right without it. ........................................ 116
5.19 Schematic bypass configuration and reference domain. ......................... 118
5.20 Reduced basis convergence results: mean error on velocity and pressure. .. 118
5.21 Output $s [ms^{-1}.10^{-2}]$ versus the parameter $\theta$ at different $N$. .......... 118
5.22 Output $s [ms^{-1}.10^{-2}]$ versus the ratios $\frac{t}{D}$ and $\frac{S}{D}$ at different $N$. ...... 119
5.23 Max output error (comparison with and without correction). ................... 124
5.24 $H^1$ velocity relative error for dual and primal problems and output relative errors. .................................................. 124
5.25 $H^1$ error, upper bound for output error (non-corrected), adjoint-corrected and non-corrected output error. .............................. 125
5.26 Output $s_N [m^2 s^{-1} \cdot 10^{-2}]$ versus $\theta$ parameter (spots dimension) and the ratios $\frac{1}{D}$ (colorbar) and $\frac{S}{D}$. Real-time “clinical” indications by a great number (i.e. hundreds) of bypass configurations (Re= $10^3$). .................................... 125

5.27 Bidimensional version of Figure 5.26. .................................. 126

6.1 Geometrical scheme for the stenosis test problem (left) and velocity absolute value $[ms^{-1}]$ for $\mu = 0.5$ (right). ......................................................... 136

6.2 Horizontal and vertical velocity $[ms^{-1}]$ for $\mu = 0.5$. ................. 137

6.3 $H^1$ relative mean error on velocity for different $\varepsilon_{max}$, imposed on all $g^j_M(x, \mu)$, considering different $\mu$ test values. ................................................. 137

6.4 $H^1$ mean relative error on velocity for different $\varepsilon_{max}$, imposed on all $g^j_M(x, \mu)$ (left) and comparison with the use of “true” $g^j(x, \mu)$ functions, without empirical interpolation (right). .......................................................... 138

6.5 Periodic channel: vertical velocity (top) and horizontal velocity (bottom) $[ms^{-1}]$, same colorbars of Figure 6.1. .......................................................... 139

6.6 Geometrical scheme for the bypass test problem (physical domain and reference one). ................................................................. 140

6.7 Haemodinamic flow: absolute value of velocity $[ms^{-1} \cdot 10^{-2}]$ in bypass subject to curved wall and graft angle ($v = 0.41$ and $\theta = \pi/3$). ................................. 141

6.8 $L^2$ and $H^1$ relative errors for pressure and velocity, respectively (max and mean). 141

6.9 Parameters distribution during off-line optimized basis assembling procedures. 142

7.1 Parameters $(D, t)$ distribution during basis assembling (left) and scheme of configuration (right). .......................................................... 163

7.2 $H^1$ and $L^2$ relative errors on velocity and pressure, first test with 2 varying parameters $(D, t)$. ............................................................. 163

7.3 Geometrical scheme for curved wall test model problem (left) and velocity (absolute value) for $\mu = 0.4$, Re = 100 (right). The colorbar is the same as in Figure 6.1. .......................................................... 164

7.4 Pressure (left) and vertical velocity (right), for $\mu = 0.7$ and Re = 100. .... 164

7.5 Absolute value of velocity (left) and vertical velocity (right), for $\mu = 0.15$ and Re = 100. ................................................................. 165

7.6 $H^1$ and $L^2$ relative mean error on velocity and pressure ($\nu = 0.1$) at different $\varepsilon_{max}$ imposed on all $g^j_M(x, \mu)$ (great number of testing configurations). .... 166

7.7 $H^1$ and $L^2$ relative max error on velocity and pressure ($\nu = 0.1$) at different $\varepsilon_{max}$ imposed on all $g^j_M(x, \mu)$ (great number of testing configuration). ..... 167

7.8 $H^1$ and $L^2$ relative error (max and mean) on velocity and pressure ($\nu = 0.04$) (testing a great number of configurations). .............................................. 167

7.9 Reduced basis online mean computational costs in comparison with finite element simulation (left) and $H^1$ reduced basis velocity error increasing $N$ (right) for channel flow at low Reynolds numbers. ................................. 168
LIST OF FIGURES

7.10 Reduced basis online mean computational costs for channel flow at higher Reynolds numbers. .................................................. 169
7.11 Geometrical scheme for the bypass test problem (physical domain and reference one). ...................................................... 170
7.12 $H^1$ relative errors on velocity with different $\epsilon_{\text{max}}$ interpolation error imposed on all $g^j_M(x, \mu)$ (testing hundreds of different configurations with 5 different parameters varying). ................................................. 170
7.13 $L^2$ relative errors on pressure with different $\epsilon_{\text{max}}$ interpolation error imposed on all $g^j_M(x, \mu)$ (testing hundreds of configurations with 5 different parameters varying). ................................................. 170
7.14 Distributed vorticity $[m^2s^{-1}]$ varying $\frac{\partial}{\partial t}$ and comparison between Stokes ($\ast$) and Navier-Stokes ($\Delta$) flows (left); distributed vorticity $[m^2s^{-1}]$ and curvature (right). ................................................................. 171
8.1 Reference domain for advection-diffusion problem. ......................... 182
8.2 Three different diffusivity coefficients: $\nu_1, \nu_2, \nu_3$. .................... 182
8.3 Control input: initial reduced basis solution (right) and final reduced basis solution (left) of state equation. Substance concentration is in $[\mu g]$. ................................................................. 186
8.4 Control input: comparison between $J(w_N, u_N)$ and $J(w_h, u_h)$ at every iteration (right) and $L^2$-error on control function trend (left). ................................................................. 187
8.5 Physical input: initial reduced basis solution (right) and final reduced basis solution (left) of state equation. Concentration is in $[\mu g]$. ................................................................. 190
8.6 Physical input: comparison between $J(w_N, u_N)$ and $J(w_h, u_h)$ at every iteration (right) and $L^2$ error on control function trend (left). ................................................................. 191
8.7 Scheme for the computational domain: subdomains and parameters. .... 192
8.8 Variations of substance concentration (in $[\mu g]$) over the observation zone when the upper ($C_3$) and lower ($C_2$) source positions change. ................................................................. 196
8.9 Scheme of the iterative process to solve control problem combined with sensitivity analysis. ................................................................. 197
8.10 Basis construction. .................................................................... 202
8.11 Max $H^1$ relative error during basis assembling and optimization. .... 202
8.12 Max and mean $H^1$ relative error on solutions. ............................ 203
8.13 Max and mean $H^1$ relative error on output. ............................... 203

A.1 Scheme for the reduced-basis multi-parameter problem. ............... 219
A.2 Mesh for the reduced-basis problem (normalized domain). ............ 220
A.3 Scheme of parameter-dependent domain $\Omega^1$ undergoing vertical shear and rotation. ................................................................. 223
A.4 Scheme for parameter-dependent domain $\Omega^3$ and reference domain undergoing both stretch and shear. ................................................................. 224
A.5 Solution using FEM-Galerkin method for a 5-parameters configuration (left) and solution using reduced basis method with $N = 60$. ................................................................. 227
A.6 Distributed error $\tilde{e}$ over the domain $\Omega$ between $\tilde{u}$ and $u_N$. 

A.7 Convergence of the difference $s_N - \tilde{s}$ increasing $N$. 

A.8 Error Bounds $s_N^+, s_N^-$ and their convergence to $s(*)$. 

A.9 Effectivity coefficient $\eta_N$ and its bounds $\eta_{\min}(=1)$ and $\eta_{\max}$. 

A.10 Effectivity coefficient $\eta_N$ with lower Bound (unity). Zoom. 

A.11 2-Parameters $(t, D)$ distribution using adaptivity in basis building off-line process. The starting configuration is in the centre of the 2-parameters space. 

A.12 5-parameters distribution using adaptivity for reduced-basis matrix assembling. 

A.13 “Shell” model scheme made up by consecutive steps.
Chapter 1

Introduction and Motivation

The theory of optimal control based on variational systems of equations is applied in order to optimize the shape of the incoming branch of the bypass (the toe) into the coronary. At this level two different options are addressed for shape design: one implements local boundary variations in the computational domain, the other, based on the theory of small perturbations, makes use of a linearized design in a reference domain. As a pre-processing stage, reduced basis methodologies based on parametrized partial differential equations are developed to provide (a) a sensitivity analysis for geometrical quantities of interest in bypass complete configurations and (b) rapid and reliable prediction of integral functional outputs. The aim of this work is (i) to provide design indications for arterial surgery in the perspective of future development for prosthetic bypasses, (ii) to develop multi-level numerical methods for optimization and shape design by optimal control, and (iii) to provide input-output relationships led by models with lower complexity (i.e. simplified flow models) and competitive computational costs to be validated with higher fidelity models in feed-back.

In this introduction multi-level geometrical approaches in the study of aorto-coronaric bypass anastomoses configurations and their optimization are briefly discussed.

1.1 Bypass clinical aspects: surgery, materials, state of the art

The surgical realization of a bypass to overcome a critically stenosed artery is a very common practice in everyday cardiovascular clinic (see Figure 1.1 [126]).

When a coronary artery is affected by a stenosis, the heart muscle cannot be properly oxygenated through blood. Aorto-coronaric anastomosis restores the oxygen amount through a bypass surgery downstream an occlusion (see Figure 1.2).

At present, different kind and shape for aorto-coronaric bypass anastomoses are available and consequently different surgery procedures are used to set up a bypass.

A bypass can be made up either by organic material (e.g. the saphena vein taken from patient’s legs or the mammary artery) or by prosthetic material. The current saphenous bypass solution requires the extraction of saphena vein with possible complications. In this respect,
CHAPTER 1. INTRODUCTION AND MOTIVATION

Figure 1.1: Simplified bypass model.

Figure 1.2: Heart, coronary arteries and bypass.

prosthetic bypasses are less invasive. They may feature very different shape for bypass anastomoses, such as, e.g., cuffed arteriovenous access grafts (Figure 1.3 and 1.4). Different cuffed models are used and analyzed such as Taylor patch [28] and Miller cuff bypass [30], but also standard end-to-side anastomoses at different graft angle [29] or other shaped carbon-fiber prostheses. For a discussion and a comparative approach based on fluid mechanics considerations see Cole et al. [27], [26].

In the cardiovascular system altered flow conditions such as separation, flow reversal, low and oscillatory shear stress areas and abnormal pulse pattern are all recognized as potentially important factors in the development of arterial diseases (White et al. [165], Loth et al. [87]). A detailed understanding of local haemodynamics phenomena and the effect of vascular wall modification on flow patterns can have useful clinical applications especially in surgical procedures or prostheses tailoring (see Steinman et al. [147], Moore et al. [99]). Concerning bypass setting, lots of different aspects and post-surgical complications have to be taken into consideration, among them we mention intimal thickening hyperplasia (near suture lines), which is a narrowing of coronaries, restenosis, surgical injury, long term graft failure. Every year 8% of all patients risk bypass occlusion, after ten years 80% bypasses must be replaced (data from HeartCenterOnLine©). Repeating procedures typically carry a higher risk of
complications. Because of all these different aspects the design and analysis of artificial arterial bypass is a very complex problem. Carbon fiber, bioartificial and collagen cuffed grafts instead of natural saphenous vein can be used in bypass shape design without needing “in loco” reconstruction. In this framework, optimal flow control and shape optimization provide a new interesting approach to the problem, aimed at improving arterial bypass graft on the basis of a better understanding of fluid dynamics aspects involved in the bypass studying. A parametric analysis on inflow conditions, branch angle and upstream geometry for bypass graft, considering also secondary flows due to the particular configuration studied, has been provided recently by Sherwin and Doorly [141].
CHAPTER 1. INTRODUCTION AND MOTIVATION

1.2 Framework: mathematical modelling in haemodynamics

Improvement in the understanding of the genesis of coronary diseases is very important as it allows to reduce surgical and post-surgical failures. It may also suggest new means in bypass surgical procedures with less invasive methods and to devise new shape in bypass configuration (see Perktold et al. [112], O’Brien and McGloughin [107]).

Generally speaking, mathematical modelling and numerical simulation can allow better understanding of phenomena involved in vascular diseases (see Abdoulaev et al. [1], Quarteroni, Tuveri and Veneziani [125], Quarteroni and Formaggia [126], Fung [39] and more recently Gerbeau and Chapelle [42]).

In this work we aim at providing design indications for bypasses surgical procedures by applying numerical optimization methods at two different geometrical levels: the former is based on the local shape of the incoming branch, the latter is devoted to a complete bridge configuration.

At each geometrical level we have developed different approaches based on either low fidelity steady Stokes flows and on high fidelity unsteady Navier-Stokes. The idea is to use complex optimization tools with low fidelity fluid model, then to verify results in feedback on the complete (high fidelity) fluid model.

At a fine geometrical level the background provided by mathematical modelling and numerical simulation has led us to apply the optimal control theory of systems governed by partial differential equations (PDEs) with the aim of optimizing the (local) shape of a simplified bypass model.

At a coarser geometrical level efficient schemes for reduced basis methodology (see Patera et al. [118]) applied to parametrized partial differential equations ($P^2DEs$) are being used to provide useful and quick indications (outputs) for bypass complete configuration in a repetitive design environment as shape design requires. With the reduced basis approach also a sensitivity analysis of the initial configuration and a study of important geometrical quantities in bypass can be obtained. (see [135] for an introduction and [111] for details). Figure 1.5 clarifies our geometrical double-level of interest for bypass design.

Figure 1.5: Bypass schemes: coarse geometrical configuration (left), parametrized by $H, L, S, t, D, \theta$ and local configuration (right), where $\Gamma_{C,0}$ represents the initial configuration, $\tilde{\Gamma}_C$ the one optimized by small perturbation techniques and $\Gamma_{C} + w$ the one optimized by local boundary variations.
The ultimate goal is to build an input-output relationship hold by different models characterized by an increasing degree of complexity (dealing with different fluid models and optimization methods) and focused on different geometrical levels. The outputs of interest are design quantities and fluid mechanics indexes, while the inputs are typically geometrical quantities. The multilevel approach is introduced in Figure 1.6.

1.2.1 On control approaches

At a local geometrical level optimal control of one (or several) aspect of the problem entails the minimization of a cost functional which describes physical quantities involved in the specific problem. Optimization process is carried out by a control function which, depending upon the context, may represent initial or boundary conditions, shape of the domain (by boundary variations), force terms, sources, etc. In computational fluid dynamics two different kind of optimal control problems can be faced.

The former is called Data Control, in which case the control is on different variables of the problem such as initial and boundary conditions, force terms, sources and coefficients. Control function can act either on the boundary or part of it, or on the whole computational domain or part of it. When control is put on the flux we have a flow control problem (Becker [14], Berggren [17], Hinze and Kunish [57], Ravindran and Hou [58], [59], Gunzburger [49], Slawig [142]).

The other field is called Domain Control, since the solution of the system of PDEs is controlled by boundary variations of the domain itself. This field is related with shape optimization and optimal shape design, see e.g. the work by Jameson et al. [66], [67], [68], Pironneau [115], Mohammadi and Pironneau [97], Kawohl et al. [70], Di Césaré [33], in the context of aerodynamics for external flows (wings and airfoils design) and internal flows (nozzles, channels).

The problem we are considering is related both with optimal shape design (see also J.Haslinger and R.A.E. Makinen [52]) and flow control (see Gunzburger [50]) in haemodynamics, involved in the observation of the evolving system and in cost functionals, such as vorticity or wall shear stress. Optimization process is carried out by a control function used as parameter in modelling the shape of the domain.

At this level two control approaches have been used: in the former the control function is used to define directly the boundary shape (local boundary variation method) in the true domain (see [123]); in the latter, the control function is used to define the mapping transformation from the reference domain to the true one. In this case the design problem becomes an optimal control problem on coefficients and the analysis can be based on small perturbation theory in fluid mechanics (see [4]). Theoretical investigation based on perturbation theory analysis and linearized shape design provides results on existence and uniqueness of solution and on the well-posedness of the problem, and permits us to better understand the problem from a theoretical point of view. In this second approach also an unsteady fluid model has been used during control process. In both cases the adjoint approach, proposed for example by J.L. Lions [82], Marchuk [95] and Agoshkov et al. [96], to get cost functionals gradient in problem.
with distributed or boundary control and observation, has been developed. In the functional optimization process a descent gradient-type method (with fixed optimized step size) is used (see Ciarlet [25]). Numerical approximation is based on the Galerkin-Finite Element Method [127]. Algorithms were developed for control and optimization as well as for the coupling between state and adjoint problems.

At the end of these first investigation stages based on optimal design a cuffed bypass is found with a shape which resembles the Taylor arterial patch [26]. A feedback validation procedure has then been implemented by solving the unsteady Navier-Stokes equations in the original configuration as well as in the final configuration obtained after applying the shape optimization processes. The original elements dealing with this part of the work are the use of small perturbation techniques in fluid mechanics combined with optimal flow control, the use of reference domain formulation to avoid mesh reconstruction at each iteration, the setting of a shape optimization problem as an optimal control problem on coefficients, representing coordinates transformations. This approach has been used both for steady and unsteady flows.

1.2.2 Reduced basis methods for pre-process and real time optimization

Especially in the field of optimization or design, where the evaluation of many different possible configurations is required – corresponding to different choices of the design parameters – even for modest-complexity problems, the computational cost is unacceptably high. To more efficiently utilize the existing computational resources, reliable methods that reduce the complexity of the problem while at the same time preserve all relevant information, are becoming very important.

Input-Output relationship

Central to every design, optimization, or control problem is the evaluation of an “input-output” relationship. The set of input parameters $\mu$, which we will collectively denote as “inputs,” identify a particular configuration of the system or of one of its components. These inputs may represent design or decision variables, such as geometry or physical properties – for example, control variables in optimization studies; or characterization variables, such as physical properties, in inverse problems. The output parameters $s(\mu)$, which we will collectively denote as “outputs”, are performance indexes for the particular input $\mu$ – for example stresses, velocity, flow rates. These outputs are typically expressed as functionals of field variables associated with a set of parametrized partial differential equations which describe the physical behavior of the system or its components. Then we are interested in calculating the outputs $s(\mu) = F(\mu)$, for many different inputs/configurations $\mu$ chosen from a parameter space $D \subset \mathbb{R}^P$, where $P$ is the number of input parameters. Here, $F$ encompasses the mathematical description of the physical problem at different levels of complexity.

For the evaluation of $F$ the underlying equations have to be solved. Usually, an analytical solution is not easy to obtain, rather a discretization procedure like the finite-element method, is used; then $F$ is replaced by $F_h$, a discrete form amenable to numerical solution. The
§1.2. MATHEMATICAL MODELLING IN HAEMODYNAMICS

basic premise, is that as the discretization “length” $h \to 0$, then $F_h \to F$, and consequently
$s_h(\mu) \to s(\mu), \forall \mu \in \mathcal{D}$ but as $h \to 0$ the cost of evaluating $F_h$ becomes prohibitive. Especially
in the context of design, control, or parameter identification where “real-time” response or many “input-output” evaluations are required, a balance between computational cost and accuracy/certainty is essential.

Computational method

Identifying the problem in the high dimensionality of the discrete problems, model-order reduction techniques have been developed. The critical observation is that instead of using projection spaces with general approximation properties — like in finite element method— we choose problem-specific approximation spaces and use these for the discretization of the original problem. Using such spaces, we can construct a model that represents with sufficient accuracy the physical problem of interest using a significantly smaller number of degrees of freedom. Depending on the choice of the global approximation spaces many possible reductions are available.

The computational methods developed in this work permit rapid and reliable evaluation of this input-output relationship induced by partial differential equations in the limit of many queries — that is, in the design, optimization, control, and characterization contexts. In designing new methods, certain qualities must be considered:

- **Efficiency** is crucial for the problems in consideration. To achieve efficiency, we shall use the reduced-basis method: a weighted-residual Galerkin-type method, where the solution is projected onto low-dimensional spaces with certain problem-specific approximation properties.

- **Relevance.** Usually in a design or optimization procedure we are not interested in the field solution, but rather in certain design measures such as the drag coefficient in the case of flow past a bluff body, or the average temperature on a surface in the case of heat conduction. The methods developed as part of this work give accurate approximations to these outputs of interest, defined as functional outputs of the field solution.

- **Reliability.** To quantify the error introduced by the reduced-basis method, an error analysis must be invoked on outputs of interest.

In our field of interest, reduced basis approximation provides not only high computational savings, a rapid (real-time) and accurate methodological pre-process to detect the essential feature of the optimization process itself, but also the study of a geometrical sensitivity analysis of a complete bypass configuration. By selecting a limited number of relevant geometrical parameters (bypass diameter $t$, artery diameter $D$, stenosis length $S$, graft angle $\theta$, bypass bridge height $H$, see Figure 1.5) and a moderate number ($N$) of sample parameters

$$\mu_k = \{t_k, D_k, S_k, \theta_k, H_k\}, \ k = 1, \ldots, N,$$

we solve the parametrized equations that govern the physical problem in a reference domain $\Omega$, properly mapped by coordinate transformations [135]. Then we build properly reduced
CHAPTER 1. INTRODUCTION AND MOTIVATION

basis functional approximation spaces for velocity and pressure to guarantee approximation and algebraic stability. For a new sample $\mu_k$ we look for a new solution which is given by a weighted combination of previously computed and stored solutions. Weights are given by the solution of a state problem on the subspace of the reduced basis by a Galerkin projection (see [111]). In this thesis reduced basis methods have been applied especially to Stokes and Navier Stokes problem. A more pedagogical introduction to elliptic operators is presented in the Appendix A in order to set up the ground for an unfamiliar reader for a better understanding when addressing the Stokes and Navier-Stokes problem, while advection-diffusion operators are considered in the conclusive part dealing with perspectives and developments. The most original contribution are concerned with the pressure treatment in Stokes and Navier-Stokes problem, the introduction of geometrical parametrization of domain by affine and non-affine maps, the study of different options to guarantee approximation stability of reduced basis approximation and, finally, basis orthonormalization to achieve algebraic stability for reduced basis. This methodological development highlights the great potential of reduced basis methods in optimal flow control and shape optimization, not only for pre-process optimization. Reduced basis can be seen as methods to solve in real-time complex problem with great computational savings without losing accuracy and fast convergence.

1.2.3 Thesis outline

The thesis is organized as follows. Part I deals with optimal flow control and local shape optimization. In Chapter 2 we deal with local shape optimization by optimal control (briefly recalled) using local boundary variations and Stokes equations. Design quantities of interest are introduced and motivated. A simplified bypass model is considered. Preliminary feedback procedures are described. In this chapter we introduce also methodological algorithms to solve our problem. In Chapter 3 we deal with small perturbation techniques in fluid mechanics. The linearized shape design problem is described using an optimal control approach in a reference domain using Stokes equations. Theoretical results on well-posedness of the problem, existence and uniqueness of solution are discussed. In Chapter 4 we extend our work considering unsteady Stokes equations and introducing a generalization of our approach. Numerical results are presented in each chapter.

Part II is devoted to reduced basis techniques for optimization problems. In Chapter 5 we introduce the reduced basis problem for Stokes equations where we focus our attention on geometrical parametrization using Cartesian geometries, and proof of stability. Preliminary results on sensitivity analysis of bypass configuration are described. Different outputs are considered and the dual residual weighted approach is introduced to improve the accuracy of the output calculation. In Chapter 6 we deal with a more complex problem introducing non-affine coordinate transformations and curved geometries. In Chapter 7 reduced basis methods are discussed for steady Navier-Stokes equations. Other
feedback results are introduced considering bypass configurations. This chapter concludes the study of the bypass optimization and provides a comparison between low fidelity models, based on Stokes equations, and high fidelity models based on Navier-Stokes equations.

In Chapter 8 several further aspects are considered, such as the possibility to use reduced basis for an optimal control problem (the example deals with advection-diffusion operators), combined also with shape sensitivity. Finally a reduced basis methodology for an unsteady version of the problem is suggested. This chapter provides some perspectives towards the solution of solving parametrized optimal control problems and some conclusions.

Appendix A provides an introduction on reduced basis techniques to explain the methodology and all its aspects dealing with elliptic operators. In this appendix is shown also how to build a parametrized geometrical configuration.

To summarize, the first part of this thesis is focused on shape optimization techniques for local configurations: local boundary variations, small perturbation techniques led by optimal flow control. The second part of the dissertation deals with the optimization of global configurations and is led by the development of reduced basis techniques in parametrized domain with an approach based on a “step-by-step” increasing complexity. The “glue” between the two parts is represented from one side by the kind of application that is considered and, from the other side, by the methodological aspects that are developed. To help the reader following the logical flow of the work, in Figure 1.7 we present the general organization of the thesis with theoretical, methodological and conceptual links highlighted. More interpretation paths can be outlined. For example, in Figure 1.8 we provide the same scheme focusing on the application point of view.
CHAPTER 1. INTRODUCTION AND MOTIVATION

Figure 1.6: A double level of interest: local shape optimization and complete configuration.
Figure 1.7: Thesis organization: theoretical and conceptual links.
Figure 1.8: Thesis organization: application to bypass optimization.
Part I

Optimal Flow Control and Shape Optimization
Chapter 2

Optimal Control and Shape Optimization by Local Boundary Variations

In this chapter we present the first approach in the study of Aorto-Coronaric bypass anastomoses. The theory of optimal control based on adjoint formulation is introduced and applied in order to optimize the shape of the zone of the incoming branch of the bypass (the toe) into the coronary using a local boundary variations method led by the gradient of a cost functional. The aim is to provide design indications in the perspective of future development for prosthetic bypasses. With a reduced model based on Stokes equations and a vorticity functional in the down-field zone of bypass, a Taylor like patch is found. A feedback procedure by Navier-Stokes fluid model based on the analysis of wall shear stress and its related indexes is proposed.

The chapter is organized as follows: firstly we make some remarks on optimal control theory, then we turn our attention to design quantities of interest for the specific bypass configuration problem and we set up a geometrical framework to analyze our problem. Then we describe a first shape optimization algorithm and provide several numerical results. For a general introduction on the bypass anastomoses problem we suggest to see Perktold et al. [80].

2.1 A control approach: looking for cost functional optimization

In this section we briefly review the optimal control theory. This will provide the ground for the set up of our shape optimization algorithms. For a general presentation of the optimal control theory the interested reader can refer to Lions [82], [83] and Aziz et al. [9]. For functional analysis elements see Yosida [167], Brezzi and Gilardi [20]. In the optimal control problems the following mathematical ingredients will set the stage:

- A control function \(u\) which belongs to a functional space \(U_{adm} \subset U\), called admissible control space.
• A state system $y(u)$ associated to the control $u$, $y$ belongs to the Hilbert space $V$ and is the solution of a PDE problem defined on the domain $\Omega$:
\[
Ay(u) = f + Bu, \tag{2.1.1}
\]
where $f$ is a given function (source term), while $A$ is an operator defined on $V$ with values in $V'$ and $B$ is a linear operator defined on $U$ and valued in $V'$, respectively. The latter equation is called state equation and represents the physical system to control and optimize. Note that $u \rightarrow y(u)$ is an affine map.

• An observation function $z(u)$ belonging to a suitable space $Z$ which is related to $y(u)$ through an operator $C$, $z(u) = Cy(u)$. Our observation will coincide with the restriction of $z(u)$ upon a convenient subset of the computational domain $\Omega$, so in our case $C$ is indeed a restriction operator.

• A cost functional $J(u)$ (to be minimized) which is a non-negative quantity depending on the observation $z(u)$.

The general formulation of an optimal control problem reads:

- *Find a function* $u \in U_{adm}$ *so that:*

\[
J(u) = \inf \{ J(v), \ \forall v \in U_{adm} \}. \tag{2.1.2}
\]

Note that the formulation (2.1.1) covers the case where the control variable is either a boundary condition or a source term. Optimal shape design can be regarded as a special case of control theory where the control is the boundary itself. More precisely, $u \rightarrow \Omega(u) \rightarrow y(u)$, where $u \rightarrow \Omega(u)$ is the parametrization of the domain by a control parameter $u$ and the problem is to minimize a functional $J(u, y(u))$, which depends on $u$ through $y(u)$, but also directly (for instance $J$ is an integral of a quantity depending on $y(u)$ extended to a region depending on $u$). See Mohammadi and Pironneau [97]. In our problem the control will be a subset of the domain boundary $\Gamma(u)$.

In the following section we will construct some cost functionals of our interest.

### 2.2 Design control quantities

As already seen in the previous section, cost functionals are related with the observation of physical quantities in the system.

Our interest is for the use of optimal control strategies in the context of improving and optimizing the shape of aorto-coronaric bypasses (see Figure 1.1). In the literature several physical quantities (often called “indexes”) have been proposed in order to assess and measure, at some extent, arteries occlusion risk (or re-occlusion after bypass surgery). These indexes have been introduced in order to enlight some specific mechanism that could be correlated to the narrowing of coronaries (phenomena called intimal thickening). A first index stems from
the observation that a crucial role is played by blood flow oscillations during the diastolic phase of every single heart beat. Another index attempts at measuring the rapid variations of the shear stress (2.2.2) on the vascular wall. In any case these and other indexes (that we are going to review briefly) could be considered as a starting point on the way to synthesize the enormous amount of information given by numerical simulation for the bypass design and they can be used to build cost functionals of interest for our problem.

The *Oscillatory Shear Index* (OSI) was introduced in 1983 by Zarins and Giddens [168] to identify the occlusion risk zones. It is defined as follows:

\[
OSI = \frac{1}{2} \left( 1 - \frac{\int_0^T \tau_w \, dt}{\int_0^T |\tau_w| \, dt} \right), \tag{2.2.1}
\]

where \([0, T]\) is the time interval of a single heart beat \((T \approx 1 \text{ sec})\) and \(\tau_w\) is the *wall shear stress*, indicated with WSS and defined for a Newtonian fluid as:

\[
\tau_w = \mu \frac{\partial u}{\partial n} \cdot \hat{t}, \tag{2.2.2}
\]

where \(u\) is the blood velocity field, \(\mu\) is the blood viscosity, \(\hat{n}\) and \(\hat{t}\) are, respectively, normal and tangential unit vector on the arterial wall. OSI measures the temporal oscillations of the shear stress *pointwise*, without taking into account the shear stress trend in the immediate neighborhood of a specific (critical) point.

Another indicator of occlusion risk zones is the *Mean Wall Shear Stress Gradient* (MWSSG) (see Veneziani [156] and Wells *et al.* [164]) defined as:

\[
MWSSG = \frac{1}{T} \int_0^T |WSSG| \, dt, \tag{2.2.3}
\]

where the *Wall Shear Stress Gradient* (WSSG) is defined as:

\[
WSSG = \frac{\partial \tau_w}{\partial t} = \nabla \tau_w \cdot \hat{t}. \tag{2.2.4}
\]

According to MWSSG, the zones featuring occlusion risk or abnormal flow pattern are those where strong variations in the shear stress along the wall occur. This index is dual to OSI, since it weighs spatial variations.

An alternative to OSI is the *Oscillatory Flow Index* (OFI), see Taylor *et al.* [149]:

\[
OFI = \frac{1}{2} \left( 1 - \frac{\int_0^T Q \, dt}{\int_0^T |Q| \, dt} \right), \tag{2.2.5}
\]

where \(Q\) is the *flow rate* across an artery section \(\Gamma_f\),

\[
Q = \int_{\Gamma_f} \rho u \cdot \hat{n} \, d\Gamma, \tag{2.2.6}
\]

being \(\rho\) the blood density. It quantifies temporal oscillations in flow rate.

A problem arising when using one of the previous functionals for optimization purposes is
that they don’t have a quadratic form. We can use these quantities for feedback processes, but we need functionals which can provide a good average in time and in space of shear stress trend. With this aim, we propose other integral quantities that can be defined along the artery wall \( \Gamma_w \) (or a critical part of it). Set for all \( t > 0 \):

\[
\Sigma(t) = \int_{\Gamma_w} \left( \frac{\partial \tau_w(t)}{\partial t} \right)^2 d\Gamma,
\]

and

\[
Q(t) = \int_{\Gamma_w} \left( \frac{\partial Q(t)}{\partial t} \right)^2 d\Gamma.
\]

Then we define:

\[
J_r = \text{mean}_{[0,T]} \Sigma(t) = \frac{1}{T} \int_0^T \Sigma(t) dt = \frac{1}{T} \int_0^T \int_{\Gamma_w} \left( \frac{\partial \tau_w(t)}{\partial t} \right)^2 d\Gamma dt,
\]

which is the \( L^2 \)-norm of the wall shear stress rate (\( L^2 - \text{WSSR} \)). Similarly, we define the analogous \( L^2 \)-norm of the flow rate \( Q \) (\( L^2 - \text{FR} \)):

\[
J_Q = \text{mean}_{[0,T]} Q(t) = \frac{1}{T} \int_0^T Q(t) dt = \frac{1}{T} \int_0^T \int_{\Gamma_f} \left( \frac{\partial Q}{\partial t} \right)^2 d\Gamma dt.
\]

These new indexes \( J_r \) and \( J_Q \) could be seen as a lumped and complete information from the previous cost functionals because they take into account both of spacial and temporal oscillation of fluid mechanics quantities.

In the perspective of using low fidelity methods for control process (to reduce computational time and costs) and then higher fidelity methods in feedback we are going also to measure other fluid mechanics quantities such as vorticity and energy dissipation (due to viscous terms) in distributed zone of flow field (and not pointwise quantities), above all when using steady fluid models. These indexes (vorticity, viscous energy dissipation) contain many information on flow properties and their links with geometry variation.

### 2.3 Mathematical modelling of the problem.

Blood is a very complex fluid which interacts with the compliant arterial vessels. Some simplificatory assumptions are in order in view of applying a control procedure. Concerning the fluid model we consider blood as a Newtonian fluid, modelled by Navier-Stokes equations (or even by the simplified Stokes model). This model is an acceptable approximation when studying blood flow in large or medium-size arteries. In coronary arteries blood is also characterized by a low velocity profile. The Reynolds number (the ratio between inertial and viscous forces) is rather low, so the flow can be regarded as laminar.

#### 2.3.1 The geometrical model

To model the incoming branch of a bypass, different bypass anastomoses models have been taken into consideration. We have simplified our model by considering a longitudinal section
in the mean plane reducing it to a two-dimensional problem. As initial configuration we have used a fiberglass model developed at the Surgical Vascular Division of Aarhus University Hospital (Denmark). See Figure 2.1. To define anastomoses geometry for numerical simulation,

![Figure 2.1: 3-D Anastomoses fiberglass model (Aarhus University Hospital, Denmark)](image)

different aspects have to be accounted for:

- Fluid dynamics phenomena in the bypass are strictly related with the shape assumed, such as the presence of blood re-circulating zones, flow separation and vorticity. Blood pulsatility due to heart beat influences our analysis and this aspect will be considered in the next chapters. The most critical zone in the bypass design is the toe and must be carefully modelled. The toe, in fact, models the incoming branch of the bypass and influences the way in which the blood approaches in the host artery from the bypass bridge device. See Figure 2.2.

- Another important aspect to keep into consideration is the graft angle between the incoming branch and the occluded artery. This aspect will be discussed in details in the second part of this thesis.

- The diameter of the occluded branch (the lower branch) is $D = 3.5$ mm and the one taken in this first investigation for the bypass is $0.96 \times D$, as suggested in Quarteroni and Formaggia [126] and used in surgery procedures. This quantity will be studied more deeply in the second part of the thesis.

A curvature in the vascular wall is introduced to model coronary arteries in order to take into consideration the presence of the heart muscle. Figure 2.2 shows the mean plane before applying the shape optimization process. These simple considerations already underline the importance of the study of a complete parametrized bypass configuration taking into account graft angle, curvature and other quantities to define the complete device scheme.

### 2.3.2 The state problem: Navier-Stokes and Stokes equations

In our optimal control process we use as state equations the steady Stokes equations which are however well suited for modelling Newtonian flows at low Reynolds number. In reality, because
of flow pulsatility we should consider the unsteady Navier-Stokes equations. Unfortunately, such as a complete fluid model would be too complex and expensive to control because of evolution terms, non-linearity in the adjoint problem and the fact that the shape of our computational domain depends on time. Nonetheless, after performing shape optimization on our reduced model governed by steady Stokes equations, we will turn our attention to the complete Navier-Stokes model for feedback procedures considering design cost functionals such as the ones presented in Section 2.2. We focus therefore on the steady Stokes problem.

The Stokes equations in a two-dimensional computational domain $\Omega$ with velocity vector $\mathbf{u} = (u, v)^T$ and pressure $p$ (which is divided by the constant blood density $\rho$) read:

$$
\begin{align*}
-\nu \Delta \mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in} \; \Omega \subset \mathbb{R}^2, \\
\nabla \cdot \mathbf{u} &= 0 \quad \text{in} \; \Omega, \\
\mathbf{u} &= 0 \quad \text{on} \; \Gamma_w, \\
u \frac{\partial \mathbf{u}}{\partial n} - p\mathbf{n} &= 0 \quad \text{on} \; \Gamma_{out}, \\
\mathbf{u} \cdot \hat{\mathbf{n}} &= 0, \; \mathbf{u} \cdot \mathbf{n} &= g_{in} \quad \text{on} \; \Gamma_{in},
\end{align*}
$$

where $\mathbf{f}$ is a force field (for example gravity $\mathbf{f} = (0, 9.8)^T \text{ms}^{-2}$), $\mathbf{n}$ and $\hat{\mathbf{t}}$ are, respectively, normal and tangential unit vector on the domain boundary $\partial \Omega$. The latter is partitioned in three components: $\Gamma_{in}$ is the inflow boundary, $\Gamma_{out}$ the outflow boundary and $\Gamma_w$ the boundary corresponding to the arterial wall; Figure 2.2 represents schematically the computational geometry and the symbols used.

Concerning boundary conditions, no-slip conditions are imposed on all vascular walls of $\Gamma_w$, over the stenosed artery portion (if we consider a complete occlusion, otherwise we have to consider this section in the same way of the inflow section but with a reduction of flow rate) and on the incoming branch of bypass. At inflow section $\Gamma_{in}$, a Hagen-Poiseuille’s velocity profile $g_{in}$ is imposed, while on outflow section $\Gamma_{out}$ of the artery a free-stress, Neumann-type, boundary condition has been imposed. Velocity values at the inflow are chosen in such a
way that the Reynolds number \( Re = \frac{\bar{u}D}{\nu} \) has order \( 10^3 \). Blood kinematic viscosity \( \nu = \frac{\mu}{\rho} \) is \( 4 \cdot 10^{-6} \text{ m}^2 \text{ s}^{-1} \), blood density \( \rho = 1 \text{ g cm}^{-3} \) and dynamic viscosity \( \mu = 4 \cdot 10^{-2} \text{ g cm}^{-1} \text{s}^{-1} \); \( \bar{u} \) is a mean inflow velocity (in absolute value) related with \( g_{in} \), while \( D \) is the arterial diameter (3.5 mm). Arterial thickness is about 0.5 mm [126].

The role of control

In our problem the control \( w \) represents the shape \( \Gamma_w \) itself or a part \( \Gamma_c \) of it (typically the incoming branch). Precisely, we write (see Figure 2.2):

\[
\Gamma_w = \Gamma_{w(rigid)} \cup \Gamma_{c(toe)},
\]

(2.3.2)

where

\[
\Gamma_{w(rigid)} = \Gamma_{w(heel)} \cup \Gamma_{w(graft)} \cup \Gamma_{w(stenosis)},
\]

(2.3.3)

then we can build our geometrical model for \( \Gamma_c \) (double-dotted in Figure 2.2) made up of \( M \) branches \( \Gamma^j_c(w) \) represented by:

\[
\Gamma_c = \bigcup_{j=0}^{M} \Gamma^j_c(w), \quad \Gamma^j_c(w) = \Gamma^0_c + w^j,
\]

(2.3.4)

where \( w^j \) is the control variable (control shape function) which changes by \( \delta w^j_k \) at each iteration of the optimization process. At the \( k \)-th iteration we have:

\[
w^j_k = \sum_{m=0}^{k-1} (\delta w^j_m),
\]

(2.3.5)

its expression will be derived in Section 2.5. Further, we take the curves (of order up to \( N \))

\[
\Gamma^j_c(s) = \sum_{i=0}^{N} \alpha^j_i f^j_i(s),
\]

(2.3.6)

which represent the initial shape, \( s \) is the horizontal abscissa that parametrizes the curve \( \Gamma_c \), \( f^j_i(s) = \sum_{k=0}^{i} a_k s^k \) are given shape polynomial functions of \( i^{th} \) order and \( \alpha^j_i \) are suitable weights (clearly each branch describing the initial shape must be properly set to guarantee a boundary with sufficient regularity).

Weak formulation for the state problem

We introduce a weak formulation for the state problem (2.3.1), see [127]: find \( u \in V = (H^1(\Omega))^2 \), \( p \in Q = L^2(\Omega) \) s.t.

\[
\begin{aligned}
a(u,v) + b(p,v) &= F(v) \quad \forall v \in X, \\
b(q,u) &= 0 \quad \forall q \in L^2(\Omega), \\
u &= u_{in} \quad \text{on } \Gamma_{in},
\end{aligned}
\]

(2.3.7)
where
\[ a(u, v) = \int_\Omega \nu \nabla u \cdot \nabla v d\Omega, \]
\[ b(p, v) = -\int_\Omega p \nabla \cdot v d\Omega, \quad F(v) = \int_\Omega f \cdot v d\Omega, \]
\[ X := \{ v : v \in (H^1(\Omega))^2, v = 0 \text{ on } \Gamma_i \cup \Gamma_w \}, \]
and \( u_m \) is \( g_in \n \) on \( \Gamma_m \) and \( 0 \) on \( \Gamma_w \). Note that the second equation in system (2.3.7) may be also written in the equivalent form:
\[ b(q, u) = 0; \] after numerical discretization, the latter formulation will lead to an algebraic system which is no longer symmetric but always positive definite ([127]). This alternative will be adopted to build the adjoint problem.

**Galerkin approximation**

The Galerkin approximation for the Stokes equations (2.3.7) is the following: find \( u_h \in V_h, \) \( p_h \in Q_h \) such that
\[
\begin{aligned}
& a(u_h, v_h) + b(p_h, v_h) = (f, v_h) - a(R_h, v_h) \quad \forall v_h \in V_h, \\
& -b(q_h, u_h) = b(q_h, R_h) \quad \forall q_h \in Q_h,
\end{aligned}
\] (2.3.8)
here \( R_h \in V_h \) and \( R_h|_{\Gamma_i} = u_{in}^h \). The spaces \( V_h \subset V \) and \( Q_h \subset Q \) are two families of subspaces of finite dimension depending on the discretization parameter \( h \). For the proof concerning the existence and uniqueness of solution of problem (2.3.8) and approximation stability aspects we suggest to see [127]. We denote
\[ \{ \varphi_j \in V_h \}, \{ \phi_k \in Q_h \}, \]
the basis functions for the spaces \( V_h \) and \( Q_h \), respectively. We can develop \( u_h \) and \( p_h \) on their basis getting
\[ u_h(x) = \sum_{j=1}^{\mathcal{N}} U_j \varphi_j(x), \quad p_h(x) = \sum_{k=1}^{\mathcal{M}} p_k \phi_k(x), \] (2.3.9)
where \( \mathcal{N} = \dim(V_h), \mathcal{M} = \dim(Q_h) \). We get the following linear system for the state problem:
\[
\begin{aligned}
& AU + B^T P = F, \\
& BU = 0,
\end{aligned}
\] (2.3.10)
where \( A \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}} \) and \( B \in \mathbb{R}^{\mathcal{M} \times \mathcal{N}} \) are matrices related to bilinear forms \( a(\cdot, \cdot) \) and \( b(\cdot, \cdot) \), with elements given by
\[ A = (a_{ij}) = (a(\varphi_j, \varphi_i)), \]
\[ B = (b_{km}) = (b(\phi_k, \varphi_m)), \]
while \( U \) and \( P \) are vectors of unknowns,
\[ U = (U_j), \quad P = (p_j). \]
Remark on stabilization

We have based our approximation on Galerkin-Finite element method and we use, for example, piecewise linear elements for velocity $\varphi$ and pressure $\phi$. This choice of elements with the same polynomial degree has to be accompanied by a stabilization procedure. We use the SUPG stabilization (see, e.g., Quarteroni and Valli [127]). We introduce a relaxation on the incompressibility constraint: find $u_h \in V_h$, $p_h \in Q_h$ such that:

\[
\begin{align*}
\left\{ \begin{array}{ll}
a(u_h, v_h) + b(v_h, p_h) = (f, v_h) & \forall v_h \in V_h, \\
b(u_h, q_h) = \Phi_h(q_h) & \forall q_h \in Q_h,
\end{array} \right.
\end{align*}
\]  

(2.3.11)

where $\Phi_h$ has a linear dependence on the test function $q_h$, but can also depend on $u_h, p_h, f$ and $h$. An example of $\Phi_h$ in the case of linear finite elements $P_1 - P_1$ is:

\[
\Phi_h(q_h) = \sum_{K \in T_h} h_K^2 \int_K \nabla p_h \cdot \nabla q_h,
\]

where $T_h$ is the finite element triangulation whose elements are indicated with $K$.

2.3.3 The observation on the system

After having described the control strategy on the boundary and introduced the state problem (and its approximation) we have to choose a cost functional. We consider a distributed observation in the down-field zone of the incoming branch of the bypass indicated as $\Omega_{wd}$ (see Figure 2.2). For this preliminary study our observation will be the vorticity $\omega = \nabla \times u = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$; $u$ is the solution of the Stokes equations (2.3.1) and we control the system by minimizing the following functional:

\[
J(\omega) = \frac{1}{2} \int_{\Omega_{wd}} |\nabla \times u|^2 \, d\Omega + \alpha \|\omega\|^2, \quad (0 < \alpha \ll 1).
\]  

(2.3.12)

where the last term provides the minimum shape deformation and guarantees existence of the solution. The symbol $\|.|\|$ denotes the norm of $L^2(\Omega_{wd})$. By this observation we can draw information on the vorticity distribution. The curl of the vorticity will provide a source term of the adjoint problem, whose solution represents the sensitivity of the cost functional to the observation (related with state solution) and will be used during our optimization procedure to get the gradient of the cost functional allowing computational savings. For further aspects on vorticity reduction problems see Berggren [17].

We underline that we consider not only what happens on the vascular wall (i.e. on part of the domain boundary) but also in the fluid domain. For this reason this investigation is also a flow control problem and not only a shape optimization problem, because an optimal shape design problem is usually based only on the boundary observation. In this approach we have also taken into consideration another important aspect, concerning wall stiffness and elasticity. In fact we can rewrite the cost functional of interest into a different form, replacing the term $\alpha \|\omega\|^2$ of (2.3.12) with the following one including the minimization of the wall
deformation (which will depend on mechanical properties of the wall), in the zone where the bypass merges with the artery:

\[ J(w) = \frac{1}{2} \int_{\Omega_{\text{w}}^f} | \nabla \times \mathbf{u} |^2 \, d\Omega + \frac{E_0}{\xi} \sum_j \left( \int_{\Gamma_{\xi}(w)} d\Gamma - \tilde{l}_j \right)^2. \]  

(2.3.13)

Here \( \tilde{l}_j \) is the original length of the vascular wall \( \Gamma_{\xi}(w) \), \( \xi \) is a suitable weight, while \( E_0 \) is Young’s elasticity modulus \((3 \cdot 10^6 \text{ dyne cm}^{-2})\). We are thus considering a multi-objective optimization of a prosthetic device. The last term in (2.3.13) describes the elastic wall behavior and provides a measure for elastic energy and wall deformation. The two quantities on the right end side of (2.3.13) have to be weighed by the choice of \( \xi \).

The choice of a vorticity based functional to extract information can be better understood looking at Figures 2.3 and 2.4 where vortices structures in the downfield zone of a bypass are shown.

Figure 2.3: Vorticity in the down-field zone of a bypass (from a numerical simulation carried out by Fischer, Lee, Loth; University of Illinois and Argonne National Laboratories by permission).

Figure 2.4: Vorticity in the downfield zone of a bypass: 2D details at the incoming branch (from a numerical simulation carried out by Fischer, Lee, Loth; University of Illinois and Argonne National Laboratories by permission).
2.3.4 The adjoint problem

As already mentioned, during the optimization process we must solve another PDE problem, the so-called adjoint problem. Considering the cost-functional (2.3.12), the state problem (2.3.7) and replacing test functions \(v\) and \(q\) with \(p\) and \(\sigma\), respectively, the associate Lagrangian reads:

\[
\mathcal{L}(u, p; \sigma; w) = J(w) - a(u, p) - b(p, p) + (f, p) + b(\sigma, u),
\]

(2.3.14)

whose derivative with respect to \((u, p)\) imposed equal to zero gives us the adjoint problem (see Gunzburger [50]):

\[
\mathcal{L}_{u,p}(u, p; \sigma; w)[\delta u, \delta p] = - \int_{\Omega} \nu \nabla \delta u \cdot \nabla pd\Omega + \int_{\Omega} \delta p \nabla \cdot pd\Omega - \int_{\Omega} \sigma \nabla \cdot \delta u d\Omega +
\]

\[
+ \int_{\Omega_{wd}} (\nabla \times u) \cdot (\nabla \times \delta u) d\Omega = 0, \quad \forall [\delta u, \delta p],
\]

which, using the notation already introduced for the state problem, can be written in the weak form as: find \(p \in X, \sigma \in L^2(\Omega)\) s.t.

\[
\begin{align*}
\begin{cases}
a(p, q) - b(\sigma, q) = H(q) \quad \forall q \in X, \\
b(\sigma, p) = 0 \quad \forall \sigma \in L^2(\Omega),
\end{cases}
\end{align*}
\]

(2.3.15)

where \(p\) and \(\sigma\) denote the adjoint velocity and pressure, respectively. We have replaced the test functions \(\delta u\) and \(\delta p\) with \(q\) and \(\sigma\) respectively, while the observation term is

\[
H(q) = \int_{\Omega_{wd}} (\nabla \times u) \cdot (\nabla \times q) d\Omega =
\]

\[
= \int_{\Omega_{wd}} q \cdot (\nabla \times \nabla \times u) d\Omega + \int_{\Gamma_{wd}} \mathbf{n} \cdot (q \times \nabla \times u) d\Gamma,
\]

using in the last step Green’s identities. This term is related with an internal source defined in the down-field zone and its boundary, see for details Berggren [17] where a similar problem is studied. A no-slip condition is imposed on the vascular wall \(\Gamma_w(w)\) and also on \(\Gamma_{in}\), while a “free-stress” condition \(\nu \frac{\partial p}{\partial n} + \sigma \mathbf{n} = 0\) is imposed on \(\Gamma_{out}\).

The discretization of this problem is carried out using the Galerkin-Finite Element method as introduced for the state problem (2.3.8): find \(p_h \in V_h, \sigma_h \in Q_h\) such that

\[
\begin{align*}
\begin{cases}
a(p_h, q_h) - b(\sigma_h, q_h) = H(q_h) \quad \forall q_h \in V_h, \\
b(\sigma_h, p_h) = 0 \quad \forall \sigma_h \in Q_h,
\end{cases}
\end{align*}
\]

(2.3.16)

An important role is played by the adjoint solution to get the quantity \(J'(w)\), which is used to provide indications on the shape modification of \(\Gamma'_w\). In fact, \(J'(w) = G(p, u, w)\): a detailed analysis will follow in Section 2.5 after introducing some notions on optimization algorithm (Section 2.4).
We anticipate here, just to clarify, that at $k$–th iteration of the optimization algorithm described in Section 2.4.3 we will have

$$\delta w_k^j = -\xi_j(s)\lambda_k^j J_k^j(w_k^j),$$

where $\lambda_k^j$ is a relaxation parameter for descent gradient-based method, called step-size (see Quarteroni, Sacco and Saleri [129]), while $\xi_j(s)$ is a weight for the shape variation, used to satisfy congruence geometrical conditions such as shape continuity and guarantee that the internal arterial diameter cannot go below a certain threshold.

The cost functional depends on $J_k(w_k)$ to underline that at each iteration the integral is computed in a zone of the domain $(\Omega_{wd})_k$ whose shape is changing at each iteration. During the shape optimization process we have two degrees of freedom available $(x, y)$ in the mean plane of the bypass. We only consider a normal displacement with respect to the original boundaries (as done also by Gunzburger [50] and Mohammadi and Pironneau [97]). For this reason the weight parameter $\xi_j(s)$ is a function of linear abscissa only.

### 2.4 Optimization and control algorithm

Before the description of the algorithm for the solution of the discretized control problem (state and adjoint equations) introduced in Sections 2.3.2-2.3.4 we recall some elements dealing with cost functional optimization and geometrical topology to manage with shape variations.

#### 2.4.1 Gradient method

Let $U$ be a Hilbert space, $(., .)$ its associated scalar product, and $w \in U \rightarrow J(w) \in \mathbb{R}$ then $J'(w)$ is a linear operator from $U$ to $\mathbb{R}$ such that

$$J(w + \delta w) = J(w) + J'(w)\delta w + o(\|\delta w\|_U)$$

At the basis of gradient methods is the Taylor expansion

$$J(w + \rho z) = J(w) + \rho(\text{Grad}_w J(w), z) + o(\|\rho\| z\|_U) \quad \forall w, z \in U, \quad \forall \rho \in \mathbb{R},$$

(2.4.1)

where $\text{Grad}_w J(w)$ is the element of $U$ defined by

$$(\text{Grad}_w J(w), z) = J'_w(w)z \quad \forall z \in U,$$

given by Riesz’ Theorem. By taking $z = -\text{Grad}_w J(w)$ in (2.4.1), with $0 < \rho \ll 1$ we find:

$$J(w + \rho z) - J(w) = -\rho\|\text{Grad}_w J(w)\|_U^2 + o(\rho\|\text{Grad}_w J(w)\|_U).$$

Hence if $\rho$ is small enough the first term on the right-hand side will dominate the remainder and the sum will be negative:

$$\rho\|\text{Grad}_w J(w)\|_U^2 > o(\rho\|\text{Grad}_w J(w)\|_U) \Rightarrow J(w + \rho z) < J(w).$$
Thus the sequence defined by:

$$w_{n+1} = w_n - \rho \text{Grad}_w J(w_n), \ n = 0, 1, 2, \ldots$$ (2.4.2)

makes the sequence of real numbers $J(w_n)$ monotonically decreasing. We have the following result:

**Theorem 2.4.1** *(from Pironneau [70]).* If $J$ is continuously differentiable, bounded from below, and $+\infty$ at infinity, then all accumulation points $w^*$ of $w_n$, generated by (2.4.2) satisfies

$$\text{Grad}_w J(w^*) = 0$$

This is the so-called *optimality condition* of the first order of the problem. If $J$ is convex then $w^*$ is a minimum; if $J$ is strictly convex the minimum is unique. By taking the best $\rho$ in the descent direction $z_n = -\text{Grad}_w J(w_n)$,

$$\rho_n = \text{argmin}_\rho \{ J(w_n + \rho z_n) \},$$

that is

$$J(w_n + \rho_n z_n) = \min_\rho \{ J(w_n + \rho z_n) \},$$

we obtain the so-called *method of steepest descent with optimal step size*. Indeed, thanks to the analysis in [116], it is enough to find $\rho_n$ with the following property (Armijo rule) to guarantee convergence: find $\rho$ such that for $0 < \alpha < \beta < 1$:

$$-\rho \beta \|z\|^2 < J(w_n + \rho z) - J(w_n) < -\rho \alpha \|z\|^2$$

to have $w_{n+1} = w_n + \rho z$. An approximate Armijo rule takes only one line of slope $\alpha \|z\|^2$ and first finds the largest $\rho$ of the form $\rho = \rho_0 2^{\pm k}$ which gives a decrement for $J$ below the line for $\rho$ and above the line for $2\rho$: choose $\rho_0 > 0$, $\alpha \in (0, 1)$ and find $\rho = \rho_0 2^{\pm k}$ where $k$ is the smallest signed integer ($k$ can be negative) such that

$$J(w_n + \rho z) - J(w_n) < -\rho \alpha \|z\|^2$$

and

$$-2\rho \alpha \|z\|^2 \leq J(w_n + 2\rho z) - J(w_n).$$

A good choice is $\rho_0 = 1$, $\alpha = \frac{1}{2}$.

### 2.4.2 Shape variation and control

In functional spaces, as in finite dimension, optimization methods require the gradient of the cost functional $J$ and for this we need to define an underlying Hilbert structure for the parameter space, in our case the shapes (see Zolésio [144]). Two strategies can be pursued:

- All the admissible shapes are obtained by a mapping transformation from a reference domain $\tilde{\Omega} : \Omega = T(\tilde{\Omega})$. Then the parameter is $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$. This approach will be used and illustrated in detail in the next two chapters.
It is possible to guarantee the Hilbert structure (the existence of a scalar product) only for small variations of the shape $\partial \Omega$. Dealing with local shape variations $\delta w$ defined about a reference boundary $\Gamma$ we consider:

$$\Gamma'(w) = \{ x + \delta w(x) \hat{n}(x) : x \in \Gamma \},$$

where $\hat{n}$ is the outer normal to $\Gamma$ at $x$ and $\Omega$ is the domain which is on the left side of the oriented boundary $\Gamma'(w)$. The Hilbert structure is placed on $w$, for instance $w^2_\Omega$. Considering a small perturbation $\Sigma'$ of $\Sigma \subset \Gamma$ given by $\delta w = -\lambda J'(w)$, using a gradient method, we have:

$$\Sigma' = \{ x - \lambda J'(w) \hat{n} : x \in \Sigma \}$$

where $w$ is a function of $x \in \Sigma$, $\Sigma$ is endowed with sufficient regularity, and $\lambda$ is a positive number destined to tend to zero. We denote the “new” domain $\Omega' = \Omega(\Sigma')$. For any $f \in H^1(C)$, where $C$ is a subset of $\Omega \cup \Omega'$ we have

$$\int_{\Omega'} f - \int_{\Omega} f = \int_{\Sigma} \delta w f + o(\lambda). \quad (2.4.3)$$

This strategy has been used in this chapter and illustrated in the following sections.

### 2.4.3 The iterative process

The generic step of an algorithm for the solution of the control problem described in Section 2.3 can be devised as follows. Note that the pedix $h$ referring to discretized solution is omitted for simplicity of notation and the index $k$ refers to iteration index.

- Solve the state problem (2.3.8) with velocity components $u_k$ and $v_k$ in the domain $\Omega_k$ with the moving boundary $\Gamma_{c_k}^d(w_{k-1})$ obtained from the previous iteration of the control cycle.

- Compute the vorticity $\nabla \times u_k$ in the domain $(\Omega_{wd})_k$.

- Evaluate the cost functional (2.3.12) or (2.3.13).

- Solve the adjoint problems (2.3.16) for $p_k$ and $\sigma_k$ with source term $H$ depending on $(\nabla \times u_k)$.

- Use the stopping test on the quantity $J_k'(w_k)$ related with $u_k$ and the adjoint state $(p_k, \sigma_k)$:

$$||J_k'(w_k)||_\mu \leq tol, \quad (2.4.4)$$

for a suitable tolerance $tol$.

- Estimate the shape variation $\delta w_{k}^j$ ($\delta w_{0}^j = 0$) on the boundary $\Gamma_{c_k}^j(w_k)$ by a descent gradient-type method (see Kawohl et al. [70]). Then using (2.4.2):

$$\delta w_{k}^j |_{\Gamma_{c_k}^j} = -\lambda_k J_k'(w_k) = -\tilde{\lambda}_k G_k(p_k, u_k, w_k) \quad (2.4.5)$$

with $0 < \lambda_k \ll 1$, which causes a vertical shape displacement $\delta w_k^j$ in (2.3.4).


- Update the boundary shape:

\[ \Gamma_{c,k+1}(w_k) \simeq \Gamma_{c} + w_{k}^j + \delta w_{k}^j = \Gamma_{c} + \sum_{m=0}^{k} (\delta w_{m}^j). \]  \hspace{1cm} (2.4.6)

- Modify the boundary and re-construct or adapt the computational grid (see Section (2.4.4)).

Figure 2.5 shows a general scheme for a flow control and shape optimization problem (discretization aspects are understood in the scheme focused on optimization algorithm). For more on optimal shape design techniques by boundary variations and shape parameters see e.g. Di Césaré [33].

### 2.4.4 Mesh strategy

During the shape optimization process, introduced above, the mesh is subject to be modified by boundary variations. There are two different strategies for mesh treatment at every step:

- Mesh reconstruction. This process is computationally expensive, even if new elements can be added.

- Mesh stretching and elements adaptivity after boundary variations operated by small deformation. This approach is faster and cheaper than the previous one. Mesh is usually stretched only in the computational domain surrounding the boundary \( \Gamma_{c} \) subject to shape changes. This technique is better because, when applicable, it guarantees the continuity of the solution (for both state and adjoint problem) during the optimization process (at different iterations).

Our approach used both strategies: the mesh was completely re-built only after a certain number of iterations, when the elements lost some of their properties set to guarantee regularity, while at each step mesh was stretched by locally boundary variations and adapted. In this process regularity conditions were put on grid quality such as on the control of minimum angles or maximum side-lengths. See Hecht *et al.* [55] and [53]. To preserve regular meshes during shape deformation we may reduce the step-size of the gradient method (to limit shape deformation) or introduce some regularization procedures.

### Smoother

To avoid boundary oscillations (and irregular shapes) and due to the fact that the gradient method produces shape variations \( \delta w_{k} \) which have necessarily less regularity than the original parametrization we can apply the following smoothing operator over the shape, defined in \( \Omega_{k} \), as proposed and demonstrated by Mohammadi and Pironneau [97]:

\[
\begin{align*}
(I - \epsilon \Delta) \delta w &= \delta \tilde{w}, \\
\delta w &= 0 \text{ on } \Gamma_{in} \cup \Gamma_{out}, \\
\frac{\partial \delta w}{\partial n} &= 0 \text{ on } \Gamma_{c}, \\
\delta w &= \delta \tilde{w} \text{ on } \Gamma_{w} \setminus \Gamma_{c}
\end{align*}
\]  \hspace{1cm} (2.4.7)
where $\delta \tilde{w}$ is the shape variation to be regularized and $\delta w$ the smoothed one, while $\epsilon$ is a viscosity parameter to be set to avoid mesh degeneration. For this reason a control over geometrical quantities such as minimum angle or maximum side-length has been used. We use this smoothing procedure at each iteration between steps (2.4.5) and (2.4.6); the value of the “viscosity” parameter at convergence is set to zero, so that we have an identity operator

Figure 2.5: Optimal shape design by flow control: scheme of the algorithm.
2.5 An equivalent interpretation of the problem

The cost functional (2.3.12) represents the square of the vorticity in the down-field (host vessel) zone of the bypass. In the adjoint problem (2.3.15) we have considered the source field

\[ H(q) = \int_{\Omega_{wd}} q \cdot (\nabla \times \nabla \times u) d\Omega + \int_{\Gamma_{wd}} \mathbf{n} \cdot (q \times \nabla \times u) d\Gamma. \]

Let us consider the particular case where $\Omega_{wd} \equiv \Omega$ and the state problem has Dirichlet conditions on the whole boundary or, alternatively, as it happens in our case, a free-stress Neumann boundary condition on the outflow, where $\mathbf{n} = (1, 0)^T$ and Dirichlet conditions on inflow and arterial walls. Then the boundary term of $H(q)$ is zero. Note that the boundary term of $H(q)$ can be written also as $\int_{\omega_{wd}} (\nabla u \cdot \mathbf{n}) q d\Gamma$.

The following identity holds:

\[ \nabla \times \omega = \nabla \times \nabla \times u = \nabla(\nabla \cdot u) - \Delta u \]

and since we are considering incompressible flows ($\nabla \cdot u = 0$) it would have been equivalent, from a physical point of view, using the source term $-\Delta u$ for the adjoint problem (2.3.15) instead of $\nabla \times \nabla \times u$. The corresponding new cost functional (multiplied by the viscosity $\nu$) becomes:

\[ J(u) = \frac{1}{2\nu} \int_\Omega |\nabla u|^2 d\Omega, \]

and represents the minimization of the energy of the system dissipated by viscous stresses (related with viscous drag). From now we will therefore consider this new functional.

To analyze this new problem (see also [97]), let us consider a domain $\Omega'$ representing a small perturbation of $\Omega$. Its boundary $\Gamma' = \partial\Omega'$ can be represented as follows,

\[ \Gamma' = \{ x + \delta w(x) \mathbf{n}(x) \mid x \in \Gamma = \partial\Omega \}, \]

where $\delta w$ is regular (for example $\delta w \in C^2$) and small. We define also:

\[ \delta \mathbf{u} = u(\Omega') - u(\Omega) \equiv \mathbf{u}' - \mathbf{u} \]

where $\mathbf{u}$ has been extended from $\Omega$ to $\Omega'$ such that $\mathbf{u} = 0$ in $\Omega' \setminus \Omega$. This extension is possible because we have zero Dirichlet boundary condition on the wall subject to shape variations. We will see that $\delta \mathbf{u}$ (and hence $\mathbf{u}'$) can be obtained solving an additional problem. We introduce

\[ \delta J \equiv J(\Omega', \mathbf{u}') - J(\Omega, \mathbf{u}) = \frac{1}{2\nu} \left( \int_{\Omega'} |
abla \mathbf{u}'|^2 d\Omega - \int_\Omega |
abla \mathbf{u}|^2 d\Omega \right), \]
we have stressed the domain on which the cost functional is defined. We rewrite \( \delta J \) as

\[
\delta J = \frac{1}{2} \nu \chi \left( \int_{\Omega} |\nabla u|^2 \, d\Omega \right) = \frac{1}{2} \nu \int_{\Omega} |\nabla u|^2 \, d\Omega + \nu \int_{\Omega} \nabla \delta u \cdot \nabla u \, d\Omega + o(\delta \Omega, \delta u),
\]

(2.5.1)

where \( \delta \Omega = \Omega' - \Omega \). We suppose that \( \nabla u \) is smooth, and we rewrite the first term of (2.5.1), related with \( \delta \Omega \) as an "extension" of the value of \( |\nabla u| \) on \( \Gamma \):

\[
\frac{1}{2} \nu \int_{\delta \Omega} |\nabla u|^2 \, d\Omega = \frac{1}{2} \nu \int_{\Gamma} \delta w |\nabla u|^2 \, d\Gamma + o(||\delta w||_{C^2}) = \frac{1}{2} \nu \int_{\Gamma} \delta w |\nabla u|^2 \, d\Gamma + o(||\delta w||_{C^2}),
\]

(2.5.2)

having exploited the fact that the tangential derivative \( \frac{\partial u}{\partial \nu} = 0 \) since \( u = 0 \) on \( \Gamma \). To rewrite the second term of (2.5.1) it is possible to show that \( \delta u \) is a solution to the following problem (now referring to the domain considered in Section 2.3) representing small flow perturbation in \( \Omega \). Then rewriting problem (2.3.1) we have: find \( \delta u \) and \( \delta p \) in \( \Omega \) such that

\[
\begin{cases}
-\nu \Delta \delta u + \nabla \delta p = 0 \quad \text{in} \; \Omega \subset \mathbb{R}^2, \\
\nabla \cdot \delta u = 0 \quad \text{in} \; \Omega, \\
\delta u = 0 \quad \text{on} \; \Gamma_u \backslash \Gamma_c, \\
\delta u = 0 \quad \text{on} \; \Gamma_{in}, \\
\nu \frac{\partial \delta u}{\partial \nu} - \delta p \hat{n} = 0 \quad \text{on} \; \Gamma_{out}, \\
\delta u = -\delta w \frac{\partial u}{\partial \nu} \quad \text{on} \; \Gamma_c,
\end{cases}
\]

(2.5.3)

Here we assume \( \Gamma_c \subset \Gamma \) as a portion of the boundary subject to shape optimization (perturbation is only on \( \Gamma_c \) which becomes \( \Gamma'_c \)). The only non-obvious relation is the boundary condition on \( \Gamma_c \) (we recall that \( u \) is already known). By a Taylor expansion we rewrite \( u'|_{\Gamma_c} = u'(x + \delta w \hat{n}) \) as

\[
u \Delta \delta u + \nabla \delta p = 0 \quad \text{in} \; \Omega \subset \mathbb{R}^2,
\]

\[
\nabla \cdot \delta u = 0 \quad \text{in} \; \Omega,
\]

\[
\delta u = 0 \quad \text{on} \; \Gamma_u \backslash \Gamma_c,
\]

\[
\delta u = 0 \quad \text{on} \; \Gamma_{in},
\]

\[
\nu \frac{\partial \delta u}{\partial \nu} - \delta p \hat{n} = 0 \quad \text{on} \; \Gamma_{out},
\]

\[
\delta u = -\delta w \frac{\partial u}{\partial \nu} \quad \text{on} \; \Gamma_c,
\]

\[
(2.5.3)
\]

which is set to zero to satisfy non-slip conditions between fluid and wall on the new boundary \( \Gamma'_c \). Being \( u'(x) = u'|_{\Gamma_c} \) we get:

\[
u \Delta \delta u + \nabla \delta p = 0 \quad \text{in} \; \Omega \subset \mathbb{R}^2,
\]

\[
\nabla \cdot \delta u = 0 \quad \text{in} \; \Omega,
\]

\[
\delta u = 0 \quad \text{on} \; \Gamma_u \backslash \Gamma_c,
\]

\[
\delta u = 0 \quad \text{on} \; \Gamma_{in},
\]

\[
\nu \frac{\partial \delta u}{\partial \nu} - \delta p \hat{n} = 0 \quad \text{on} \; \Gamma_{out},
\]

\[
\delta u = -\delta w \frac{\partial u}{\partial \nu} \quad \text{on} \; \Gamma_c,
\]

(2.5.3)

Now we consider (2.3.15) in the weak form and with the new right-hand-side \( -\nu \Delta u \) (instead of \( \nabla \times \nabla \times u \)), we multiply by \( (\delta u, \delta p) \) and integrate by parts, obtaining:

\[
\begin{cases}
\nu \int_{\Omega} \nabla p \cdot \nabla \delta u \, d\Omega - \nu \int_{\Gamma_c} \frac{\partial u}{\partial \nu} \cdot \delta u \, d\Gamma + \int_{\Omega} \sigma \nabla \cdot \delta ud\Omega = \nu \int_{\Omega} \nabla u \cdot \nabla \delta u \, d\Omega - \nu \int_{\Gamma_c} \frac{\partial u}{\partial \nu} \cdot \delta u \, d\Gamma,
\end{cases}
\]

(2.5.4)

Using equation (2.5.3) multiplied by \( (p, \sigma) \) we have:

\[
\begin{cases}
\nu \int_{\Omega} \nabla p \cdot \nabla \delta u \, d\Omega - \int_{\Omega} \delta p \nabla \cdot p \, d\Omega = 0, \\
\end{cases}
\]

(2.5.5)
and so, comparing (2.5.4) and (2.5.5) the only term surviving is

$$ -\nu \int_{\Gamma_e} \frac{\partial p}{\partial n} \cdot \delta u d\Gamma = \nu \int_{\Omega} \nabla u \cdot \nabla \delta u d\Omega - \nu \int_{\Gamma_e} \frac{\partial u}{\partial n} \cdot \delta u d\Gamma, $$

where $\delta u|_{\Gamma_e} = -\delta w \frac{\partial u}{\partial n}|_{\Gamma_e}$, so

$$ -\nu \int_{\Gamma_e} (\frac{\partial p}{\partial n} - \frac{\partial u}{\partial n}) \cdot \delta u d\Gamma = \nu \int_{\Omega} \nabla u \cdot \nabla \delta u d\Omega. \quad (2.5.6) $$

Finally from (2.5.2) and (2.5.6) we can rewrite (2.5.1) as:

$$ \delta J = \frac{1}{2} \nu \int_{\Gamma_e} \delta w \frac{\partial u}{\partial n}^2 d\Gamma + \nu \int_{\Gamma_e} \delta w \left( \frac{\partial p}{\partial n} - \frac{\partial u}{\partial n} \right) \cdot \frac{\partial u}{\partial n} d\Gamma, $$

or

$$ \delta J = \nu \int_{\Gamma_e} \delta w \left( \frac{\partial p}{\partial n} - \frac{1}{2} \frac{\partial u}{\partial n} \right) \cdot \frac{\partial u}{\partial n} d\Gamma. $$

In our algorithm we are interested in $J'(w)$ and recalling that

$$ \delta J = \int_{\Gamma_e} J'(w) \delta w d\Gamma $$

we can say that the new $\delta w|_{\Gamma_e} = -\rho J'(w) = -\rho \nu \left( \frac{\partial p}{\partial n} - \frac{1}{2} \frac{\partial u}{\partial n} \right) \cdot \frac{\partial u}{\partial n} = -\rho G(p, u, w)$. In this analysis we can better understand the role of the adjoint variables to get the gradient of cost functional and the new shape.

### 2.6 Some numerical results

In this section we present some results carried out by local boundary variations algorithm to optimize a preliminary bypass configuration, to obtain vorticity reduction and a feedback procedure to prevent wall shear stress oscillations with the functional (2.2.9). Numerical tests and simulations have been carried out using *Bamg* [55], a Bi-dimensional Anisotropic Mesh Generator and *FreeFem* [54], a finite element Library developed at INRIA, the French National Institute for Research in Computer Science and Control, with the development of algorithms based on control theory and shape optimization. Several remarks carried out from intensive preliminary numerical investigations are in order:

- The wall curvature in the host artery (to model arteries near the heart muscle) increases vorticity.

- The graft angle of the incoming branch of the bypass influences vorticity, by reducing the angle also vorticity is reduced in the down-field observation zone (keeping fixed the configuration provided by other quantities such as arterial and bypass diameter).

- Bypass configuration causes an increase of vorticity in the downfield area of about 35% than the previous idealized 2D configuration without bypass bridge and coronary occlusion.
After obtaining the new optimal shape for the incoming branch in 25 iterations, vorticity reduction, in our specific modelling context led by optimal control, is about of 45% in comparison with the starting idealized 2D configuration.

Figures 2.6-2.10 provide an account of numerical results and show the optimal shape of the bypass obtained using steady Stokes equations in an optimal control problem with cost functionals (2.3.12) and (2.3.13). As we can see the shape that is found resembles the Taylor patch configuration [26], a little cuffed shape. The effect of the Taylor patch is to reduce gradually the average velocity of the blood as it approaches the distal anastomosis, since the cross-sectional area of the bypass is steadily becoming larger. This prevents the sudden deceleration experienced in the conventional model with the fluid returning to the host vessel. There is a gradual reduction in the momentum of the blood while approaching the junction, in fact, the blood is guided more smoothly through the vessel thanks to the gradually changing geometry. Flow disturbances are abated, undesirable flow separation at the toe of the bypass diminished.

Figure 2.6: Complete bypass configuration (bridge) before optimal shape design process: iso-velocity $[m s^{-1} \cdot 10^{-2}]$.

Figure 2.7: Complete bypass configuration at the end of shape optimization: iso-velocity, same colorbar of the previous picture.
2.7 Feedback procedures and results interpretation

A feedback procedure has been implemented by solving the unsteady Navier-Stokes equations in the original configuration as well as in the final configuration obtained after applying the optimal shape design process on a reduced model, based on steady Stokes equations. The full Navier-Stokes problem reads, for all $0 < t < T$:

\[
\begin{align*}
\frac{\partial u}{\partial t} - \nu \Delta u + (u \cdot \nabla)u + \nabla p &= f \quad \text{in } \Omega, \\
\nabla \cdot u &= 0 \quad \text{in } \Omega, \\
u \frac{\partial u}{\partial n} - \nu \hat{n} &= 0 \quad \text{on } \Gamma_{\text{out}}, \\
u \frac{\partial u}{\partial n} - \hat{n} &= 0 \quad \text{on } \Gamma_{\text{in}}, \\
\end{align*}
\]  

(2.7.1)
We used Hagen-Poiseuille inflow pulsatory condition (Womersley profile with period $T = 1s$). Time discretization is based on backward Euler method with a time step $\Delta t = 0.01s$. Figures 2.11-2.12 show unsteady flows at different time steps.

During the feedback procedure we used the functional (2.2.9) to provide useful information about unsteady fluid dynamics phenomena such as wall shear stress oscillations during pulsatory systolic and diastolic phases. This quadratic functional is complete because it keeps into consideration WSS variations in time and the functional is not pointwise but defined on $\Gamma_c$, the vascular wall we are modelling by optimal shape design. In the specific modelling context that we have considered, shape optimization process guarantees a reduction of 45% of vorticity using steady Stokes flow model and a reduction of 25% in wall shear stress oscillations in time (on the vascular wall $\Gamma_c$ we are modelling). In general, the flow at the distal junction exhibits considerable spatial and temporal variations. With the adaptation operated by shape optimization technique we find, as already seen, a graft which resembles the Taylor patch configuration in which the anastomotic flow is less disturbed, a less adverse shear stress distribution prevails and furthermore flow separation is reduced [26]. Intimal thickening hyperplasia, which causes restenosis, should be alleviated at the toe in the new bypass configuration proposed.

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{figure2.11.png}
\caption{Fluid dynamics unsteady phenomena into the incoming branch of bypass at $t = 0.3s$, velocity $[m/s \cdot 10^{-2}]$.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{figure2.12.png}
\caption{Fluid dynamics unsteady phenomena into the incoming branch of bypass at $t = 0.5s$.}
\end{figure}

In Figure 2.13 an integral quantity of interest is plotted, it is the wall shear stress variation in time $\Sigma(t)$ (2.2.7) on $\Gamma_c$, its calculation has been done to get information about functional $J_c$ (2.2.9) which provides integral information in time $(0, T)$ for $\Sigma(t)$. As we can see the oscillatory behavior of $\Sigma(t)$ on $\Gamma_c$ is strongly reduced after applying shape optimization.
It is also possible to develop algorithm to apply optimal control and shape optimization to fully unsteady incompressible Navier-Stokes equations (see Laporte [75] and Le Tallec [76]) and eventually to consider the coupled fluid-structure problem (Moubachir and Zolésio [100]).

In our case at this step the interest is to develop efficient algorithms based on reduced fluid models to provide preliminary design indications concerning surgical prosthesis realizations to be verified in feedback and supported by other efficient optimization methods such as reduced basis applied to parametrized domains.

We have also seen that our control approach based on a vorticity functional is related (and equivalent from a physical point of view) with the one based on the minimization of viscous dissipation. For this reason we can conclude that the use of the unsteady functional $J_r$ related with shear stresses (and so the gradient of velocity, considering a Newtonian fluid) in feedback is strictly related with the functional used in the shape optimization process.

As conclusion of this chapter we introduce a theoretical explanation of results, achieved after the optimization approach, and related with fluid mechanics phenomena in the bypass, especially in the host artery.

Applying the 2D version of Stokes Theorem (Green Theorem, see for example Adams [3]) we can say that the cost functional calculated in down-field zone of the bypass is given by:

$$J(u) = \int_{\Omega_{wd}} \left( \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) d\Omega = \int_{\Gamma_{wd}} (u dx + v dy) d\Gamma,$$

in our case the boundary conditions on $\Gamma_{wd}$ in the down-field zone provides a contribution only at the inflow (the intersection between the bypass and the host artery) and at the outflow (of bypass configuration considered). See scheme in Figure 2.14.
The intersection zone between the bypass and the artery is becoming larger and the mean velocity of the blood is diminished so the term $udx$ is becoming smaller and if the fluid is less disturbed also the term $vdy$ (representing a distorsion of the main stream) is small in the down-field zone. The relationship (2.7.2) tells us that while studying (stationary) vorticity we can get information on velocity field and viscous dissipation as well. In this framework we can unify results already available in literature and trying to better understand our own optimization results.
Chapter 3

Shape Design by Optimal Control using Small Perturbation Theory

In this chapter we present another shape design approach using small perturbation theory. The theory of optimal control based on adjoint formulation is applied to a state problem representing the flow perturbation induced by shape variation (first order problem) in order to optimize the incoming branch of the bypass (the toe) into the coronary. We use a method based on the reference domain formulation described by a map whose coefficients have to be optimized so that we show how to solve a shape design problem without dealing with shape functions and boundary variations (that would invoke mesh adaptation, re-meshing and smoothing). Rather, we will have to solve an optimal control problem on the coefficients of the transformation map.

3.1 An alternative approach to local boundary variations

This second proposed approach for local shape design is based on a map from the real domain Ω to a (rectangular) reference one Ō using a variable transformation:

\[ \tilde{x} = x, \quad \tilde{y} = \frac{1}{f(x, \varepsilon)} y, \quad (3.1.1) \]

where \( f(x, \varepsilon) \) represents, for example, in our case, the upper shape of the bypass and can be developed as

\[ f(x, \varepsilon) = f_0(x) + \varepsilon f_1(x) + \varepsilon^2 f_2(x) + ..., \quad (3.1.2) \]

being \( f_0(x) \) the given (non-optimized) shape and \( \varepsilon \) is a small parameter. Assuming that the state problem has a solution \( u, p \) that is infinitely differentiable with respect to \( \varepsilon \):

\[ u = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + ..., \]

\[ p = p_0 + \varepsilon p_1 + \varepsilon^2 p_2 + ..., \]

and using small perturbation techniques (see Van Dyke [154]), we can derive the equations for \( u_k, p_k \) starting from the state problem and then mapping \( \Omega \) to the reference domain \( \tilde{\Omega} \). Here
(u₀, p₀) represents the solution of the state problem in the domain Ω₀ with upper boundary \( f₀(x) \), while (u₁, p₁) represents the flow perturbation due to the first order shape variation \( f₁(x) \). At this point we can use optimal control techniques to solve the problem for \( u₁, p₁ \) (the first corrections), the function \( f₁(x) \) which represents a perturbation in the shape \( f₀(x) \) (weighted by \( ε \)) is another unknown for the problem and is used as control variable. In this context Masmoudi et al. [47] have investigated a complementary approach based on high order derivatives and Taylor expansion of cost functional with respect to shape parameters.

The structure of this chapter is the following: in Section 3.2 we recall the problem based on the same fluid model of the previous chapter (a generalization to unsteady flows will be introduced in the next chapter) and we develop small perturbation techniques in Section 3.3, then we deal with shape optimization problem in Section 3.4 and with the optimal control framework in Section 3.5. Then we study theoretical aspects of the problem such as existence and uniqueness of solutions in Section 3.6, we propose an algorithm in Section 3.7 and we describe numerical results in Section 3.8. A comparison between local boundary variations method and small perturbation is provided in Section 3.9.

3.2 Problem statement

Let Ω be a bounded domain of \( \mathbb{R}^2 \), \( Γ ≡ ∂Ω \) is the boundary of Ω, \( Ω = Ω ∪ ∂Ω \), \( x = (x, y)^T \) is a point of \( Ω \). In the sequel aggregation of vector quantities \( u \) with scalar quantities \( p \) are indicated with \( Q \) (\( Q = (u, p) \)), \( Φ \) or \( \Phi \).

We consider again an idealized, two-dimensional bypass bridge configuration as in Figure 3.1 and the domain of Figure 3.2, where the dotted line represents the geometry of the complete anastomosis; \( Γ_{w₂} \) is the section of the original artery, \( Γ_{in} \) is the new anastomosis inflow after bypass surgery, \( Γ_{out} \) is the anastomosis outflow. We consider a boundary value problem for the Stokes equations, see for example Zeytounian [169], used as in the previous chapter to

Figure 3.1: Idealized, 2-D bypass bridge configuration (left) and detail of the sensible part for the optimization process (right), where the dotted curve represents a possible shape variation.
§3.2. PROBLEM STATEMENT

A model low Reynolds blood flow in the optimization process. The problem reads: find \( u, p \) s.t.

\[
\begin{cases}
-\nu \Delta u + \nabla p = f & \text{in } \Omega, \\
\nabla \cdot u = 0 & \text{in } \Omega, \\
\n u = u_{in} & \text{on } \Gamma_{in}, \\
\n u = 0 & \text{on } \Gamma_{w1} \cup \Gamma_{w3}, \\
\n -p\hat{n} + \nu \frac{\partial u}{\partial n} = g_{out} & \text{on } \Gamma_{out} \cup \Gamma_{w2},
\end{cases}
\]

(3.2.1)

where \( \hat{n} = (n_1, n_2)^T \) is the outward unit normal vector on \( \Gamma \), \( f = f(x, y) \), \( u_{in} = u_{in}(x, y) \), \( g_{out} = g_{out}(x, y) \) are given vector functions, \( \nu = \text{const} > 0 \) and \( u_f = \{u_{in} \text{ on } \Gamma_{in}, 0 \text{ on } \Gamma_{w1} \cup \Gamma_{w3}\} \). In the following we may need to impose some additional restriction on \( p \) (for example \( \int_\Omega p \, d\Omega = 0 \) if \( \Gamma_{in} = \Gamma \)).

The weak statement of (3.2.1) reads: find \( u \in (H^1(\Omega))^2, p \in L^2(\Omega) \) s.t.

\[
\begin{cases}
a(u, v) = b(p, v) + G(v) & \forall v \in X, \\
b(q, u) = 0 & \forall q \in L^2(\Omega), \\
u = u_f & \text{on } \Gamma_{in} \cup \Gamma_{w1} \cup \Gamma_{w3},
\end{cases}
\]

(3.2.2)

where with \( v \) and \( q \) we indicate test functions and:

\[
a(u, v) = \int_{\Omega} \nu \nabla u \cdot \nabla v \, d\Omega,
\]

\[
b(p, v) = \int_{\Omega} p \nabla \cdot v \, d\Omega, \\
G(v) = \int_{\Omega} f \cdot v \, d\Omega + \int_{\Gamma_{out} \cup \Gamma_{w2}} g_{out} \cdot v \, d\Gamma,
\]

\[
X := \{v : v \in (H^1(\Omega))^2, v = 0 \text{ on } \Gamma_{in} \cup \Gamma_{w1} \cup \Gamma_{w3}\}.
\]

For further mathematical aspects of the problem see, for example, Lions [85]. The subset \( \Gamma_{c,\varepsilon} \) of \( \Gamma_{w1} \) is parametrized by a function \( f(x, \varepsilon) \) of \( x \in [x_1, x_2] \) and of a small parameter \( \varepsilon \in [-\varepsilon_0, \varepsilon_0], \varepsilon_0 = \text{const} \). More precisely we assume that \( f(x, \varepsilon) \) can be developed as follows:

\[
f(x, \varepsilon) = f_0(x) + \varepsilon f_1(x) + \varepsilon^2 f_2(x) + \ldots,
\]

(3.2.3)
where \( f_k \in W^{1,\infty}(x_1, x_2) \), for \( k = 0 \), (we recall that \( W^{1,\infty}(x_1, x_2) \) is the space of functions \( f_k \in L^\infty(x_1, x_2) \) such that the distributional derivative of first order of \( f_k \) belongs to \( L^\infty(x_1, x_2) \)) while for \( k \geq 1 \) \( f_k \in W^{0,\infty}_1(x_1, x_2) \) with \( f_k(x_1) = f_k(x_2) = 0, k \geq 1 \). Here the function \( f_0(x) > 0 \) describes the original subset \( \Gamma_c,0 \) (corresponding to \( \Gamma_c,\varepsilon \) for \( \varepsilon = 0 \)) of the boundary \( \Gamma_{w0} \equiv \partial \Omega_0 \) of a given non-optimized domain \( \Omega_0 \), see Figure 3.3 (in perturbation theory this is often called "unperturbed domain"), while \( f_k(x), k \geq 1 \), could be unknown when dealing with control problem (see Section 3.4).

Although \( a(., .), b(., .) \) and \( G(.) \) depend on the parametrization \( f \) of the part \( \Gamma_c,\varepsilon \), this dependence has been understood for simplicity of notations.

### 3.3 The problem for the perturbation functions

Let us introduce the reference (simple-shaped) domains \( \tilde{\Omega}_1 = \{0 < \tilde{x} < A, 0 < \tilde{y} < \tilde{\beta}_1 = \beta\} \), \( \tilde{\Omega}_2 = \{0 < \tilde{x} < A, -\tilde{\beta}_2 < \tilde{y} < 0\} \), and \( \tilde{\Omega} = \tilde{\Omega}_1 \cup \tilde{\Omega}_2 \) (see Figure 3.4). Then we assume that \( f(x, \varepsilon) > 0 \) and consider the following variable transformation:

\[
T_f : \tilde{\Omega}_1 \cup \tilde{\Omega}_2 \to \tilde{\Omega}, \quad \tilde{x} = T_f(x),
\]

such that \( T_f \) is the identity in \( \Omega_2 \), while \( T_f(x, y) = (x, \frac{\beta}{f(x,y)} y) \) in \( \Omega_1 \).

We set \( \tilde{\kappa} = (\tilde{x}, \tilde{y})^T \) and define

\[
\tilde{u}(\tilde{\kappa}) := u \circ T_f^{-1}(\tilde{\kappa}) = u(\tilde{x}, \tilde{y}f(\tilde{x}, \varepsilon)/\beta).
\]
where \( \mathbf{u} = (\tilde{u}, \tilde{v})^T \). Then,
\[
dx \, dy = \frac{f(\tilde{x}, \varepsilon)}{\beta} \, d\tilde{x} \, d\tilde{y}
\]
and the following relations hold:
\[
\begin{align*}
\frac{\partial \phi}{\partial y}(\tilde{x}) &= \frac{\partial \phi}{\partial y}(\tilde{x}) - \frac{\partial L(\tilde{x}, \varepsilon)}{\partial y} \frac{\partial \phi}{\partial y}(\tilde{x}), \\
\frac{\partial \phi}{\partial x}(\tilde{x}) &= \frac{\partial \phi}{\partial x}(\tilde{x}) - \frac{\partial L(\tilde{x}, \varepsilon)}{\partial x} \frac{\partial \phi}{\partial y}(\tilde{x}) (\text{with } f_x := \frac{df}{dx}),
\end{align*}
\]
(3.3.1)
\[
\begin{align*}
\hat{D}(f)\mathbf{u}(\tilde{x}) &= ((\nabla \cdot \mathbf{u}) \circ T^{-1}_f)(\tilde{x}) = \frac{\partial u}{\partial x} - \frac{\partial \tilde{u}}{\partial y} - \frac{\partial f_x(\tilde{x})}{\partial y} \frac{\partial \phi}{\partial y}(\tilde{x}) + \frac{\partial f_x(\tilde{x})}{\partial x} \frac{\partial \phi}{\partial x}(\tilde{x}), \\
\hat{R}(f)\mathbf{u}(\tilde{x}) &= ((\nabla \times \mathbf{u}) \circ T^{-1}_f)(\tilde{x}) = \frac{\partial u}{\partial x} - \frac{\partial \tilde{v}}{\partial y} - \frac{\partial f_x(\tilde{x})}{\partial y} \frac{\partial \phi}{\partial x}(\tilde{x}) - \frac{\partial f_x(\tilde{x})}{\partial x} \frac{\partial \phi}{\partial y}(\tilde{x}),
\end{align*}
\]
(3.3.2)
Then in \( \hat{\Omega} \) we have:
\[
\hat{D}(f)\mathbf{u} = m_2 \hat{\nabla} \cdot \mathbf{u} + m_1 \hat{D}(f)\mathbf{u}, \quad \hat{R}(f)\mathbf{u} = m_2 \hat{\nabla} \times \mathbf{u} + m_1 \hat{R}(f)\mathbf{u},
\]
where \( \hat{\nabla} \phi := \left( \frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y} \right) \), while \( m_s \) is the characteristic function of \( \Omega_s \) \((s = 1, 2)\). Now the functional spaces \( X \) and \( L^2 \) are defined on the domain \( \hat{\Omega} \). To simplify our notation from now on we will set (unless otherwise specified):
\[
\mathbf{x} = \tilde{x}, \quad \mathbf{u}(x, y) = \tilde{u}(\tilde{x}, \tilde{y}), \quad u = \tilde{u}, \quad v = \tilde{v}, \ldots, \quad \mathcal{D} = \hat{\mathcal{D}}, \mathcal{R} = \hat{\mathcal{R}}, \quad \Omega \equiv \hat{\Omega}, \quad \Gamma_{w_k} \equiv \hat{\Gamma}_{w_k},
\]
\[
\Omega_1 \equiv \hat{\Omega}_1, \Omega_2 \equiv \hat{\Omega}_2
\]
Then problem (3.2.2) in the new variables and in the new domains reads as follows:
\[
\begin{align*}
\left\{
\begin{array}{ll}
a(f; \mathbf{u}, \mathbf{v}) &= b(f; p, v) + G(f; v) \quad \forall \mathbf{v} \in X, \\
b(f; q, \mathbf{u}) &= 0 \quad \forall q \in L^2(\Omega), \\
\mathbf{u} &= \mathbf{u}_f \quad \text{on } \Gamma_{in} \cup \Gamma_{w_1} \cup \Gamma_{w_2}.
\end{array}
\right.
\end{align*}
\]
(3.3.3)
We have emphasized the dependence of \( a(f; \cdot, \cdot) \), \( b(f; \cdot, \cdot) \), and \( G(f; \cdot) \) on the shape function \( f \). Precisely, we have:
\[
a_1(f; \mathbf{u}, \mathbf{v}) = a_1(f; \mathbf{u}, \mathbf{v}) + a_2(\mathbf{u}, \mathbf{v}),
\]
\[
a_1(f; \mathbf{u}, \mathbf{v}) = \int_{\Omega_1} \left( \frac{f \nu}{\beta} \left( \frac{\partial u}{\partial x} - \frac{f_x u}{\partial y} \right) \cdot \left( \frac{\partial v}{\partial x} - \frac{f_x v}{\partial y} \right) + \frac{\beta^2}{\beta} \frac{\partial u}{\partial x} \cdot \frac{\partial v}{\partial y} \right) d\Omega,
\]
\[
a_2(\mathbf{u}, \mathbf{v}) = \int_{\Omega_2} \left( \frac{\nu}{\beta} \frac{\partial \mathbf{u}}{\partial x} \cdot \frac{\partial \mathbf{v}}{\partial y} + \frac{\partial \mathbf{u}}{\partial y} \cdot \frac{\partial \mathbf{v}}{\partial y} \right) d\Omega,
\]
\[
b_1(f; p, \mathbf{v}) = \int_{\Omega_1} \left( \frac{f \nu}{\beta} D(f) \mathbf{v} d\Omega \right) + b_2(p, \mathbf{v}) = \int_{\Omega_2} \nu \hat{\nabla} \cdot \mathbf{v} d\Omega,
\]
\[
G_1(f; \mathbf{v}) = \int_{\Omega_1} \left( \frac{f}{\beta} \mathbf{v} \cdot d\Omega \right) + \int_{(\Gamma_{out} \cup \Gamma_{w_2}) \cap \partial \Omega_1} \mathbf{g}_{out} \cdot \mathbf{v} d\Gamma,
\]
\[
G_2(\mathbf{v}) = \int_{\Omega_2} \mathbf{f} \cdot \mathbf{v} d\Omega + \int_{(\Gamma_{out} \cup \Gamma_{w_2}) \cap \partial \Omega_2} \mathbf{g}_{out} \cdot \mathbf{v} d\Gamma.
\]
Assume that the problem (3.3.3) has a solution \( u, p \) that is infinitely differentiable with respect to \( \varepsilon \):

\[
\begin{align*}
u &= u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \ldots \\
p &= p_0 + \varepsilon p_1 + \varepsilon^2 p_2 + \ldots
\end{align*}
\]  

(3.3.4)

where \( p_k \in L^2(\Omega), u_k \in X, k \geq 0 \). Using (3.2.3), (3.3.4) and small perturbation techniques we can derive the equations for \( u_k, p_k, k \geq 0 \). In particular, for \( k = 0 \) \( u_0 \) and \( p_0 \) satisfy the following zero-order problem, for which velocity and pressure are obtained without considering a shape variation:

\[
\begin{align*}
a(f_0; u_0, v) &= b(f_0; p_0, v) + G(f_0; v) \forall v \in X, \\
b(f_0; q, u_0) &= 0 \forall q \in L^2(\Omega), \\
u_0 &= u_f \text{ on } \Gamma_{in} \cup \Gamma_{w_1} \cup \Gamma_{w_3}.
\end{align*}
\]  

(3.3.5)

Note that the test functions \( v, q \) in (3.3.3) can be assumed to be independent of \( \varepsilon \) in the sequel.

Correspondingly we define:

\[
\mathcal{R}_{obs,0} := \mathcal{R}(f_0)u_0
\]  

(3.3.6)

representing the vorticity mapped into the simple (reference) domain. For \( k = 1 \) the functions \( u_1 \) and \( p_1 \), representing small perturbations in the velocity and pressure fields, respectively, are the solution of the equations:

\[
\begin{align*}
a(f_0; u_1, v) &= b(f_0; p_1, v) + \frac{\partial}{\partial \varepsilon} b(f_0; p_0, v) |_{\varepsilon=0} + \\
&+ \frac{\partial}{\partial \varepsilon} G(f; v) |_{\varepsilon=0} - \frac{\partial}{\partial \varepsilon} a(f_0; u_0, v) |_{\varepsilon=0} \forall v \in X, \\
b(f_0; q, u_1) &= \frac{\partial}{\partial \varepsilon} b(f_0; q, u_0) |_{\varepsilon=0} = 0 \forall q \in L^2(\Omega), \\
u_1 &= 0 \text{ on } \Gamma_{in} \cup \Gamma_{w_1} \cup \Gamma_{w_3},
\end{align*}
\]  

(3.3.7)

where we have some new terms depending on the new configuration \( (f_1) \) but evaluated with respect to the zero-order solution \( u_0 \) and \( p_0 \):

\[
\frac{\partial}{\partial \varepsilon} b(f_0; p_0, v) |_{\varepsilon=0} = b(f_1; p_0, v) = \int_{\Omega_1} f_1 \beta p_0 D(f_0) v d\Omega + \int_{\Omega_1} f_0 \beta p_0 D(f_1) v d\Omega,
\]

with

\[
D_j(f_1, v) = \frac{\partial}{\partial \varepsilon} D(f) v |_{\varepsilon=0} = -\left[y \left( \frac{f_{1,xf_0}-f_{0,xf_1}}{f_0^3} \right) \frac{\partial u}{\partial y} + \frac{\beta f_1}{f_0} \frac{\partial v}{\partial y} \right]
\]

\[
D_j(f_1, u_0) = \frac{\partial}{\partial \varepsilon} D(f) u_0 |_{\varepsilon=0} := D_j f_1 \text{ in the sequel},
\]

\[
G(f; v) |_{\varepsilon=0} = G_1(f_1; v) = \int_{\Omega_1} f_1 f \cdot v d\Omega,
\]

\[
\frac{\partial}{\partial \varepsilon} a(f_0; u_0, v) |_{\varepsilon=0} = a(f_1; u_0, v) = \int_{\Omega_1} f_{1,\nu} \left( \frac{\partial u_0}{\partial x} - \frac{y f_{0,x} \partial u_0}{f_0} \right) \frac{\partial v}{\partial x} - \frac{y f_{0,x} \partial v}{f_0} + \frac{\beta^2}{f_0^3} \frac{\partial u_0}{\partial y} \frac{\partial v}{\partial y} d\Omega + \\
&- \int_{\Omega_1} \frac{f_{0,\nu}}{f_0^3} \left( f_{1,xf_0}-f_{0,xf_1} \right) \left( \frac{\partial u_0}{\partial y} \frac{\partial v}{\partial x} - \frac{y f_{0,x} \partial v}{f_0} \right) + \left( \frac{\partial u_0}{\partial x} - \frac{y f_{0,x} \partial u_0}{f_0} \right) \frac{\partial v}{\partial y} d\Omega.
\]
3.4. THE SHAPE OPTIMIZATION PROBLEM

\[ - \int_{\Omega_1} \frac{f_0}{\beta} \left( \frac{2\beta^2 f_1}{f_0^3} \right) \frac{\partial u_0}{\partial y} \cdot \frac{\partial v}{\partial y} \, d\Omega. \]

Therefore the problem for \( u_1, p_1 \) reads as follows: find \( u_1 \in X, p_1 \in L^2(\Omega) \) s.t.:

\[
\begin{cases}
  a(f_0; u_1, v) - b(f_0; p_1, v) = b(f_1; f_0; p_0; v) + G_1(f_1; v) - a(f_1; u_0, v) \forall v \in X, \\
  b(f_0; q; u_1) + b(f_1; q; u_0) = 0 \forall q \in L^2(\Omega),
\end{cases}
\]

This is a generalized Stokes problem (see Galdi [40]) representing the perturbations in the system due to the induced shape variation. By a similar technique we can derive the equations for \( u_k, p_k \) with \( k \geq 2 \). However we will not further carry on this development in this work. At the end of the chapter in Section 3.10 we are going to discuss how to deal with higher orders and non-linear problems. The general guidelines to build the problem at different order is based on the idea of replacing in the state equations the development of velocity, pressure and shape and then solving the problem at various orders of \( \varepsilon \), see also Hinch [56] and Kevorkian [71]. In particular, our shape optimization and control problem will be considered as a first-order problem and it will deal with flow perturbation.

### 3.4 The shape optimization problem

Suppose now that in (3.3.7) the function \( f_1(x) \) is unknown and so are \( u_1, p_1 \). To complete problem (3.3.7) we will have to formulate some “additional equations”, or, alternatively, we should require that \( f_1 \) be determined by minimizing a suitable cost functional.

Problem (3.2.2) can be supplemented by the “additional equation”:

\[ C(f, u, p) = 0 \]  

where \( C \) is an operator (linear or nonlinear) defined on \( H^1_0(x_1, x_2) \times X \times L^2(\Omega) \). (We consider now \( f \in H^1_0 \) for convenience). We assume \( C \) have a regular dependence on \( f, u, p \). Using the representations (3.2.3) and (3.3.4) we derive from (3.4.1) the following equation:

\[ C(f, u, p) = C_0(f_0, u_0, p_0) + \varepsilon C_1(f_1, u_1, p_1) + \mathcal{O}(\varepsilon^2) = 0, \quad \forall \varepsilon \in [-\varepsilon_0, \varepsilon_0] \]  

where

\[ C_1(f_1, u_1, p_1) := \frac{\partial C}{\partial \varepsilon}(f, u, p)|_{\varepsilon=0}. \]  

If we assume, for example, that the data of our problems are such that \( C_0(f_0, u_0, p_0) = 0 \), then we can use

\[ C_1(f_1, u_1, p_1) = 0 \]  

as additional equation to complete (3.3.7). An alternative approach would consist in replacing the exact controllability equation (3.4.4) by the following minimization problem:

\[ \inf_{f_1} \int_{\Omega} \frac{f_0}{\beta} |C_1(f_1, u_1, p_1)|^2 d\Omega, \]  

where we assume that \( C_1 \) has image in \( L^2(\Omega) \). Note that (3.4.5) is a weak statement of (3.4.4).
In the next sections we apply the approach described above for the completion of (3.3.7); moreover we use the following special choice of (3.4.1):

$$\mathcal{C}(f, u) := (\nabla \times u) \circ T_f^{-1} - R_{obs,\varepsilon} \text{ in } \Omega_{wd} \subseteq \Omega,$$

(3.4.6)

where $\Omega_{wd}$ is a suitable subset of $\Omega$ in which we want our additional equation (or our “control”) to take place. In our application $\Omega_{wd}$ is the down-field zone of the domain, where the bypass is hosted into the artery.

Moreover

$$R_{obs,\varepsilon} = R_{obs,0} + \varepsilon R_{obs,1} + \varepsilon^2 R_{obs,2} + \ldots,$$

with $R_{obs,0} := ((\nabla \times u_0) \circ T_{f_0}^{-1})$.

(3.4.7)

Then we have: $\mathcal{C}_0(f_0, u_0) = 0$ (to consider the additional equation only for the first-order problem), while the equation (3.4.4) reads:

$$\mathcal{C}_1(f_1, u_1) = R(f_0)u_1 + m_1 \mathcal{R}_f f_1 - R_{obs,1} = 0 \text{ in } \Omega_{wd},$$

(3.4.8)

where

$$R(f_0)u_1 = (\nabla \times u_1) \circ T_{f_0}^{-1}(x, y) = \frac{\partial v_1}{\partial x} - \frac{y f_{0,x} \partial v_1}{f_0} - \frac{\beta \partial u_1}{f_0},$$

$$\mathcal{R}_f f_1 := \mathcal{R}_f(f_1, u_0) = -y \left(\frac{f_{1,x} f_0 - f_{0,x} f_1}{f_0^2}\right) \frac{\partial v_1}{\partial y} + \frac{\beta f_1 \partial u_0}{f_0^2},$$

which represents two contributes to vorticity due to flow perturbation ($u_1$) and shape perturbation $f_1$, respectively. Finally, $R_{obs,1}$ is a given function (for example $R_{obs,1} = -R_{obs,0}$ to reduce vorticity by a shape optimization process on $f_1$ and flow perturbation $u_1$ and $p_1$).

Therefore we have the problem on the domain reported in Figure 3.4: find $u_1 \in X$, $p_1 \in L^2(\Omega)$, $f_1 \in H^1_0(x_1, x_2)$ s.t.

$$\begin{cases}
a(f_0; u_1, v) = b(f_0; p_1, v) + b(f_1; p_0, v) + G_1(f_1; v) - a(f_1; u_0, v) & \forall v \in X, \\
b(f_0; q, u_1) + b(f_1; q, u_0) = 0 & \forall q \in L^2(\Omega), \\
R(f_0)u_1 + m_1 \mathcal{R}_f f_1 - R_{obs,1} = 0 & \text{in } \Omega_{wd}.
\end{cases}$$

(3.4.9)

Problem (3.4.9) is an “exact controllability problem”. These kind of problems have solutions in some particular cases only. For this reason we replace (3.4.9) by the following optimal control problem: find $u_1 \in X$, $p_1 \in L^2(\Omega)$, $f_1 \in H^1_0(x_1, x_2)$ s.t.

$$\begin{cases}
a(f_0; u_1, v) - b(f_0; p_1, v) = b(f_1; p_0, v) + G_1(f_1; v) - a(f_1; u_0, v) & \forall v \in X, \\
b(f_0; q, u_1) + b(f_1; q, u_0) = 0 & \forall q \in L^2(\Omega), \\
\inf_{f_1} \frac{\alpha}{2} \|f_1\|_{H^1_0(x_1, x_2)}^2 + \gamma_1 J_1(f_1, v_1),
\end{cases}$$

(3.4.10)

where

$$J_1(f_1, u_1) = \frac{1}{2} \int_\Omega \frac{m_{wd} f_0}{\beta} (R(f_0)u_1 + m_1 \mathcal{R}_f f_1 - R_{obs,1})^2 d\Omega,$$

(3.4.11)

$\alpha = \text{const} \geq 0$ is a small regularization parameter, $\gamma_1 > 0$ is a weight coefficient, $m_{wd}$ is the characteristic function of $\Omega_{wd}$. Problem (3.4.10) is called “first-order problem” since it provides the solution of the first order terms in the developments (3.2.3) and (3.3.4).
Note that the third equation from (3.4.9) is considered in (3.4.10) in the least square sense; then when \( \alpha = 0 \) (3.4.10) represents the weak statement of problem (3.4.9). Otherwise the solution \( v_1 = v_1(\alpha) \), \( p_1 = p_1(\alpha) \), \( f_1 = f_1(\alpha) \) of (3.4.10) will represent an approximate (regularized) solution of (3.4.9).

We will also consider a generalized optimal control problem, still given by (3.4.10); however now instead of \( J_1 \) we use

\[
J(f_1, u_1, p_1) = \gamma_1 J_1(f_1, u_1) + \gamma_2 J_2(f_1, u_1, p_1),
\]

Here \( \gamma_2 = \text{const} \geq 0 \) is a weight coefficient, while \( J_2(f_1, u_1, p_1) \) is an additional functional that we assume to be quadratic. A few examples of \( J_2(f_1, u_1, p_1) \) follow.

**Example 3.4.1**

\[
J_2(f_1, u_1, p_1) := J_2(u_1, p_1) = \frac{1}{2} \left( ||p_1 - p_{\text{out},1}||^2_{L^2(\Gamma_{\text{out}})} + \int_{\Gamma_{\text{out}}} |u_1 - u_{\text{out},1}|^2 d\Gamma \right) \tag{3.4.12}
\]

where \( p_{\text{out}}, u_{\text{out}} \) are given.

**Example 3.4.2** Now let \( \Omega_{\text{obs}} \) be a “small” subdomain of \( \Omega \) with boundary \( \Gamma_{\text{obs}} \) and \( \Gamma_{\text{w3}} \subset \Gamma_{\text{obs}} \):

\[
\Omega_{\text{obs}} = (0, A) \times (-\beta_2, -\beta_2 + \delta), \ \text{with} \ 0 < \delta \ll \beta_1.
\]

![Figure 3.5: Subdomain \( \Omega_{\text{obs}} \).](image)

If \( \Omega = \Omega_1 \) (i.e. \( \beta_2 = 0 \)), we can take

\[
J_2(f_1, u_1, p_1) := J_2(u_1) = \frac{1}{2} \int_{\Omega_{\text{obs}}} |u_1 - u_{\text{obs},1}|^2 f_0 d\Omega, \tag{3.4.13}
\]

where \( u_{\text{obs}} \) is a prescribed vector-function.

**Example 3.4.3** Take \( \Omega_1 = \Omega \), and use: \( \int_{\Omega} p_1 d\Omega = 0 \),

\[
J_2(f_1, u_1, p_1) := J_2(u_1) = \frac{1}{2} \int_{\Gamma_{\text{w3}}} |\frac{\partial u_1}{\partial n} - g_{1,\text{obs}}|^2 d\Gamma, \tag{3.4.14}
\]
CHAPTER 3. SHAPE DESIGN USING SMALL PERTURBATION

for a given \(g_{1,\text{obs}}\).

Example 3.4.4 Finally we may consider:

\[
\Omega_1 = \Omega,
\]

\[
J_2(f_1, u_1, p_1) := J_2(u_1) = \frac{1}{2} \int_{\Omega_{\text{obs}}} \beta f_0 \frac{\partial u_1}{\partial y} - \|g_{2,\text{obs}}\|^2 \ d\Omega,
\]

(3.4.15)

with a given \(g_{2,\text{obs}}\).

Remark 3.4.1 The term \(J_2\) can be considered as an “overdetermination” of the problem, as in principle it allows the control of further quantities of physical interest. Using such a \(J_2\) makes easier to prove uniqueness but more troublesome to analyze existence for the optimal control problem.

3.5 The variational equations of the optimal control problem

While considering (3.4.10) we have taken into consideration the simple domain \(\tilde{\Omega}\) of Figure 3.4. Another possibility consists of using the new variable transformation

\[
x = T_{f_0}^{-1}(\tilde{x}), \ \tilde{x} \in \tilde{\Omega}, \ x \in \Omega_0,
\]

(3.5.1)

which is the identity in \(\tilde{\Omega}_2\), while \(T_{f_0}^{-1}(\tilde{x}, \tilde{y}) = (\tilde{x}, f_0(\tilde{x}) \tilde{y})\) in \(\tilde{\Omega}_1\). After applying (3.5.1) we will work in the “unperturbed” domain \(\Omega_0\) (see Figure 3.6) where the expressions for the bilinear forms in (3.4.10) become simpler. Indeed, with the help of transformation (3.5.1), problem

(3.4.10) reads upon its reformulation in \(\Omega_0\): find \(u := u_1, \ p := p_1, \ f := f_1\) s.t.

\[
\begin{cases}
    a_0(u, v) - b_0(p, v) = b_f(f; p_0, v) + G_1(f; v) - a_f(f; u_0, v) \ \forall v \in X \\
    b_0(q, u) + b_f(f; q, u_0) = 0 \ \forall q \in L^2(\Omega), \\
    \inf_f \frac{1}{2} \|f\|^2_{H_0^1(x_1, x_2)} + J(f, u, p),
\end{cases}
\]

(3.5.2)

\footnote{From now on we denote \(u_1 = u, \ p_1 = p, \ f_1 = f\) however we should keep in mind that now \(u, p, f\) represent the “first corrections” of \(u_0, \ p_0, \ f_0\) on the unperturbed domain.}
where

\[
a_0(u, v) = \int_{\Omega} \nu (\frac{\partial u}{\partial x} \cdot \frac{\partial \nu}{\partial x} + \frac{\partial u}{\partial y} \cdot \frac{\partial \nu}{\partial y}) d\Omega.
\]

\[
b_0(p, v) = \int_{\Omega} p \nabla \cdot v d\Omega,
\]

\[
b_f(f, p, v) = \int_{\Omega,1} p_0 D_f(f, v) d\Omega + \int_{\Omega,1} \frac{f}{f_0} p_0 \nabla \cdot v d\Omega,
\]

\[
\mathcal{D}_f(f, v) = -[y \left( \frac{f_x f_0 - f_0 x f}{f_0^2} \right) \frac{\partial u}{\partial y} + \frac{f}{f_0} \frac{\partial v}{\partial y}],
\]

\[
\mathcal{D}_f(f, v_0) = \mathcal{D}_f,
\]

\[
G_1(f; v) = \int_{\Omega,1} \frac{f}{f_0} f \cdot v d\Omega,
\]

\[
a_f(f; u_0, v) = \int_{\Omega,1} \frac{f \nu}{f_0} \nabla u_0 \cdot \nabla v d\Omega - \int_{\Omega,1} \nu y \left( \frac{f_x f_0 - f_0 x f}{f_0^2} \right) \left( \frac{\partial u_0}{\partial y} \cdot \frac{\partial v}{\partial x} + \frac{\partial u_0}{\partial x} \cdot \frac{\partial v}{\partial y} \right) d\Omega + \int_{\Omega,1} \frac{2 f \nu \partial u_0}{f_0} \cdot \frac{\partial v}{\partial y} d\Omega.
\]

\[
J_1(f, u) = \frac{1}{2} \int_{\Omega_0} m_{wd} \nabla \times u + m_1 \mathcal{R} f - \mathcal{R}_{\text{obs,1}}^2 d\Omega,
\]

\[
\mathcal{R} f := \mathcal{R}(f, u_0) = -y \left( \frac{f_x f_0 - f_0 x f}{f_0^2} \right) \frac{\partial u_0}{\partial y} + \frac{f}{f_0} \frac{\partial u_0}{\partial y}.
\]

\[
\nabla \times u = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}, \nabla \cdot u = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}
\]

and \( J_2(f, u, p) \) are given by corresponding expressions. At this step we have developed our problem in two levels (zero and first order) and rewritten the first order problem in a simpler way on the domain \( \Omega_0 \). In order to derive the operator form of problem (3.5.2) and formulate the problem in an optimal control setting we introduce the following functional spaces:

\[
\mathbb{X} \subseteq (L^2(\Omega))^2 \subseteq \mathbb{X}^*, \mathbb{H}^p \subseteq L^2(\Omega) \subseteq \mathbb{H}^p^*,
\]

\[
\mathbb{H}_f \subseteq L^2(x_1, x_2) \subseteq \mathbb{H}_f^*,
\]

\[
\mathbb{W} := \mathbb{X} \times \mathbb{H}_f \subseteq \mathbb{H}_0 := (L^2(\Omega))^2 \times L^2(\Omega) \subseteq \mathbb{W}^*,
\]

where \( \mathbb{X}^*, \mathbb{H}^p_*, \mathbb{H}^*_f \) and \( \mathbb{W}^* \) are the dual spaces of \( \mathbb{X}, \mathbb{H}^p, \mathbb{H}_f \) and \( \mathbb{W} \), respectively. Let us reformulate (3.5.2) in the following form: find \( \Phi := (u, p) \in \mathbb{W} = (\mathbb{X} \times \mathbb{H}_f), f \in \mathbb{H}_f, \) s.t.

\[
\left\{ \begin{array}{l}
\mathcal{L}(\Phi, \hat{\Phi}) = B(f, \hat{\Phi}) \forall \hat{\Phi} = (v, q) \in \mathbb{W}, \\
\inf_{f \in \mathbb{H}_f} \frac{1}{2} ||f||_{\mathbb{H}_f}^2 + J(f, \Phi),
\end{array} \right. \tag{3.5.3}
\]

where

\[
\mathcal{L}(\Phi, \hat{\Phi}) := a_0(u, v) - b_0(p, v) + b_0(q, v),
\]

\[
\mathcal{L}(\Phi, \hat{\Phi}) := a_0(u, v) - b_0(p, v) + b_0(q, v),
\]

\[
\mathcal{L}(\Phi, \hat{\Phi}) := a_0(u, v) - b_0(p, v) + b_0(q, v),
\]
\[ B(f, \hat{\Phi}) := b_f(f, p_0, v) + G_1(f, v) - a_f(f, u_0, v) - b_f(f, q, u_0). \]

Should \( \Phi \) be a solution of (3.5.3), then
\[ \alpha(f, \hat{f})_{\mathbb{H}_f} + \langle J'_b(f, \Phi), \Phi_f \rangle + \langle J'_f(f, \Phi), \hat{f} \rangle = 0, \tag{3.5.4} \]
for any \( \hat{f} \in \mathbb{H}_f \) (\( \hat{f} \) is the independent variation), where \( \Phi_f \in \mathbb{W} \) satisfies the following equation:
\[ \mathcal{L}(\Phi, \hat{\Phi}) = B(f, \hat{\Phi}) \forall \hat{\Phi} \in \mathbb{W}. \tag{3.5.5} \]

In (3.5.4), \( J'_b = \frac{\partial J}{\partial \Phi} \) and \( J'_f = \frac{\partial J}{\partial f} \) are partial derivatives of \( J \), while \( \langle Q, \Phi \rangle \) is the duality between \( \mathbb{W} \) and \( \mathbb{W}^* \) and \( \langle g, f \rangle \) the duality between \( \mathbb{H}_f \) and \( \mathbb{H}_f^* \). Then we can rewrite (3.5.3) as a system of “optimality conditions”:
\[ \begin{cases} \mathcal{L}(\Phi, \hat{\Phi}) = B(f, \hat{\Phi}) \forall \hat{\Phi} \in \mathbb{W}, \\ \alpha(f, \hat{f})_{\mathbb{H}_f} + \langle J'_b(f, \Phi), \Phi_f \rangle + \langle J'_f(f, \Phi), \hat{f} \rangle = 0 \forall \hat{f} \in \mathbb{H}_f. \end{cases} \tag{3.5.6} \]

The element \( \Phi_f \) can be eliminated from (3.5.6) by introducing the adjoint problem: find \( Q := (q, \sigma)^T \in \mathbb{W} \) s.t.
\[ \mathcal{L}^*(Q, \hat{W}) := \mathcal{L}(\hat{W}, Q) = \langle J'_b(f, \Phi), \hat{W} \rangle \forall \hat{W} \in \mathbb{W}. \tag{3.5.7} \]

Since \( \Phi_f \in \mathbb{W} \) we can choose \( \hat{W} = \Phi_f \) in (3.5.7), yielding
\[ \langle J'_b(f, \Phi), \Phi_f \rangle = \mathcal{L}(\Phi, \hat{\Phi}) = B(f, Q) \tag{3.5.8} \]
and the system of variational equations (3.5.6) reads now as follows:
\[ \begin{cases} \mathcal{L}(\Phi, \hat{\Phi}) = B(f, \hat{\Phi}) \forall \hat{\Phi} \in \mathbb{W}, \\ \mathcal{L}^*(Q, \hat{W}) = \langle J'_b(f, \Phi), \hat{W} \rangle \forall \hat{W} \in \mathbb{W}, \\ \alpha(f, \hat{f})_{\mathbb{H}_f} + B(f, Q) + \langle J'_f(f, \Phi), \hat{f} \rangle = 0 \forall \hat{f} \in \mathbb{H}_f. \end{cases} \tag{3.5.9} \]

The first equation is the state equation, the second is the adjoint one and the third equation is the optimality condition. Let us define the following operators (see Lions and Magenes [84], Lions [82] and Agoshkov [6]):
\[ L : \mathbb{W} \to \mathbb{W}^*, \ (L_a \Phi, \hat{\Phi})_{\mathbb{H}_0} := \mathcal{L}(\Phi, \hat{\Phi}), \forall \Phi, \hat{\Phi} \in \mathbb{W}, \]
\[ L^* : \mathbb{W} \to \mathbb{W}^*, \ (\hat{W}, L^* Q)_{\mathbb{H}_0} = (L \hat{W}, Q)_{\mathbb{H}_0}, \forall \hat{W}, Q \in \mathbb{W}, \]
\[ B : \mathbb{H}_f \to \mathbb{W}^*, \ (B f, \Phi)_{\mathbb{H}_0} = B(f, \Phi) \forall \Phi. \]
\[ \Lambda_w : \mathbb{W}^* \to \mathbb{W}^*, \ (\Lambda_w J_b(f, \Phi), \hat{W})_{\mathbb{H}_0} := \langle J'_b(f, \Phi), \hat{W} \rangle, \]
\[ \Lambda_f : \mathbb{H}_f^* \to \mathbb{H}^*_f, \ (\Lambda_f J_f(f, \Phi), \hat{f})_{L^2(x_1, x_2)} := \langle J'_f(f, \Phi), \hat{f} \rangle. \]

Now the system (3.5.9) can be written in operator form as follows:
\[ \begin{cases} L_a \Phi = B f \ (\text{in } \mathbb{W}^*), \\ L^* Q = \Lambda_w J_b(f, \Phi) \ (\text{in } \mathbb{W}^*), \\ \alpha \Lambda_a f + B^* Q + \Lambda_f J_f(f, \Phi) = 0 \ (\text{in } (\mathbb{H}_f)^*), \end{cases} \tag{3.5.10} \]
where $\Lambda_c$ is the extension to $\mathbb{H}_f$ of the following operator $\Lambda_{c,0}$ (defined on $H^2 \cap \mathbb{H}_f$):

$$\Lambda_{c,0} f := -f_{xx} + f,$$

**Remark 3.5.1** The system (3.5.10) with a cost functional $J(\Phi) = \| C \Phi - \Psi \|_{H_{obs}}^2$, where $C : \mathbb{W} \rightarrow \mathbb{H}_{obs}$ is a given operator and $\Psi \in \mathbb{H}_{obs}$ a given observation function, is analyzed in Agoshkov [6]. In this case $J' = 0$ and $\Lambda_w J'_w (f, \Phi) = C^*(C \Phi - \Psi)$ and we find a “classical” optimal control problem in the sense of J.L. Lions’ formulation.

### 3.6 Existence and uniqueness results

We analyze the particular cases where the cost functional $J$ is chosen as outlined by Example 3.4.1 to obtain some uniqueness and existence results. See also Isakov [61].

Let $J$ be the functional $J_2$ of Example 3.4.1. Then

$$J(f, \Phi) = J(f, u, p) = \frac{\gamma_1}{2} \int_{\Omega_0} m_{wd} |\nabla \times u + m_1 R_f f - R_{obs,1}|^2 d\Omega + \frac{\gamma_2}{2} \int_{\Gamma_{out}} \left( |p - p_{out}|^2 + |u - u_{out}|^2 \right) d\Gamma$$

(3.6.1)

To study the problem in this case we assume that $\Omega_{wd} = \Omega_0$ and we define:

$$X := \{ u : u \in (H^2(\Omega))^2, u = 0 \text{ on } \Gamma_{in} \cup \Gamma_{w_1} \cup \Gamma_{w_2} \},$$

$$\mathbb{H}^p := H^1(\Omega_0), \quad \mathbb{H}_f := H^2(x_1, x_2) \cap H^1_0(x_1, x_2).$$

Here we consider $H^2$ as velocity space in order to be allowed to use the uniqueness continuation theorem (see Weck [162], Fabre and Lebeau [35]). The derivatives $J'_w (f, \Phi)$ and $J'_f (f, \Phi)$ become

$$\langle J'_w (f, \Phi), \dot{\Phi} \rangle = \gamma_1 \int_{\Omega_0} m_{wd} (\nabla \times u + m_1 R_f f - R_{obs,1}) \cdot (\nabla \times v) d\Omega +$$

$$+ \gamma_2 \int_{\Gamma_{out}} (p - p_{out}) q d\Gamma + \gamma_2 \int_{\Gamma_{out}} (u - u_{out}) \cdot v d\Gamma,$$

$$\langle J'_f (f, \Phi), \dot{f} \rangle = \gamma_1 \int_{\Omega_0} m_{wd} (\nabla \times u + m_1 R_f f - R_{obs,1}) R_f \dot{f} d\Omega,$$

$$\forall \dot{\Phi} = (v, q) \text{ and } \forall \dot{f}.$$
where
\[ F(f, v) := b_f(f, p_0, v) + G_1(f, v) - a_f(f, u_0, v), \]
and for every \( f, u_f = u_f(f) \), \( p_f = p_f(f) \) denote the solution of the system given by the first and second equations in (3.6.2) corresponding to a right end side \( f = \hat{f} \). The system (3.5.9) is: find \( u_f \in X, p_f \in \mathbb{H}^p \)

\[
\begin{align*}
    a_0(u_f, v) &= b_0(p_f, v) + F(f, v) \forall v \in X, \\
    b_0(q, u_f) + b_f(f; q, u_0) &= 0 \forall q \in \mathbb{H}^p(\Omega), \\
    a_0(\dot{q}, q) &= -b_0(\dot{\sigma}, \dot{q}) + \gamma_1 \int_{\Omega_0} m_{wd}(\nabla \times u_f + m_1 R_f f - R_{obs, 1}) \cdot (\nabla \times \dot{q}) d\Omega + \\
    &+ \gamma_2 \int_{\Gamma_{out}} (u_f - u_{out}) \cdot \dot{q} d\Gamma \forall \dot{q} \in X, \\
    -b_0(\dot{\sigma}, q) &= \gamma_2 \int_{\Gamma_{out}} (p_f - p_{out}) \dot{\sigma} d\Gamma \forall \dot{\sigma} \in \mathbb{H}^p, \\
    a(f, \dot{f})_{\Omega_f} + F(\dot{f}, q) - b_f(f; \sigma, u_0) + \\
    &+ \gamma_1 \int_{\Omega_0} m_{wd}(\nabla \times u_f + m_1 R_f f - R_{obs, 1}) m_1 R_f \dot{f} d\Omega = 0 \forall \dot{f} \in \mathbb{H}_f. 
\end{align*}
\] (3.6.3)

In the sequel we assume that the generalized Stokes problem (3.3.7) has a unique solution for any given \( u_0, p_0 \) (the solution in the unperturbed domain \( \Omega_0 \)) and for each \( f \in \mathbb{H}_f \). (See Galdi [40], Fabre and Lebeau [36], Girault and Raviart [43]).

Consider now the problem (3.6.3) for \( \alpha > 0 \).

**Proposition 3.6.1** For any \( \alpha > 0 \) problem (3.6.3) has a unique solution for any given \( R_{obs, 1} \).

**Proof.** Following Agoshkov [6], we formally invert \( L \) and \( L^* \) in the first and second equations of (3.5.10) then we substitute \( \Phi, \tilde{Q} \) into the third equation and we obtain the following weak problem: for \( f \in \mathbb{H}_f \):

\[
\alpha(f, \tilde{f})_{\mathbb{H}_f} + (A_f, \tilde{A}_f)_{L^2(x_1, x_2)} = (G, \tilde{A}_f)_{L^2(x_1, x_2)} \forall \tilde{f} \in \mathbb{H}_f, \tag{3.6.4}
\]

where \( A \) is a linear operator, which depends on previous operators from variational equations, while \( G \) will depend on the data. More precisely from (3.6.2) we obtain:

\[
(f, \tilde{f})_{\mathbb{H}_f} = (A_f f, \tilde{f})_{L^2(x_1, x_2)}, \\
(A_f, \tilde{A}_f)_{L^2(x_1, x_2)} = \gamma_1 \int_{\Omega} m_{wd}(\nabla \times u + m_1 R_f f) \cdot (\nabla \times u_f + m_1 R_f \tilde{f}) d\Omega + \\
+ \gamma_2 \int_{\Gamma_{out}} (pp_f + u \cdot u_f) d\Gamma, \\
(G, \tilde{A}_f)_{L^2(x_1, x_2)} = \gamma_1 \int_{\Omega} m_{wd} R_{obs, 1} \cdot (\nabla \times u_f + m_1 R_f \tilde{f}) d\Omega + \gamma_2 \int_{\Gamma_{out}} (p_{out} p_f + u_{out} \cdot u_f) d\Gamma,
\]

where \( \Phi = (u, p) = L^{-1} B f, \Phi_{\tilde{f}} = (u_{\tilde{f}}, p_{\tilde{f}}) = L^{-1} B \tilde{f}, \forall \tilde{f} \in \mathbb{H}_f. \)

We see that problem (3.6.4) has a unique solution which satisfies \( \|f\|_{\mathbb{H}_f} \leq \|G\|^2/(2\alpha) < \infty. \)

Correspondingly, we can construct \( u, p, q, \sigma \), which together with \( f \) provides the unique
Consider now the problem (3.6.3) with $\alpha = 0$.

**Proposition 3.6.2** Assume that: i) The solution of the generalized Stokes problem satisfies 
\[ \left( \frac{\partial v_0}{\partial y} \right)^2 + \left( \frac{\partial u_0}{\partial y} \right)^2 > 0 \] at $y = 0$, $x \in (x_1, x_2)$ ii) problem (3.6.3) has a solution. Then the solution of (3.6.3) is unique in the class $(H^2(\Omega))^2 \times H^1(\Omega) \times W^{1,\infty}(x_1, x_2)$. 

**Proof.** Let $(u_1, \ldots, f_1)$ and $(u_2, \ldots, f_2)$ be two solutions of (3.6.3). Then for $u = u_1 - u_2, \ldots, f = f_1 - f_2$ from (3.6.2) we obtain:

\[
\begin{aligned}
a_0(u, v) &= b_0(p, v) + F(f, v) \forall v \in X, \\
b_0(q, u) + b^f(f; q, u_0) &= 0 \forall q \in H^p(\Omega), \\
\nabla \times u + m_1R_ff &= 0 \text{ in } \Omega, \\
p &= 0, \quad u = 0 \text{ on } \Gamma_{out}.
\end{aligned}
\] (3.6.5)

Consider the second and third equation from (3.6.5) in $\Omega_{2,0}$

\[ \nabla \cdot u = 0, \quad \nabla \times u = 0 \text{ in } \Omega_{2,0}. \]

Then $\Delta u = 0$ in $\Omega_{2,0}$. Considering $v$ with $\text{supp}(v) \subseteq \Omega_{2,0}$ from the first equation of (3.6.5) we find $\nabla p = 0$, then $p = \text{const}$ in $\Omega_{2,0}$ and $-p\mathbf{n} + \nu\frac{\partial u}{\partial n} = 0$ on $\Gamma_{out}$. Since $p = 0$ on $\Gamma_{out}$ then $p = 0$ in $\Omega_{2,0}$ and $\nu\frac{\partial u}{\partial n} = 0$ on $\Gamma_{out}$ too. Consequently, $u$ satisfies:

\[ \Delta u = 0 \text{ in } \Omega_{2,0}, \quad u = \nu\frac{\partial u}{\partial n} = 0 \text{ on } \Gamma_{out}. \]

This problem has only the trivial solution $u = 0$ in $\Omega_{2,0}$. Since $u \in (H^2(\Omega))^2$ then

\[ u = \frac{\partial u}{\partial n} = 0 \text{ on } \Gamma_0 := \{(x, y) : y = 0, x_1 < x < x_2\}. \]

Consider now the second and third equations from (3.6.5) in $\Omega_{1,0}$:

\[
\begin{aligned}
\nabla \cdot u - \left[ yf(x, 0, f_0) \frac{\partial u_0}{\partial y} + f \frac{\partial u}{\partial y} \right] &= 0 \text{ in } \Omega_{1,0}, \\
\nabla \times u - \left[ yf(x, 0, f_0) \frac{\partial v_0}{\partial y} - f \frac{\partial u}{\partial y} \right] &= 0 \text{ in } \Omega_{1,0}.
\end{aligned}
\] (3.6.6)

On $\Gamma_0$ we have:

\[ \nabla \cdot u - \frac{f}{f_0} \frac{\partial v_0}{\partial y} = 0, \quad \nabla \times u + \frac{f}{f_0} \frac{\partial u_0}{\partial y} = 0, \]

\[ |f(x)| = f_0 \left[ \left( \frac{\nabla \cdot u}{f_0} \right)^2 + \left( \frac{\nabla \times u}{f_0} \right)^2 \right]^{1/2} \text{ on } \Gamma_0, \]
chapter 3. shape design using small perturbation

(The dependence of the right end side on \(x\) and \(y\) is understood). Since \(u = \frac{\partial u}{\partial n} = \frac{\partial u}{\partial y} = 0\) on \(\Gamma_0\), then

\[
\nabla \cdot u \mid_{y=0} = \frac{\partial u}{\partial x} \mid_{y=0} = 0, \quad \nabla \times u \mid_{y=0} = \frac{\partial v}{\partial y} \mid_{y=0} = 0, \quad x \in (x_1, x_2).
\]

i.e. \(f(x) = 0\). Therefore, \(u = 0, p = 0\) too.

Let us note once more, that if \(\gamma_2 > 0\) we account for the cost functional \(J_2\), then we overde-
terminate the initial problem with (3.4.4) for \(\alpha = 0\). Therefore in this case usually we have
uniqueness results, but not existence results. However in some physical problems the above
overdeterminations (and the presence of term \(\alpha \|f\|_W\)) have a physical sense, therefore in
these cases we can consider the optimal control problems like (3.4.5). This problem becomes
independent of the initial problem (where we have only \(J_1\)). Here, we have also existence re-
sults and can name these optimal control problems as the “optimal shape design problems”.
Nevertheless, it is interesting to investigate well posedness of the above variational problems
when \(\alpha = \gamma_2 = 0\).

3.7 Iterative processes

In this section we propose some iterative processes which are well suited for solving the vari-
ational equations obtained in the previous sections and we will use them for our numerical
tests.

Consider the problem (3.5.10); if for \(k = 0, 1, \ldots f^{(k)}\) is known, then \(f^{(k+1)}\) can be determi-
nated by solving the following equations ([6]):

\[
\begin{aligned}
L \Phi^{(k)} &= B f^{(k)}, \\
L^* Q^{(k)} &= \lambda_w J_f(f^{(k)}, \Phi^{(k)}), \\
\Lambda_c w^{(k)} &= B^* Q^{(k)} + J_f(f^{(k)}, \Phi^{(k)}), \\
f^{(k+1)} &= f^{(k)} - \tau_k (\alpha f^{(k)} + w^{(k)}),
\end{aligned}
\]

(3.7.1)

where \(\{\tau_k\}\) is a family of parameters whose determination follows from the theory of extremal
problems (see Vasiliev [155]), the general theory of iterative processes ( Marchuk [94], Quar-
teroni and Valli [127], Quarteroni, Sacco and Saleri [129]), and the ill-posed problems theory
( see Tikhonov and Arsenin [151], Vainikko and Veretennikov [153]). The step (3.7.1) would
read as follows for the variational form (3.5.9) of problem (3.5.10):

\[
\begin{aligned}
\mathcal{L}(\Phi^{(k)} + \hat{\Phi}, \hat{\Phi}) &= B(f^{(k)}), \quad \forall \hat{\Phi} \in \mathbb{W}, \\
\mathcal{L}(\hat{W}, Q^{(k)}) &= \langle J_f(f^{(k)}, \Phi^{(k)}), \hat{W} \rangle, \quad \forall \hat{W} \in \mathbb{W}, \\
\langle w^{(k)} , f \rangle_{\mathbb{H}_f} &= B(\hat{f}, Q^{(k)}) + \langle J^*_f(f^{(k)}, \Phi^{(k)}), \hat{f} \rangle, \quad \forall \hat{f} \in \mathbb{H}_f, \\
f^{(k+1)} &= f^{(k)} - \tau_k (\alpha f^{(k)} + w^{(k)}). 
\end{aligned}
\]

(3.7.2)

Consider now problem (3.6.2) (with \(\Omega_{wd} \subseteq \Omega\)). The iterative process (3.7.2) for this problem
reads as follows:

\[
\begin{aligned}
a_0(u^{(k)}, v) &= b_0(p^{(k)}, v) + F(f^{(k)}, v) \quad \forall v \in X, \\
b_0(q, u^{(k)}) + b_f(f^{(k)}; q, u_0) &= 0 \quad \forall q \in \mathbb{H}^p(\Omega), \\
a_0(\dot{q}, q^{(k)}) &= -b_0(\sigma^{(k)}, \dot{q}) + \gamma_1 \int_{\Omega_0} m_{wd}(\nabla \times u^{(k)}) + m_1 R_f f^{(k)} - R_{obs,1}), \\
(\nabla \times \dot{q})d\Omega + \gamma_2 \int_{\Gamma_{out}} (u^{(k)} - u_{out}) \cdot \dot{q}d\Gamma &= 0 \quad \forall \dot{q} \in X, \\
-b_0(\sigma, q^{(k)}) &= \gamma_2 \int_{\Gamma_{out}} (p^{(k)} - p_{out})\sigma d\Gamma \quad \forall \sigma \in \mathbb{H}^p, \\
(w^{(k)}, \hat{f})_{\mathbb{H}_f} &= F(f, q^{(k)}) - b_f(f; \sigma^{(k)}, u_0) + \\
+ \gamma_1 \int_{\Omega_0} m_{wd}(\nabla \times u^{(k)}) + m_1 R_f f^{(k)} - R_{obs,1}) m_1 R_f \hat{f}d\Omega \quad \forall \hat{f} \in \mathbb{H}_f, \\
f^{(k+1)} &= f^{(k)} - \tau_k (\alpha f^{(k)} + w^{(k)}), \quad k = 0, 1, \ldots
\end{aligned}
\]  

Consider now the \textit{finite dimensional case} in which the functions \(f, \{f^{(k)}\}, \hat{f}\) all are sought for in a finite-dimensional subspace \(\mathbb{H}_{f,N} \subset \mathbb{H}_f\) of dimension \(N < \infty\), whose basis \(\varphi_i \in \mathbb{W}^{1,\infty}(x_1, x_2), i = 1, 2, \ldots, N\). Then the following theorem holds.

**Theorem 3.7.1**. Assume that \(\Omega_{ad} = \Omega, \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 > 0\) at \(y = 0, x \in (x_1, x_2)\). Then:

1. The problem (3.6.2) is correctly solvable (solution exists and is unique) for \(\alpha \geq 0\) and all \(N < \infty\);

2. The iterative process (3.7.3) is convergent for any \(\alpha > 0, N < \infty\) and provided the parameters \(\tau_k > 0, k = 0, 1, 2, \ldots\) are small enough;

3. If \(\alpha\) is sufficiently small while \(k\) is sufficiently large, then \(\{u^{(k)}, p^{(k)}, f^{(k)}\}\) can be taken as an approximate solution of problem (3.6.2).

**Proof**:

1. The existence of the solution for \(\alpha > 0\) has been proved early. Let us consider the case \(\alpha = 0\). Since \(f = \sum_{i=1}^{N} a_i \varphi_i \in \mathbb{H}_{f,N}\) then in the form (3.6.4) with \(\alpha = 0\) we conclude that this equation is correctly solvable (because the problem (3.6.2) can have only unique solution in \(X \times \mathbb{H}^p \times \mathbb{H}_f\), see Proposition 3.6.2). We assume the generalized Stokes problem to be correctly solvable for any given \(f \in \mathbb{H}_f\). Hence the problem (3.6.2) is correctly solvable too.

2. If \(\alpha > 0\) then the bilinear form on the left hand side of (3.6.4) is coercive and continuous with respect to the norm \(\|f\|_{A,\alpha} = \sqrt{\alpha \|f\|_{\mathbb{H}_f}^2 + \|Af\|_{L^2(x_1, x_2)}^2}\). Then according to the general theory of iterative algorithm the process given by

\[
(f^{(k+1)}, \hat{f})_{\mathbb{H}_f} = (f^{(k)}, \hat{f})_{\mathbb{H}_f} - \tau (\alpha (f^{(k)}; \hat{f})_{\mathbb{H}_f} + (Af^{(k)}, A\hat{f})_{L^2(x_1, x_2)}) - \\
-(G, A\hat{f})_{L^2(x_1, x_2)}, \quad k = 0, 1, \ldots
\]

is convergent for small \(\tau > 0\). Hence the process (3.7.3) is convergent also and

\[
\|u^{(k)} - u\|_X + \|p^{(k)} - p\|_{\mathbb{H}^p} + \|f - f^{(k)}\|_{\mathbb{H}_f} \to 0, \quad k \to \infty.
\]

(3.7.4)
If $\Lambda^{-1} A^* A \in [C_1, C_2], C_1, C_2 = \text{const}$, and $\tau_h = 2/(2\alpha + C_1 + C_2)$ then (3.7.4) becomes (see Agoshkov [6]):

$$\|u^{(k)} - u\|_{\mathcal{X}} + \|p^{(k)} - p\|_{\mathcal{H}^p} + \|f - f^{(k)}\|_{\mathcal{H}^E} \leq C \left( \frac{C_2 - C_1}{2\alpha + C_1 + C_2} \right)^k \rightarrow 0, k \rightarrow \infty. \quad (3.7.5)$$

3. Let $u_0, p_0, f_0$ be a solution of (3.6.2) when $\alpha = 0$. According to the theory of ill-posed problems (Tikhonov and Arsenin [151], Vainikko and Veretennikov [153]) we have: $\|f_0 - f_0\|_{\mathcal{H}^p} \rightarrow 0$ as $\alpha \rightarrow 0^+$, where $(f_0, u_0, p_0)$ is the solution of (3.6.2) for $\alpha > 0$. Hence

$$\|u_0 - u_{\alpha}\|_{\mathcal{X}} + \|p_0 - p_{\alpha}\|_{\mathcal{H}^p} \rightarrow 0, \text{ as } \alpha \rightarrow 0^+. \quad (3.7.6)$$

Then owing to (3.7.4) we conclude that the statements of the theorem holds true also.

Then we have

$$\|u^{(k)} - u\|_{\mathcal{X}} + \|p^{(k)} - p\|_{\mathcal{H}^p} \rightarrow 0, \text{ as } \alpha \rightarrow 0^+. \quad (3.7.7)$$

The simple schemes in Figure 3.7 can be considered as examples of the above problems when $f \in \mathbb{H}_{f,N}$ for small $N$ (the dimension of $\mathbb{H}_{f,N}$).

Figure 3.7: Domain $\Omega$ with $N$ shape functions: (a) $N = 1$, $f = \beta_1 + a \varphi_0(x), \varphi_0 = x(x_2 - x)$; (b) $N = 3$, $f = \beta_1 + \sum_{i=1}^{3} a_i \phi_i$.

### 3.8 Test problem and numerical results

We consider some test problems on simplified arterial bypass configurations. Numerical simulations have been carried out using also in this case the *Bamg* –Bi-dimensional Anisotropic Mesh Generator– and the finite element library *FreeFem*, used to develop algorithms based on control theory and adjoint formulation for generalized Stokes problem. For application of finite element method to incompressible flow see also Gresho and Sani [46]. In this section we present numerical results using as cost functional the $L^2$ norm of the vorticity in the down-field zone of the new incoming branch of the bypass.

Wall curvature was considered only in the zone of the incoming branch of the bypass ($-2 \leq x \leq 0$) where we set $f_0 = 1 - \sin(\pi x)$; in other parts we used piecewise constant function. The graft angle of the bypass incoming branch (which influences vorticity) is equal to zero (between the artery and the new incoming branch there isn’t a relative angle). This aspect will be studied better in the next chapters dealing with a complete bypass configurations and
The original configuration is very similar to the one studied by Lee, Loth and Fisher (see Lee et al. [77] and Figure 3.8). Velocity values $u_m$ at the inflow are chosen in such a way that the Reynolds number $Re = \frac{\bar{u}D}{\nu}$ has order $10^3$. We recall that blood kinematic viscosity $\nu = \frac{\mu}{\rho}$ is equal to $4 \times 10^{-6} m^2 s^{-1}$, blood density $\rho = 1 g cm^{-3}$ and dynamic viscosity $\mu = 4 \times 10^{-2} g cm^{-1}s^{-1}$; $\bar{u}$ is a mean inflow velocity (absolute value) related with $u_m$, while $D$ is the arterial diameter (3.5 mm) [126].

Figures 3.9-3.11 provide a preliminary account of numerical results and show how the shape of the bypass using generalized steady Stokes equations in an optimal control problem is smoothed out at the corner, which represents a singularity. Figure 3.9 refers to the original configuration; whereas Figure 3.10 to the configuration obtained after 25 iterations of the optimization algorithm (the vorticity has been reduced by about the 30%). Figure 3.11 shows the most sensible zone of the bypass with respect to the cost functional to be minimized. In the picture, first order shape variation (correction) $f_1y/\beta$ has been plotted in the reference domain. This quantity is related with the adjoint solution, representing the shape sensitivity with respect to the gradient of the cost functional: the “hot” zone is the one at the corner.

Figure 3.9: Idealized 2-D bypass configuration before optimal shape design process: isovelocity $[ms^{-1}, 10^{-2}]$. 

Figure 3.8: Example of configuration proposed and studied by Lee, Loth and Fisher [77].
Figure 3.10: Bypass configuration at the end of shape optimization using first corrections: iso-velocity [$m s^{-1} \cdot 10^{-2}$].

Figure 3.11: First correction $\frac{\partial y}{\partial \sigma}$, quantity which is related with the adjoint solution $(\mathbf{q}, \sigma)$ and represents the shape variation in bypass configuration (reported in the reference domain) [$m \cdot 10^{-2}$].

The results that are obtained by applying shape optimization to first order problem provide a regularization of the shape between the incoming branch of the bypass and the host artery. These results provide a validation of our methodological approach: the shape found is the one presented as “optimized” (minimizing wall shear stress) by Lei, Archie et Kleinstreuer in [78] as an alternative one with respect to the Taylor patch, reducing wall shear stress. We recall that the vorticity cost functional is related with the viscous dissipation dealing with incompressible flows and that the distributed vorticity is related with the velocity field too in the boundary of the region we considered. See Section 2.7.
3.9 A comparison between the two methods

At the end of this first investigation stage, preliminary reported in [4] for the small perturbation theory, we present below some numerical results obtained by applying both the optimal control by local boundary variation (Figure 3.12) and the small perturbations techniques (Figure 3.13), starting from the same configuration. Results by local boundary variations give a shape more similar to Taylor patch (see for a comparison [78]). We have to consider that in this case we are not dealing with a first order perturbation problem. The idea we propose is to iterate our algorithm and restart the optimization process with a new initial shape \( \tilde{f}_0 \) which is the optimized one \( \tilde{f}_0 = f_0 + \varepsilon f_1 \) from the previous iterative process. Figure 3.14 shows the shape sensitivity with respect with the cost functional for the problem studied by local boundary variations, while Figure 3.15 shows the reduction of the cost functional (3.4.11) during the optimization processes and a comparison between the two methods. The vorticity reduction during the linearized shape design process (by small perturbation) is shown in Figure 3.15 for the first iterative process. It is possible to iterate the process to get the same configuration provided by local boundary variations approach (Figure 3.12) and a more consistent vorticity reduction as in the local boundary variations case.

The method proposed in this chapter has allowed us (i) to study a shape optimization problem as a control problem on a set of coefficients representing shape functions, (ii) to set the problem in the theoretical framework of a generalized Stokes problem and (iii) to avoid all the topological aspects related with shape optimization with nodal variations in the computational mesh.

Figure 3.12: Bypass configuration (velocity \([m \cdot s^{-1} \cdot 10^{-2}]\), same colorbar of Fig. 3.10) near the incoming branch after shape optimization by boundary variations (30 iterations, 40% vorticity reduction).
CHAPTER 3. SHAPE DESIGN USING SMALL PERTURBATION

Figure 3.13: Bypass configuration (velocity $[m s^{-1} \cdot 10^{-2}]$, same colorbar of Fig. 3.10, 25 iterations and 30% vorticity reduction) after design by small perturbations.

Figure 3.14: Bypass configuration and its shape sensitivity in reference domain with respect to the cost functional (quantity related with adjoint solution) to underline the most sensible zone related with observation $[m \cdot 10^{-2}]$.

3.10 Further developments

In this chapter we have focused on the problem of determining the first corrections for the shape design of simplified two-dimensional bypass configurations. We recall here that we are in the perspective of using low order models for optimal control and shape optimization algorithms, but this methodology may be applied to fully unsteady incompressible Stokes (as we will do in the next chapter) and Navier-Stokes equations and consider the coupled fluid-structure problem. The setting of the problem in a three-dimensional geometry will provide more realistic design indications concerning surgical prosthesis realizations. The development of tools for geometry reconstruction from medical data (medical imaging and other non-invasive means) and their integration with numerical simulation could provide
improvements in disease diagnosis procedures and more realistic models. Also the integration of numerical models with experimental ones could allow to get a more complete study framework, see for example Bertolotti et al. [18].

An important application of the numerical method developed in this chapter is the possibility to realize the iterative process for solving initial non-linear problems. For that it is sufficient to consider $f = f_0 + \varepsilon f_1$, where $f_0$ is the initial configuration and $f_1$ the computed first correction, as the new $f_0$, then to calculate a new first correction and so on.

A further useful development may be devoted to build domain decomposition methods (see Quarteroni and Valli [128]) based on optimal control approaches. What we are going to develop in the next chapters of this thesis, after a generalization of the approach we have proposed here, are efficient schemes for reduced-basis methodology approximations (see Patera, Maday, Prud’homme, Rovas, Veroy and Turinici in [118] and/or [119]) to study complete parametrized bypass configuration. This methodology will be more efficient for the use in a repetitive design environment as optimal shape design methodology requires.

Figure 3.15: Distributed vorticity $[m^2 s^{-1}]$ reduction achieved during the two optimization processes: shape design by local boundary variations and small perturbations.
Chapter 4

Shape Design by Optimal Control with Unsteady Flows and Perturbation Method

In this chapter we present an approach for the study of the Aorto-Coronaric bypass anastomoses using unsteady Stokes equations. Our aim is still to develop optimization methods based on simplified fluid models. This approach can be seen as a methodological generalization of the results of the previous chapter; as a matter of fact, the shape optimization process takes into account unsteady flow phenomena over a period $T$ to build an optimized shape which is not time dependent. The small perturbation method, coupled with optimal control, is extended to unsteady flows, moreover existence and uniqueness results are generalized to the unsteady case.

4.1 Introduction

In this chapter we still apply optimal control for the shape optimization of Aorto-Coronaric bypass anastomoses. We consider unsteady (rather than steady) Stokes equations with the aim of providing a better understanding of the blood flow dynamics. We analyze the “first correction” method which is derived by applying a perturbation method to the initial unsteady problem in a space-time domain $\Omega \times (0,T)$ with $\Omega \subset \mathbb{R}^2$. The boundary $\partial \Omega$ of $\Omega$ is parametrized by a suitable function $f$. In this chapter we extend the approach and the results already introduced in the previous chapter (and preliminary reported in [4]) to the case where the non-stationary Stokes equations are used. An outline of this chapter is as follows. In Section 4.2 we formulate our problem, in Section 4.3 we deal with the problem of perturbed functions in the generalized unsteady Stokes equations framework. In Section 4.4 we introduce the shape optimization problem and its equations in the optimal control framework dealing with unsteady cost functionals (Section 4.5). In Section 4.6 uniqueness and existence results are extended, then in Section 4.7 an iterative optimization algorithm is proposed. Section 4.8 deals with a test problem and numerical results, finally some conclusions are drawn in
Section 4.9.

4.2 The non-stationary problem

Let $\Omega$ be a bounded domain of $\mathbb{R}^2$ with boundary $\Gamma$, $\mathbf{x} = (x, y)^T$ is a point of $\Omega$, $t \in [0, T]$, $T < \infty$, is the time variable. As before we consider the idealized, two-dimensional bypass bridge configuration of Figure 4.1 and the domain on Figure 4.2, where the dotted line represents the geometry of the complete anastomosis; $\Gamma_{w_2}$ is the section of the original artery, $\Gamma_{in}$ is the new anastomosis inflow after bypass surgery, $\Gamma_{out}$ is the anastomosis outflow.

![Diagram](image)

Figure 4.1: Idealized, 2-D bypass bridge configuration (left) and detail of the sensible part for the optimization process (right). The dotted curve represents the portion of the boundary that is subjected to change.

![Diagram](image)

Figure 4.2: Main notations: $\Omega = \Omega_1 \cup \Omega_2$, $\Gamma_w = \Gamma_{w_1} \cup \Gamma_{w_2} \cup \Gamma_{w_3}$, $\Gamma_0 = \partial \Omega_1 \cap \partial \Omega_2$

We consider the following boundary-value problem for the unsteady Stokes equations, used
to model blood flow at low Reynolds number: find $u, p$ s.t.

$$
\begin{align*}
\frac{\partial u}{\partial t} - \nu \Delta u + \nabla p &= f \text{ in } \Omega \times (0, T), \\
\nabla \cdot u &= 0 \text{ in } \Omega \times (0, T), \\
\mathbf{u} &= \mathbf{u}_{in} \text{ on } \Gamma_{in}, \quad \mathbf{u} = \mathbf{0} \text{ on } \Gamma_{w1} \cup \Gamma_{w2} \quad \forall t \in (0, T), \\
-p\mathbf{n} + \nu \frac{\partial u}{\partial n} &= \mathbf{g}_{out} \text{ on } \Gamma_{out} \cup \Gamma_{w2} \quad \forall t \in (0, T), \\
\mathbf{u} &= \mathbf{u}^* \text{ at } t = 0 \text{ in } \Omega,
\end{align*}
$$

where $\mathbf{u}_t := \frac{\partial \mathbf{u}}{\partial t}$, $\mathbf{u}^*$ is a given vector function such that $\nabla \cdot \mathbf{u}^* = 0$ in $\Omega$, $\mathbf{n} = (n_1, n_2)^T$ is the outward unit normal vector on $\Gamma$, $f = f(x, y, t)$, $\mathbf{u}_{in} = \mathbf{u}_{in}(x, y, t)$, $\mathbf{g}_{out} = \mathbf{g}_{out}(x, y, t)$ are given vector functions, $\nu = const > 0$ and $\mathbf{u}_f = \{\mathbf{u}_{in} \text{ on } \Gamma_{in}; \mathbf{0} \text{ on } \Gamma_{w1} \cup \Gamma_{w2}\}$. As seen previously, the subset $\Gamma_{c, \varepsilon}$ of $\Gamma_{w1}$ is parametrized by a function $f(x, \varepsilon)$ of $x \in [x_1, x_2]$ and $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$, $\varepsilon_0 = const$ is a small parameter. More precisely we assume that $f(x, \varepsilon)$ is time independent and can be developed as follows:

$$
f(x, \varepsilon) = f_0(x) + \varepsilon f_1(x) + \varepsilon^2 f_2(x) + \ldots,
$$

where $f_k \in \mathbb{W}^{1,\infty}(x_1, x_2)$, for $k = 0$, and $f_k \in \mathbb{W}^{1,\infty}_0(x_1, x_2)$, for $k \geq 1$, so that $f_k(x_1) = f_k(x_2) = 0, k \geq 1$. Here the function $f_0(x) > 0$ describes the original subset $\Gamma_{c,0}$ of the boundary of the “unperturbed domain”, $\Gamma_{w0} \equiv \partial \Omega_0$ (see Figure 4.3), while $f_k(x), k \geq 1$, could be unknown when considering a control problem (see Section 4.4).
We introduce the following real Hilbert spaces:

\[ X := \{ v : v \in (H^1(\Omega))^2, v = 0 \text{ on } \Gamma_{in} \cup \Gamma_{w_1} \cup \Gamma_{w_3} \}, \]

\[ H := L^2(\Omega \times (0, T)) \equiv H^*, \]

\[ Y := L^2(0, T; \mathbb{R}^2), u(x, y, T) = 0 \} \]

The weak statement of Eq.(4.2.1) reads: find \( u \in L^2(0, T; (H^1(\Omega))^2) \), \( p \in H \) s.t.

\[
\begin{cases}
  a(u, v) - b(p, v) = G(v) \quad \forall v \in W, \\
  b(q, u) = 0 \quad \forall q \in H, \\
  u = u_f \text{ on } \Gamma_{in} \cup \Gamma_{w_1} \cup \Gamma_{w_3} \quad \forall \ t \in \ (0, T),
\end{cases}
\]

(4.2.3)

we recall that with \( v \) and \( q \) we indicate test functions and:

\[
a(u, v) = \int_0^T \int_{\Omega} \nu \nabla u \cdot \nabla v \, dt - \int_0^T \int_{\Omega} u \cdot \nu_v \, d\Omega dt,
\]

\[
b(p, v) = \int_0^T \int_{\Omega} p \nabla \cdot v \, d\Omega dt,
\]

\[
G(v) = \int_0^T \int_{\Omega} f \cdot v \, d\Omega dt + \int_0^T \int_{\Gamma_{out} \cap \Gamma_{w_2}} g_{out} \cdot v \, d\Gamma dt + \int_{\Omega} u^* \cdot v(x, y, 0) \, d\Omega.
\]

The forms \( a(., .) \), \( b(., .) \) and \( G(., .) \) depend on the parametrization \( f \) of \( \Gamma_{c, \varepsilon} \), however this dependence will be understood for simplicity of notations.

4.3 The unsteady problem for the perturbed functions

Let us introduce the reference (simple-shaped) domains \( \bar{\Omega}_1 = \{ 0 < \bar{x} < A, 0 < \bar{y} < \beta_1 \equiv \beta \}, \)

\( \bar{\Omega}_2 = \{ 0 < \bar{x} < A, -\beta_2 < \bar{y} < 0 \}, \) and \( \bar{\Omega} = \bar{\Omega}_1 \cup \bar{\Omega}_2 \) (see Figure 4.4). Then we assume that \( f(x, \varepsilon) > 0 \) and consider the following variable transformation:

\[
T_f : \bar{\Omega}_1 \cup \bar{\Omega}_2 \rightarrow \bar{\Omega}, \ \bar{x} = T_f(x);
\]

\( T_f \) is the identity in \( \Omega_2 \), while \( T_f(x, y) = (x, \frac{\beta}{f(x, \varepsilon)} y) \) in \( \Omega_1 \). We set \( \bar{x} = (\bar{x}, \bar{y}) \) and define

\[
\tilde{u}(\bar{x}, t) := u \circ T_f^{-1}(\bar{x}, t) = u(\bar{x}, \bar{y} f(\bar{x}, \varepsilon)/\beta, t).
\]

where \( \tilde{v} = (\tilde{u}, \tilde{v}) \). Then,

\[
dx dy = \frac{f(\bar{x}, \varepsilon)}{\beta} d\bar{x} d\bar{y}
\]

and for every \( \tilde{\phi} : \bar{\Omega} \times (0, T) \rightarrow \mathbb{R} \) and \( \phi = \tilde{\phi} \circ T_f \) the following relations hold (with \( f_x := df/dx \)):

\[
\frac{\partial \phi}{\partial y}(\bar{x}, t) = \frac{\beta}{f(\bar{x}, \varepsilon)} \frac{\partial \phi(\bar{x}, t)}{\partial y}, \quad \frac{\partial \phi}{\partial x}(\bar{x}, t) = \frac{\partial \phi(\bar{x}, t)}{\partial \bar{x}} - \frac{y}{f(\bar{x}, \varepsilon)} \frac{f_x(\bar{x}, \varepsilon)}{\partial \bar{y}},
\]

(4.3.1)
Then in \( \tilde{\Omega} \) we have:
\[
\tilde{\mathcal{D}}(f) \tilde{u} = m_2 \tilde{\nabla} \cdot \tilde{u} + m_1 \tilde{\mathcal{D}}(f) \tilde{u}, \quad \tilde{\mathcal{R}}(f) \tilde{u} = m_2 \tilde{\nabla} \times \tilde{u} + m_1 \tilde{\mathcal{R}}(f) \tilde{u},
\]
where \( \tilde{\nabla} \phi := (\frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y}) \), while \( m_s \) is the characteristic function of \( \tilde{\Omega}_s \) \((s = 1, 2)\). To simplify the notations from now on we will set (unless otherwise specified):
\[
\mathbf{x} = \tilde{\mathbf{x}}, \quad \mathbf{u}(x,y,t) = \tilde{\mathbf{u}}(\tilde{x},\tilde{y},t), \quad u = \tilde{u}, \quad v = \tilde{v}, \ldots,
\]
\[
\mathcal{D} = \tilde{\mathcal{D}}, \quad \mathcal{R} = \tilde{\mathcal{R}}, \quad \Omega = \tilde{\Omega}, \quad \Gamma_{w_k} = \tilde{\Gamma}_{w_k}.
\]
Then problem (4.2.3) in the new reference frame \( \tilde{\Omega} \), now renamed \( \Omega \), reads as follows:
\[
\begin{cases}
    a(f; \mathbf{u}, \mathbf{v}) - b(f; p, \mathbf{v}) = G(f; \mathbf{v}) \quad \forall \mathbf{v} \in W, \\
    b(f; q, \mathbf{u}) = 0 \quad \forall q \in H, \\
    \mathbf{u} = u_f \text{ on } \Gamma_{in} \cup \Gamma_{w_1} \cup \Gamma_{w_3} \forall t \in (0, T).
\end{cases}
\]
(4.3.3)

We have emphasized the dependence of \( a(f; \cdot, \cdot), b(f; \cdot, \cdot), \) and \( G(f; \cdot) \) on \( f \). Precisely, upon writing \( \Omega_1 \) instead of \( \tilde{\Omega}_1 \) and \( \Omega_2 \) instead of \( \tilde{\Omega}_2 \) for simplicity of notation we have (unless otherwise specified, integration is carried out with respect to \( d\Omega dt \)):
\[
\begin{align*}
    a_1(f; \mathbf{u}, \mathbf{v}) &= a_1(f; \mathbf{u}, \mathbf{v}) + a_2(\mathbf{u}, \mathbf{v}), \\
    a_1(f; \mathbf{u}, \mathbf{v}) &= \int_0^T \int_{\Omega_1} \frac{f}{\beta} \left( \frac{\partial \mathbf{u}}{\partial x} \frac{\partial \mathbf{u}}{\partial y} + \frac{\partial \mathbf{v}}{\partial x} \frac{\partial \mathbf{v}}{\partial y} \right) - \int_0^T \int_{\Omega_1} \frac{f}{\beta} \mathbf{u} \cdot \mathbf{v}_t, \\
    a_2(\mathbf{u}, \mathbf{v}) &= \int_0^T \int_{\Omega_2} \frac{f}{\beta} \mathbf{u} \cdot \mathbf{v}_t, \\
    b_1(f; p, \mathbf{v}) &= b_1(f; p, \mathbf{v}) + b_2(p, \mathbf{v}), \\
    b_1(f; p, \mathbf{v}) &= \int_0^T \int_{\Omega_1} \frac{f}{\beta} p \mathcal{D}(f) \mathbf{v}, \quad b_2(p, \mathbf{v}) = \int_0^T \int_{\Omega_2} p \tilde{\nabla} \cdot \mathbf{v}, \\
    G_1(f; \mathbf{v}) &= G_1(f; \mathbf{v}) + G_2(\mathbf{v}), \\
    G_1(f; \mathbf{v}) &= \int_0^T \int_{\Omega_1} \frac{f}{\beta} \mathbf{f} \cdot \mathbf{v} + \int_0^T \int_{(\Gamma_{out} \cup \Gamma_{w_2}) \cap \partial \Omega_1} \mathbf{g}_{out} \cdot \mathbf{v} d\Gamma dt + \\
    &+ \int_{\Omega_1} \frac{f}{\beta} \mathbf{u}_s(x,y) \cdot \mathbf{v}(x,y,0) d\Omega, \\
    G_2(\mathbf{v}) &= \int_0^T \int_{\Omega_2} \mathbf{f} \cdot \mathbf{v} + \int_0^T \int_{(\Gamma_{out} \cup \Gamma_{w_2}) \cap \partial \Omega_2} \mathbf{g}_{out} \cdot \mathbf{v} d\Gamma dt + \\
    &+ \int_{\Omega_2} \mathbf{u}_s(x,y) \cdot \mathbf{v}(x,y,0) d\Omega.
\end{align*}
\]
Assume that problem (4.3.3) has a solution \( u, p \) that is infinitely differentiable with respect to \( \varepsilon \):

\[
\begin{align*}
\mathbf{u} &= \mathbf{u}_0 + \varepsilon \mathbf{u}_1 + \varepsilon^2 \mathbf{u}_2 + \ldots \\
p &= p_0 + \varepsilon p_1 + \varepsilon^2 p_2 + \ldots
\end{align*}
\] (4.3.4)

where \( p_k \in H, u_k \in Y, k \geq 1 \), the test functions \( \mathbf{v}, q \) appearing in (4.3.3) and in the sequel can be assumed as independent of \( \varepsilon \).

Using (4.2.2)-(4.3.4) and the small perturbation technique we can deduce the equations satisfied by \( u_k, p_k, k \geq 0 \). In particular, for \( k = 0 \), \( u_0 \) and \( p_0 \) satisfy

\[
\begin{align*}
a(f_0; \mathbf{u}_0, \mathbf{v}) - b(f_0; p_0, \mathbf{v}) &= G(f_0; \mathbf{v}) \quad \forall \mathbf{v} \in W, \\
b(f_0; q, \mathbf{u}_0) &= 0 \quad \forall q \in H, \\
\mathbf{u}_0 &= \mathbf{u}_f \text{ on } \Gamma_{in} \cup \Gamma_{w1} \cup \Gamma_{w3}, \forall t \in (0, T).
\end{align*}
\] (4.3.5)

Correspondingly we define:

\[
\mathcal{R}_{obs,0} := \mathcal{R}(f_0)u_0.
\] (4.3.6)

We introduce some further functional spaces \( \mathbb{H}^p \) and \( \mathbb{H}_f \) for \( p \) and \( \{f_k\} \), respectively, which satisfy:

\[
\mathbb{H}^p \subseteq H \subseteq \mathbb{H}^p, \quad \mathbb{H}^p \subseteq L^2(x_1, x_2) \subseteq \mathbb{H}^p,
\]

\[
W := W \times \mathbb{H}^p \subseteq \mathbb{H}_0 := L^2(\Omega \times (0, T))^2 \times L^2(\Omega \times (0, T)) \subseteq \mathbb{W}^*.
\]

Then we set:

\[
\mathbb{Y} := Y \times \mathbb{H}^p \subseteq \mathbb{H}_0 \subseteq \mathbb{Y}^*.
\]

Then for \( k = 1 \) the functions \( u_1, p_1 \), considered as the components of the vector-function \( \Phi_1 := (u_1, p_1) \in \mathbb{Y}, f_1 \in \mathbb{H}_f \), satisfy the equation:

\[
\mathcal{L}(\Phi_1, \dot{\Phi}) = B(f_1, \dot{\Phi}) \quad \forall \dot{\Phi} := (\mathbf{v}, q) \in W,
\] (4.3.7)

where

\[
\mathcal{L}(\Phi_1, \dot{\Phi}) := a_0(f_0; u_1, \mathbf{v}) - b_0(f_0; p_1, \mathbf{v}) + b_0(f_0; q, \mathbf{u}_1),
\]

\[
B(f_1, \dot{\Phi}) := b_f(f_1; p_0, \mathbf{v}) + G_1(f_1; \mathbf{v}) - a_f(f_1; \mathbf{u}_0, \mathbf{v}) - b_f(f_1; q, \mathbf{u}_0),
\]

\[
b_f(f_1; p_0, \mathbf{v}) := \frac{\partial}{\partial \varepsilon} b(f; p_0, \mathbf{v}) |_{\varepsilon=0} = \int_0^T \int_{\Omega_1} \frac{f_1}{\beta} p_0 \mathcal{D}(f_0) \mathbf{v} +
\]

\[
+ \int_0^T \int_{\Omega_1} \frac{f_0}{\beta} p_0 \mathcal{D}(f_1, \mathbf{v}),
\]

\[
\mathcal{D}_f(f_1, \mathbf{v}) := \frac{\partial}{\partial \varepsilon} \mathcal{D}(f) \mathbf{v} |_{\varepsilon=0} = - \left[ y \left( \frac{f_1}{f_0} \mathbf{f}_0 \right) \frac{\partial \mathbf{u}}{\partial y} + \frac{\beta f_1}{f_0^2} \frac{\partial \mathbf{v}}{\partial y} \right],
\]

\[
\mathcal{D}_f(f_1, \mathbf{u}_0) := \frac{\partial}{\partial \varepsilon} \mathcal{D}(f) \mathbf{u}_0 |_{\varepsilon=0} := \mathcal{D}_f f_1 \text{ in the sequel},
\]

\[
G_1(f_1; \mathbf{v}) := \frac{\partial}{\partial \varepsilon} G(f; \mathbf{v}) |_{\varepsilon=0} = \int_0^T \int_{\Omega_1} \frac{f_1}{\beta} \mathbf{f} \cdot \mathbf{v} + \int_{\Omega_1} \frac{f_1}{\beta} \mathbf{u}_0(x, y) \cdot \mathbf{v}(x, y, 0) d\Omega,
\]
§4.4. THE SHAPE OPTIMIZATION PROBLEM: UNSTEADY COST FUNCTIONALS

\[ a_f(f_1; u_0, v) := \frac{\partial}{\partial \epsilon} a(f; u_0, v)|_{\epsilon=0} = \]

\[ = \int_0^T \int_{\Omega_1} \frac{f_1 \nu}{\beta} \left( \frac{\partial u_0}{\partial x} - \frac{y f_0, x f_0}{f_0} \frac{\partial u_0}{\partial y} \right) \cdot \left( \frac{\partial v}{\partial x} - \frac{y f_0, x f_0}{f_0} \frac{\partial v}{\partial y} \right) + \frac{\beta^2}{f_0^2} \frac{\partial u_0}{\partial y} \cdot \frac{\partial v}{\partial y} \right) - \int_0^T \int_{\Omega_1} \frac{f_0 \nu}{\beta} \left( \frac{2 \beta^2 f_1}{f_0^2} \frac{\partial u_0}{\partial y} \cdot \frac{\partial v}{\partial y} \right) - \int_0^T \int_{\Omega_1} \frac{f_1}{\beta} u_0 \cdot v_t. \]

Problem (4.3.7) is a weak statement for the non-stationary Stokes problem. In the sequel we assume that this problem has a unique solution for any given \( u_0, p_0 \) (the solution in the unperturbed domain \( \Omega_0 \)), and for each \( f_1 \in \mathbb{H}_f \).

4.4 The shape optimization problem: unsteady cost functionals

In this section our mathematical derivation will be formal, in the sense that we allow ourselves the regularity that is necessary for our manipulation. Suppose now that the (time independent) function \( f_1 \) in problem (4.3.7) is unknown and so are \( u_1, p_1 \). To complete problem (4.3.7) we will have to either provide some additional equations or require that \( f_1 \) be determined by minimizing a suitable (time dependent) cost functional.

In general terms, problem (4.3.7) will be supplemented by the additional equation:

\[ C(f, u, p) = 0 \quad (4.4.1) \]

where \( C \) is an operator (linear or nonlinear) defined on \( H^1_0(x_1, x_2) \times Y \times \mathbb{H}^p \). (We consider now \( f \in H^1_0(\Omega) \) for convenience). We assume \( C \) to depend smoothly on its variables \( f, u, p \).

Using the representations (4.2.2) and (4.3.4) we derive from (4.4.1) the following equation:

\[ C(f, u, p) = C_0(f_0, u_0, p_0) + \varepsilon C_1(f_1, u_1, p_1) + \mathcal{O}(\varepsilon^2) = 0, \quad \forall \varepsilon \in [-\varepsilon_0, \varepsilon_0] \quad (4.4.2) \]

where

\[ C_1(f_1, u_1, p_1) := \frac{\partial C}{\partial \varepsilon}(f, u, p)|_{\varepsilon=0}. \quad (4.4.3) \]

If we assume that the data of our problems are such that \( C_0(f_0, u_0, p_0) = 0 \), then we can replace (4.4.2) by the approximate equation

\[ C_1(f_1, u_1, p_1) = 0 \quad (4.4.4) \]

and use it to complete (4.3.7), yielding an exact controllability problem. An alternative approach would consist in replacing (4.4.4) by the following equivalent minimization problem:

\[ \inf_{f_1} \int_0^T \int_{\Omega} \frac{f_0}{\beta} |C_1(f_1, u_1, p_1)|^2 d\Omega dt, \quad (4.4.5) \]
where we assume that $C_1$ has image in $\mathbb{R}^p$. In the next sections we apply the approaches (4.4.4)-(4.4.5) described above for the completion of (4.3.7) and we will move from the following special choice of (4.4.1):

$$C(f, u) := ((\nabla \times u) \circ T_f^{-1})(x, y, t) - R_{obs, c}(x, y, t) \text{ in } \Omega_{wd} \subseteq \Omega \forall t \geq 0,$$  \hspace{1cm} (4.4.6)

where we recall that $\Omega_{wd}$ is a suitable subset of $\Omega$ in which we want our additional equation (or our “control”) to take place. Moreover

$$R_{obs, c} = R_{obs, 0} + \varepsilon R_{obs, 1} + \varepsilon^2 R_{obs, 2} + \ldots, \quad R_{obs, 0} := ((\nabla \times u_0) \circ T_{f_0}^{-1}).$$  \hspace{1cm} (4.4.7)

Then we have: $C_0(f_0, u_0) = 0$, while the equation (4.4.4) corresponding to the special choice (4.4.6) reads:

$$C_1(f_1, u_1) = R(f_0)u_1 + m_1 R_f f_1 - R_{obs, 1} = 0 \text{ in } \Omega_{wd}, \forall t,$$  \hspace{1cm} (4.4.8)

where

$$R(f_0)u_1 = (\nabla \times u_1) \circ T_{f_0}^{-1}(x, y) = \frac{\partial v_1}{\partial x} \frac{y f_{0,x}}{f_0} - \frac{\partial v_1}{\partial y} \frac{f_1}{f_0},$$

$$R_f f_1 := R_f(f_1, u_0) = -y \left( \frac{f_{1,x}}{f_0} \frac{f_0}{f_0} \frac{\partial v_1}{\partial y} \frac{f_1}{f_0} \frac{\partial u_1}{\partial y} + \beta f_1 \frac{\partial u_0}{\partial y} \frac{f_1}{f_0} \frac{\partial v_1}{\partial y} \right).$$

In conclusion we consider the following problem: find $f_1 = (u_1, p_1) \in \mathbb{Y}$, $f_1 \in H_0^1(x_1, x_2)$ s.t.

$$\left\{ \begin{array}{l}
C_1(f_1, u_1) = \mathcal{L}(\Phi_1, \tilde{\Phi}) = B(f_1, \tilde{\Phi}) \forall \tilde{\Phi} = (\nu, q) \in \mathbb{W}, \\
R(f_0)u_1 + m_1 R_f f_1 - R_{obs, 1} = 0 \text{ in } \Omega_{wd} \forall t \geq 0,
\end{array} \right.$$  \hspace{1cm} (4.4.9)

where $R_{obs, 1}$ is a given function (for example in our case it could be $R_{obs, 1} = -R_{obs, 0}$). Problem (4.4.9) is an “exact controllability problem”. These problems have solutions in some particular cases only. For this reason we replace (4.4.9) by the following generalized optimal control problem: find $f_1 = (u_1, p_1) \in \mathbb{Y}$, $f_1 \in H_0^1(x_1, x_2)$ s.t.

$$\left\{ \begin{array}{l}
\mathcal{L}(\Phi_1, \tilde{\Phi}) = B(f_1, \tilde{\Phi}) \forall \tilde{\Phi} = (\nu, q) \in \mathbb{W}, \\
\inf_{f_1} = \frac{\alpha}{2} \|f_1\|^2_{H_0^1(x_1, x_2)} + J(f_1, u_1, p_1),
\end{array} \right.$$  \hspace{1cm} (4.4.10)

where

$$J(f_1, u_1, p_1) = \gamma_1 J_1(f_1, u_1) + \gamma_2 J_2(f_1, u_1, p_1) + \gamma_3 J_3(f_1, u_1, p_1),$$

$$J_1(f_1, u_1) = \frac{1}{2} \int_0^T \int_{\Omega_{wd}} \frac{f_0}{m_{wd}} \frac{\partial v_1}{\partial y} \frac{f_1}{f_0} \frac{\partial u_1}{\partial y},$$

$$\alpha = const \geq 0$$ is a small regularization parameter, $\gamma_1, \gamma_2, \gamma_3$ are non-negative constant weights, $m_{wd}$ is the characteristic function of $\Omega_{wd}$. The functional $J_1(f_1, u_1)$ allows the control of the vorticity over a period $T$ which is, as already seen, a relevant clinical index related also with velocity and viscous dissipation (more specifically to shear stresses), while $J_2(f_1, u_1, p_1)$ and $J_3(f_1, u_1, p_1)$ are additional functionals that are assumed to be quadratic. Note that the second equation in (4.4.9) is considered in (4.4.10) in the least square sense; then (4.4.10) for $\alpha = 0, \gamma_2 = \gamma_3 = 0$ provides the weak statement of problem (4.4.9). For
\( \alpha > 0 \) the solution \( v_1 = v_1(\alpha) \), \( p_1 = p_1(\alpha) \), \( f_1 = f_1(\alpha) \) of (4.4.10) represents an approximate (regularized) solution of (4.4.9).

An example of \( J_2(f_1, u_1, p_1) \) follows:

\[
J_2(f_1, u_1, p_1) = J_2(u_1, p_1) :=
\frac{1}{2} \left( |p_1 - p_{\text{out}, 1}|^2_{L^2(\Gamma_{\text{out}} \times (0, T))} + |u_1 - u_{\text{out}, 1}|^2_{L^2(\Gamma_{\text{out}} \times (0, T))^2} \right)
\]

The functional \( J_3 \) is introduced in order to enhance the smoothness in time of \( u_1, p_1 \). There we will take:

\[
J_3(f_1, u_1, p_1) = J_3(\tau; p_1) := \frac{1}{2} \| J(p_1 - p_{\text{out}, 1}) \|^2_{L^2(\Gamma_{\text{out}} \times (0, T))},
\]

where

\[
J(p) := \frac{p(x, y, t) - p(x, y, t - \tau)}{\tau},
\]

where \( \tau \geq 0 \) is a parameter, while it is assumed that the functions \( p_1, p_{\text{out}, 1} \) be extended by parity to the negative values of \( t \) (i.e. \( p(x, y, t) := p(x, y, -t) \) as \( t < 0 \)). If \( p_1, p_{\text{out}, 1} \in H^1(0, T; L^2(\Omega)) \) and \( \tau \to 0 \) then \( J_3 \to \frac{1}{2} \| p_{1,t} - (p_{\text{out}, 1,t}) \|^2_{L^2(\Gamma_{\text{out}} \times (0, T))} \). Otherwise said, \( J_3 \) in fact imposes a regularity restriction on \( p_1 \). If \( \tau \to \infty \), then \( J_3 \to 0 \) and no additional restriction holds on \( p_1 \). If \( 0 < \tau < \infty \) then the introduction of \( J_3 \) can be regarded as a tool that yields a regularization condition for \( p_1 \) on \( \Gamma_{\text{out}} \times (0, T) \). (Of course there are other ways to introduce similar regularity restrictions).

### 4.5 The optimal control formulation for unsteady flows

When considering (4.4.10) we have considered the simple domain of Figure 4.4, that we have previously denoted with \( \Omega \). An alternative possibility (that we are going to follow) consists of using the new variable transformation

\[
x = T_{f_0}^{-1}(\tilde{x}), \quad \tilde{x} \in \Omega, \quad x \in \Omega_0,
\]

which is the identity in \( \Omega_2 \), while \( T_{f_0}^{-1}(\tilde{x}, \tilde{y}) = (\tilde{x}, \frac{f_0(\tilde{x})}{f_1(\tilde{x})} \tilde{y}) \) in \( \Omega_1 \), then working in the “unperturbed” domain \( \Omega_0 \) (see Figure 4.5), as done in the previous chapter, the expressions for the bilinear forms in (4.4.10) become simpler.

Indeed in \( \Omega_0 \times (0, T) \) problem (4.4.10) can be reformulated as follows: find \( \Phi = (u, p) := \Phi_1 = (u_1, p_1) \in \mathcal{Y}, \quad f := f_1 \in \mathbb{H}_f^2 \), such that

\[
\begin{align*}
\mathcal{L}(\Phi, \Phi) &= B(f, \Phi) \quad \forall \Phi := (v, q) \in \mathcal{W},
\inf_{f \in \mathbb{H}_f} &= \frac{1}{2} \| f \|^2_{H_0^1(x_1, x_2)} + J(f, \Phi),
\end{align*}
\]

where

\[
\mathcal{L}(\Phi, \Phi) = a_0(u, v) - b_0(p, v) + b_0(q, v),
\]

\( ^2 \)From now on we denote \( u_1 = u, \quad p_1 = p, \quad f_1 = f \) however we should keep in mind that now \( u, p, f \) represent the “first corrections” of \( u_0, \quad p_0, \quad f_0 \) on the unperturbed domain.
be a solution of (4.5.2), then

$$ B(f, \hat{\Phi}) := b_f(f, p_0, v) + G_1(f, v) - a_f(f, u_0, v) - b_f(f, q, u_0), $$

$$ a_0(u, v) = \int_0^T \int_{\Omega_0} \nu \left( \frac{\partial u}{\partial x} \cdot \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \cdot \frac{\partial v}{\partial y} \right) - \int_0^T \int_{\Omega_0} u \cdot v_t, $$

$$ b_0(p, v) = \int_0^T \int_{\Omega_0} p \nabla \cdot v, $$

$$ b_f(f, p_0, v) = \int_0^T \int_{\Omega_{0,1}} p_0 D_f(f, v) + \int_0^T \int_{\Omega_{0,1}} \frac{f}{f_0} p_0 \nabla \cdot v, $$

$$ D_f(f, v) = - \left[ y\left( \frac{f_x f_0 - f_{0,x} f}{f_0^2} \right) \frac{\partial u}{\partial y} + \frac{f}{f_0} \frac{\partial v}{\partial y} \right], $$

$$ D_f(f, v_0) := D_f f, $$

$$ G_1(f; v) = \int_0^T \int_{\Omega_{0,1}} \frac{f}{f_0} \cdot v + \int_0^T \int_{\Omega_{0,1}} \frac{f}{f_0} u_0 \cdot v(x, y, 0) d\Omega, $$

$$ a_f(f; u_0, v) = \int_0^T \int_{\Omega_{0,1}} \frac{f v}{f_0} \nabla u_0 \cdot \nabla v - \int_0^T \int_{\Omega_{0,1}} \nu y \left( \frac{f_x f_0 - f_{0,x} f}{f_0^2} \right) \left( \frac{\partial u_0}{\partial y} \cdot \frac{\partial v}{\partial x} + \frac{\partial u_0}{\partial x} \cdot \frac{\partial v}{\partial y} \right) + $$

$$ - \int_0^T \int_{\Omega_{0,1}} \frac{2 f v}{f_0} \frac{\partial u_0}{\partial y} \cdot \frac{\partial v}{\partial y}, $$

$$ J_1(f, u, p) = \gamma_1 J_1(f, u) + \gamma_2 J_2(f, u, p) + \gamma_3 J_3(\tau; p), $$

$$ J_1(f, u) = \frac{1}{2} \int_0^T \int_{\Omega_0} m_{wd} |\nabla \times u| + m_1 R_f f - R_{obs,1}|^2, $$

$$ \mathcal{R}_f f := \mathcal{R}_f(f, u_0) = -y \left( \frac{f_x f_0 - f_{0,x} f}{f_0^2} \right) \frac{\partial v_0}{\partial y} + \frac{f}{f_0} \frac{\partial u_0}{\partial y}, $$

$J_2(f, u, p)$ and $J_3(\tau, p)$ are defined similarly. Let us derive the operator form of problem (4.5.2). Should $\hat{\Phi}$ be a solution of (4.5.2), then

$$ \alpha(f, \hat{f})_{\mathcal{H}_f} + \langle J'_\Phi(f, \hat{\Phi}, \hat{\Phi}_f) \rangle + \langle J'_f(f, \hat{\Phi}, \hat{f}) \rangle = 0, $$

for any $\hat{f} \in \mathcal{H}_f$ ($\hat{f}$ is the independent variation), where $\hat{\Phi}_f \in \mathcal{W}$ satisfies the following equation:

$$ \mathcal{L}(\hat{\Phi}_f, \hat{\Phi}) = B(f, \hat{\Phi}) \forall \hat{\Phi} \in \mathcal{W}. $$

Figure 4.5: Transformation of the simple domain into the unperturbed domain $\Omega_0$. 
4.5. THE OPTIMAL CONTROL FORMULATION FOR UNSTEADY FLOWS

In (4.5.4), \( J'_\Phi = \frac{\partial J}{\partial \Phi} \) and \( J'_f = \frac{\partial J}{\partial f} \) are partial derivatives of \( J \), while \( \langle Q, \Phi \rangle \) stands for \( \mathcal{W} \langle Q, \Phi \rangle \mathcal{W}^* \) the duality between \( \mathcal{W} \) and \( \mathcal{W}^* \) and \( (g, f) \) for the duality \( \mathcal{H}_f(g, f) \mathcal{H}^*_f \) between \( \mathcal{H}_f \) and \( \mathcal{H}^*_f \). Then we can write for (4.5.2) the system of “optimality conditions”:

\[
\begin{align*}
\mathcal{L}(\Phi, \dot{\Phi}) &= B(f, \dot{\Phi}) \quad \forall \Phi \in \mathcal{W}, \\
\alpha(f, \dot{f})_{\mathcal{H}_f} + \langle J'_\Phi(f, \Phi), \dot{\Phi} \rangle + \langle J'_f(f, \Phi), \dot{f} \rangle &= 0 \quad \forall \dot{f} \in \mathcal{H}_f. 
\end{align*}
\] (4.5.6)

The element \( \Phi_f \) can be eliminated from (3.5.6) by introducing the adjoint problem: find \( \hat{Q} := (q, \sigma) \in \mathcal{W} \) s.t.

\[
\mathcal{L}^*(\hat{Q}, \hat{W}) := \mathcal{L}(\hat{W}, \hat{Q}) = \langle J'_\Phi(f, \Phi), \hat{W} \rangle \quad \forall \hat{W} \in \mathcal{Y}. 
\] (4.5.7)

Since \( \Phi_f \in \mathcal{Y} \) we can choose \( \hat{W} = \Phi_f \) in (4.5.7), yielding

\[
\langle J'_\Phi(f, \Phi), \Phi_f \rangle = \mathcal{L}(\Phi_f, \hat{Q}) = B(f, \hat{Q}) 
\] (4.5.8)

and the system of variational equations (4.5.6) reads now as follows:

\[
\begin{align*}
\mathcal{L}(\Phi, \dot{\Phi}) &= B(f, \dot{\Phi}) \quad \forall \Phi \in \mathcal{W}, \\
\mathcal{L}^*(\hat{Q}, \hat{W}) &= \langle J'_\Phi(f, \Phi), \hat{W} \rangle \quad \forall \hat{W} \in \mathcal{Y}, \\
\alpha(f, \dot{f})_{\mathcal{H}_f} + B(f, \dot{Q}) + \langle J'_f(f, \Phi), \dot{f} \rangle &= 0 \quad \forall \dot{f} \in \mathcal{H}_f. 
\end{align*}
\] (4.5.9)

The first equation is the state equation. By means of the following operators (see Lions and Magenes [84], Lions [82] and Agoshkov [6]):

\[
\begin{align*}
L : \mathcal{Y} &\to \mathcal{W}^*, \quad (L\Phi, \dot{\Phi})_{\mathcal{H}_0} := \mathcal{L}(\Phi, \dot{\Phi}), \\
L^* : \mathcal{W} &\to \mathcal{Y}^*, \quad (L^*\dot{Q}, \dot{Q})_{\mathcal{H}_0} = (L\dot{W}, \dot{W})_{\mathcal{H}_0}, \\
B : \mathcal{H}_f &\to \mathcal{W}^*, \quad (Bf, \Phi)_{\mathcal{H}_0} = B(f, \Phi) \\
\Lambda_w : \mathcal{Y}^* &\to \mathcal{Y}^*, \quad (\Lambda_w J(f, \Phi), \hat{W})_{\mathcal{H}_0} := \langle J'_\Phi(f, \Phi), \hat{W} \rangle, \\
\Lambda_f : \mathcal{H}^*_f &\to \mathcal{H}^*_f, \quad (\Lambda_f J_f(f, \Phi), \hat{f})_{L^2(x_1, x_2)} := \langle J'_f(f, \Phi), \hat{f} \rangle,
\end{align*}
\]

the system (4.5.9) can be written in operator form as follows:

\[
\begin{align*}
L\Phi &= Bf \quad \text{(in } \mathcal{W}^*), \\
L^*\dot{Q} &= \Lambda_w J_f(f, \Phi) \quad \text{(in } \mathcal{Y}^*), \\
\alpha \Lambda_c f + B^* \dot{Q} + \Lambda_f J_f(f, \Phi) &= 0 \quad \text{(in } \mathcal{H}^*_f),
\end{align*}
\] (4.5.10)

there \( \Lambda_c \) is the extension to \( \mathcal{H}^*_f \) of the operator:

\[
\Lambda_{c,0} f := -f_{xx} + f
\]

whose domain is \( D(\Lambda_{c,0}) = H^2(\Omega) \cap \mathcal{H}_f \).
4.6 Existence and uniqueness results for unsteady case

We analyze the particular cases where the cost functional $J$ is chosen as outlined in Section 4.4, and we extend uniqueness and existence results proven to the unsteady case. Let $J$ be a weighted sum of the functionals $J_1, J_2, J_3$ of Section 4.4. Then

$$ J(f, \Phi) = J(f, u, p) = \frac{\gamma_1}{2} \int_0^T \int_{\Omega_0} m_{wd} |\nabla \times u + m_1 R_f f - R_{obs,1}|^2 \, dt \, d\Omega + \frac{\gamma_2}{2} \int_0^T \int_{\Gamma_{out}} (|p - p_{out}|^2 + |u - u_{out}|^2) \, d\Gamma dt + \frac{\gamma_3}{2} \int_0^T \int_{\Gamma_{out}} |J(p - p_{out})|^2 d\Gamma dt $$

(4.6.1)

We assume that $\Omega_{wd} = \Omega_0$ and we define the spaces:

$$ \mathbb{X} := \{ u : u \in (H^2(\Omega))^2, u = 0 \text{ on } \Gamma_{in} \cup \Gamma_{w1} \cup \Gamma_{w3} \} $$

$$ \mathbb{H} := L^2(0, T; H^1(\Omega_0)), \mathbb{H}_f := H^2(x_1, x_2) \cap H^1_0(x_1, x_2). $$

Also here we pretend that the velocity be in $H^2$ in order to use the uniqueness continuation theorem. The derivatives $J'_f(f, \Phi)$ and $J'_f(f, \Phi)$ become

$$ \langle J'_f(f, \Phi), \hat{f} \rangle = \gamma_1 \int_0^T \int_{\Omega_0} m_{wd}(\nabla \times u + m_1 R_f f - R_{obs,1}) \cdot (\nabla \times \hat{v}) \, dt \, d\Omega + $$

$$ + \gamma_2 \int_0^T \int_{\Gamma_{out}} (p - p_{out}) q \, d\Gamma dt + \gamma_2 \int_0^T \int_{\Gamma_{out}} (u - u_{out}) \cdot \hat{v} \, d\Omega dt + $$

$$ + \gamma_3 \int_0^T \int_{\Gamma_{out}} J(p - p_{out}) \cdot \hat{J}(q) \, d\Gamma dt, $$

$$ \langle J'_f(f, \Phi), \hat{f} \rangle = \gamma_1 \int_0^T \int_{\Omega_0} m_{wd}(\nabla \times u + m_1 R_f f - R_{obs,1}) R_f \hat{f} \, dt \, d\Omega, $$

$$ \forall \Phi = (v, q) \text{ and } \forall \hat{f}. $$

The system of variational equations (4.5.6) becomes: find $\Phi_f = (u_f, p_f) \in \mathbb{Y} \times \mathbb{H}$ s.t.

$$ \begin{aligned}
\mathcal{L}(\Phi_f, \hat{\Phi}) &= B(f, \hat{\Phi}) \forall \hat{\Phi} \in \mathbb{Y} \times \mathbb{H}, \\
\mathcal{L}^*(\hat{Q}, \hat{W}) &= \langle J'_f(f, \Phi), \hat{W} \rangle \forall \hat{W} \in \mathbb{Y} \times \mathbb{H}, \\
\alpha(f, \hat{f}) \hat{f} + B(f, \hat{Q}) + \\
+ m_1 R_f \hat{f} \, d\Omega dt + \gamma_2 \int_0^T \int_{\Gamma_{out}} ((p_f - p_{out}) p_f + (u_f - u_{out}) \cdot u_f) \, d\Gamma dt + \\
+ \gamma_3 \int_0^T \int_{\Gamma_{out}} J(p_f - p_{out}) \hat{J}(p_f) d\Gamma dt = 0 \forall \hat{f} \in \mathbb{H}_f,
\end{aligned} $$

(4.6.2)

where for every $\hat{f}$, $u_f = u_f(\hat{f})$, $p_f = p_f(\hat{f})$ denote the solution of the system given by the first equation in (4.6.2) corresponding to a right end side $f = \hat{f}$. The system (4.5.9) becomes: find $\Phi_f = (u_f, p_f) \in \mathbb{Y} \times \mathbb{H}, Q = (q, \sigma) \in \mathbb{Y} \times \mathbb{H}$ s.t.

$$ \begin{aligned}
\mathcal{L}(\Phi_f, \hat{\Phi}) &= B(f, \hat{\Phi}) \forall \hat{\Phi} \in \mathbb{Y} \times \mathbb{H}, \\
\mathcal{L}^*(\hat{Q}, \hat{W}) &= \langle J'_f(f, \Phi), \hat{W} \rangle \forall \hat{W} \in \mathbb{Y} \times \mathbb{H}, \\
\alpha(f, \hat{f}) \hat{f} + B(f, \hat{Q}) + \\
+ m_1 R_f \hat{f} \, d\Omega dt + \gamma_2 \int_0^T \int_{\Gamma_{out}} ((p_f - p_{out}) p_f + (u_f - u_{out}) \cdot u_f) \, d\Gamma dt + \\
+ \gamma_3 \int_0^T \int_{\Gamma_{out}} J(p_f - p_{out}) \hat{J}(p_f) d\Gamma dt = 0 \forall \hat{f} \in \mathbb{H}_f,
\end{aligned} $$

(4.6.3)
§4.6. EXISTENCE AND UNIQUENESS RESULTS FOR UNSTEADY CASE

where

$$\langle J_k(f, \Phi), W \rangle = \int_0^T \int_{\Omega} m_{wd}(\nabla \times u + m_1 R_f f - R_{obs,1}) \cdot (\nabla \times \dot{q}) d\Omega dt +$$

$$+ \gamma_2 \int_0^T \int_{\Gamma_{out}} (u_f - u_{out}) \cdot \dot{q} d\Gamma dt + \gamma_2 \int_0^T \int_{\Gamma_{out}} (p_f - p_{out}) \dot{\sigma} d\Gamma dt +$$

$$+ \gamma_3 \int_0^T \int_{\Gamma_{out}} J(p_f - p_{out}) \cdot J(\dot{\sigma}) d\Gamma dt.$$  

Consider now the problem (4.6.3) for \( \alpha > 0 \).

**Proposition 4.6.1** For any \( \alpha > 0 \) problem (4.6.3) has a unique solution for any given \( R_{obs,1} \).

**Proof.** Following [6] and what already seen in the previous chapter for steady flow, we formally invert \( L \) and \( L^* \) in the first and second equations of (4.5.10), then we substitute \( \Phi, Q \) into the third equation and we obtain the following weak problem, \( f \in \mathbb{H}_f \) satisfies:

$$\alpha(f, \hat{f})_{\mathbb{H}_f} + (Af, A\hat{f})_{L^2(x_1, x_2)} = (G, A\hat{f})_{L^2(x_1, x_2)} \forall \hat{f} \in \mathbb{H}_f. \tag{4.6.4}$$

\( A \) is a linear operator, while \( G \) depends on the data. Precisely, from (4.6.2) we obtain:

$$(f, \hat{f})_{\mathbb{H}_f} = (Af, \hat{f})_{L^2(x_1, x_2)},$$

$$(Af, A\hat{f})_{L^2(x_1, x_2)} = \gamma_1 \int_0^T \int_{\Omega} m_{wd}(\nabla \times u + m_1 R_f f) \cdot (\nabla \times u_f + m_1 R_f \dot{f}) d\Omega dt +$$

$$+ \gamma_2 \int_0^T \int_{\Gamma_{out}} (p_f + u \cdot u_f) d\Gamma dt + \gamma_3 \int_0^T \int_{\Gamma_{out}} J(p) \cdot J(p_f) d\Gamma dt,$$

$$(G, A\hat{f})_{L^2(x_1, x_2)} = \gamma_1 \int_0^T \int_{\Omega} m_{wd}R_{obs,1} \cdot (\nabla \times u) + m_1 R_f \dot{f}) d\Omega dt +$$

$$+ \gamma_2 \int_0^T \int_{\Gamma_{out}} (p_{out} + p_f) d\Gamma dt + \gamma_3 \int_0^T \int_{\Gamma_{out}} J(p_{out}) \cdot J(p_f) d\Gamma dt,$$

where \( \Phi = (u, p) = L^{-1} Bf, \Phi_f = (u_f, p_f) = L^{-1} B \hat{f}, \forall \hat{f} \in \mathbb{H}_f \).

We see that if \( \alpha > 0 \) then problem (4.6.4) has a unique solution which satisfies \( \|f\|_{\mathbb{H}_f}^2 \leq \|G\|^2/(2\alpha) < \infty \). Correspondingly, we can construct \( u, p, q, \sigma \), which together with \( f \) provide the unique solution of (4.6.3). \( \square \)

Consider now problem (4.6.3) with \( \alpha = 0 \).

**Proposition 4.6.2** Assume that:

i) The solution of the generalized non-stationary Stokes problem (4.3.7) satisfies:

$$\left( \frac{\partial v}{\partial y} \right)^2 + \left( \frac{\partial u_0}{\partial y} \right)^2 > 0 \text{ at } y = 0, \ x \in (x_1, x_2);$$
ii) problem (4.6.3) has a solution in the class \( L^2(0, T; H^2(\Omega)^2) \times L^2(0, T; H^1(\Omega)) \times W^{1,\infty}(x_1, x_2) \). Then this solution is unique.

Proof. Let \((u_1, p_1, f_1)\) and \((u_2, p_2, f_2)\) be two solutions of (4.6.3). Then for \( u = u_1 - u_2, p = p_1 - p_2, f = f_1 - f_2 \) from (4.6.2) we obtain:

\[
\begin{align*}
\begin{aligned}
a_0(u, v) - b_0(p, v) &= F(f, v) \quad \forall v \in W, \\
b_0(q, u) + b_1(f; q, u_0) &= 0 \quad \forall q \in \mathbb{H}^p, \\
\nabla \times u + m_1 \mathcal{R} f &= 0 \quad \text{in} \quad \Omega \times (0, T), \\
p &= 0, \quad u = 0 \quad \text{on} \quad \Gamma_{out} \times (0, T).
\end{aligned}
\end{align*}
\] (4.6.5)

Consider the classical form of the second and the third equation from (4.6.5) in \( \Omega_{2,0} \times (0, T) \)

\[
\nabla \cdot u = 0, \quad \nabla \times u = 0 \quad \text{in} \quad \Omega_{2,0} \times (0, T).
\]

Then \( \Delta u = 0 \) in \( \Omega_{2,0} \) \( \forall t \in (0, T) \). Considering \( v \) with \( \text{supp}(v) \subseteq \Omega_{2,0} \) from the first equation of (4.6.5) we find \( \nabla p = 0 \), then \( p = \text{const} \) in \( \Omega_{2,0} \) and \( -\nu \mathbf{n} + \nu \partial u \partial \mathbf{n} = 0 \) on \( \Gamma_{out} \) \( \forall t \). Since \( p = 0 \) on \( \Gamma_{out} \) then \( p = 0 \) in \( \Omega_{2,0} \) and \( \nu \partial u \partial \mathbf{n} = 0 \) on \( \Gamma_{out} \). Consequently, for all \( t \in (0, T) \), \( u \) satisfies:

\[
\Delta u = 0 \quad \text{in} \quad \Omega_{2,0}, \quad u = \nu \partial u \partial \mathbf{n} = 0 \quad \text{on} \quad \Gamma_{out}.
\]

Owing to the uniqueness continuation theorem this Cauchy problem has only the trivial solution \( u = 0 \) in \( \Omega_{2,0} \). Since \( u \in L^2(0, T; H^2(\Omega)^2) \) then

\[
u \frac{\partial u}{\partial \mathbf{n}} = 0 \text{ on } \Gamma_t := \{(x, y) : y = 0, x_1 < x < x_2\}, \quad \forall t \in (0, T).
\]

Consider now the second and third equations from (4.6.5) in \( \Omega_{1,0} \) in their classical form, \( \forall t \in (0, T) \):

\[
\begin{align*}
\begin{aligned}
\nabla \cdot u &= \left[ y \left( \frac{f_x f_0 - f_0 f_x}{f_0^2} \right) \frac{\partial u_0}{\partial y} + \frac{f}{f_0} \frac{\partial u_0}{\partial y} \right] = 0 \quad \text{in} \quad \Omega_{1,0}, \\
\nabla \times u &= \left[ y \left( \frac{f_x f_0 - f_0 f_x}{f_0^2} \right) \frac{\partial u_0}{\partial y} - \frac{f}{f_0} \frac{\partial u_0}{\partial y} \right] = 0 \quad \text{in} \quad \Omega_{1,0}.
\end{aligned}
\end{align*}
\] (4.6.6)

On \( \Gamma_0 \) we have:

\[
\nabla \cdot u - \frac{f}{f_0} \frac{\partial v_0}{\partial y} = 0, \quad \nabla \times u + \frac{f}{f_0} \frac{\partial u_0}{\partial y} = 0,
\]

\[
|f(x)| = f_0 \left[ \left( \frac{\nabla \cdot u}{\partial y} \right)^2 + \left( \nabla \times u \right)^2 \right]^{1/2} \quad \text{on} \quad \Gamma_0,
\]

the dependence of the right end side on \( x \) and \( y \) is understood). Since \( u = \frac{\partial u}{\partial \mathbf{n}} = \frac{\partial u}{\partial y} = 0 \text{ on } \Gamma_0 \), then

\[
\nabla \cdot u |_{y=0} = \frac{\partial u}{\partial x} |_{y=0} = 0, \quad \nabla \times u |_{y=0} = \frac{\partial u}{\partial y} |_{y=0} = 0, \quad x \in (x_1, x_2).
\]

i.e. \( f(x) = 0 \). Therefore, \( u = 0 \), \( p = 0 \) too, \( \forall t \in (0, T) \). \( \square \)
4.7 The iterative process in the non-stationary case

In this section we propose some iterative processes for solving the space-time variational equations obtained in the previous sections. These algorithms are an extension of those introduced in the previous chapter. Consider the problem (4.5.10); if for \( k = 0, 1, \ldots \) \( f^{(k)} \) is known, then \( f^{(k+1)} \) can be determined by solving the following equations [6]:

\[
\begin{align*}
L\Phi^{(k)} &= Bf^{(k)}, \\
L^*Q^{(k)} &= \Lambda_w J^*(f^{(k)}, \Phi^{(k)}), \\
\Lambda_w u^{(k)} &= B^*Q^{(k)} + \Lambda_f J_f(f^{(k)}, \Phi^{(k)}), \\
f^{(k+1)} &= f^{(k)} - \tau_k(\alpha f^{(k)} + w^{(k)}),
\end{align*}
\]

where \( \{\tau_k\} \) is a family of parameters whose determination follows from the theory of extremal problems (see Vasiliev [155]), the general theory of iterative processes (see Marchuk [94], Quarteroni and Valli [127], Quarteroni, Sacco and Saleri [129]), and the ill-posed problems theory (Tikhonov and Arsenin [151], Vainikko and Veretennikov [153]). Its variational form reads as:

\[
\begin{align*}
a_0(u^{(k)}, v) - b_0(p^{(k)}, v) &= F(f^{(k)}, v) \forall v \in W, \\
b_0(q, u^{(k)}) &= -b_f(f^{(k)}; q, u_0) \forall q \in \mathbb{H}^p, \\
a_0(q, q^{(k)}) + b_0(q^{(k)}, \hat{q}) &= G_k(q) \forall q \in Y, \\
- b_0(\sigma, q^{(k)}) &= g_k(\sigma) \forall \sigma \in \mathbb{H}^p, \\
(w^{(k)}, f)_{\mathbb{H}_f} &= d_k(f) \forall f \in \mathbb{H}_f, \\
f^{(k+1)} &= f^{(k)} - \tau_k(\alpha f^{(k)} + w^{(k)}), \quad k = 0, 1, \ldots,
\end{align*}
\]

where

\[
F(f^{(k)}, v) = b_f(f^{(k)}, p_0, v) + G_1(f^{(k)}, v) - a_f(f^{(k)}, u_0, v),
\]

\[
G_k(q) = \gamma_1 \int_0^T \int_{\Omega_0} m_{\text{wd}}(\nabla \times u^{(k)} + m_1 \mathcal{R}_f f^{(k)} - \mathcal{R}_{\text{obs},1}) \cdot (\nabla \times q) d\Omega dt + \\
+ \gamma_2 \int_0^T \int_{\Gamma_{\text{out}}} (u^{(k)} - u_{\text{out}}) \cdot \hat{q} d\Gamma dt,
\]

\[
g_k(\sigma) = \gamma_2 \int_0^T \int_{\Gamma_{\text{out}}} (p^{(k)} - p_{\text{out}}) \hat{\sigma} d\Gamma dt + \gamma_3 \int_0^T \int_{\Gamma_{\text{out}}} \mathcal{J}(p^{(k)} - p_{\text{out}}) \cdot \hat{\sigma} d\Gamma dt,
\]

\[
d_k(f) = F(\tilde{f}, q^{(k)}) - b_f(\tilde{f}; q^{(k)}, u_0) + \\
+ \gamma_1 \int_0^T \int_{\Omega_0} m_{\text{wd}}(\nabla \times u^{(k)} + m_1 \mathcal{R}_f f^{(k)} - \mathcal{R}_{\text{obs},1}) m_1 \mathcal{R}_f \tilde{f} d\Omega \Omega dt.
\]

Consider now the \textit{finite dimensional case} in which the function \( f, \{f^{(k)}\}, \tilde{f} \) all are sought for in a finite-dimensional subspace \( \mathbb{H}_{f,N} \subset \mathbb{H}_f \) of dimension \( N < \infty \), whose basis \( \varphi_i \in \mathbb{W}^{1,\infty}(x_1, x_2), i = 1, 2, \ldots, N \). Then the following theorem holds true.
Theorem 4.7.1 Assume that:

\[ \Omega_{\text{wd}} = \Omega, \]
\[ \left( \frac{\partial v_0}{\partial y} \right)^2 + \left( \frac{\partial u_0}{\partial y} \right)^2 > 0 \text{ at } y = 0, \; x \in (x_1, x_2). \]

Then:

1. For \( \alpha \geq 0 \) and any \( N < \infty \) the problem (4.6.2) is well posed (i.e. it admits a unique solution that depends continuously on the data);

2. The iterative process (4.7.2) is convergent for any \( \alpha > 0, \; N < \infty \), provided the parameters \( \tau_k > 0, \; k = 0, 1, 2, \ldots \) are small enough;

3. If \( \alpha \) is sufficiently small while \( k \) is sufficiently large, then \( \{u^{(k)}, p^{(k)}, f^{(k)}\} \) can be regarded as an approximate solution of problem (4.6.2).

Proof:

1. For \( \alpha > 0 \) the existence was proven earlier in Proposition 4.6.1. Let us consider the case \( \alpha = 0 \), since \( f = \sum_{i=1}^{N} a_i \varphi_i \in \mathbb{H}_{f,N} \), then (4.6.4) is well posed (because problem (4.6.2) can have only unique solution in \( \mathcal{X} \times \mathbb{H}^p \times \mathbb{H}_{f} \), see Proposition 4.6.2). We assume the unsteady Stokes problem to be well posed for given \( f \in \mathbb{H}_{f} \). Hence the problem (4.6.2) is well posed too.

2. If \( \alpha > 0 \) then the bilinear form on the left hand side of (4.6.4) is coercive and continuous with respect to the norm

\[ \|f\|_{A,\alpha} = \sqrt{\alpha \|f\|_{\mathbb{H}_{f}}^2 + \|A f\|_{L^2(x_1,x_2)}^2}. \]

Then the process given by

\[ (f^{(k+1)}, \hat{f})_{\mathbb{H}_f} = (f^{(k)}, \hat{f})_{\mathbb{H}_f} - \tau \alpha (f^{(k)}, \hat{f})_{\mathbb{H}_f} + (A f^{(k)}, A \hat{f})_{L^2(x_1,x_2)} - (G, A \hat{f})_{L^2(x_1,x_2)}, \; k = 0, 1, \ldots \]

is convergent for small \( \tau > 0 \). Hence also the process (3.7.3) is convergent and

\[ \|u^{(k)} - u\|_{\mathcal{Y}} + \|p^{(k)} - p\|_{\mathbb{H}^p} + \|f - f^{(k)}\|_{\mathbb{H}_f} \to 0, \; k \to \infty. \quad (4.7.3) \]

If \( A^{-1} A^* A \in [C_1, C_2], C_1, C_2 = const \), choosing \( \tau_k = 2/(2\alpha + C_1 + C_2) \) we obtain (see Agoshkov [6]):

\[ \|u^{(k)} - u\|_{\mathcal{Y}} + \|p^{(k)} - p\|_{\mathbb{H}^p} + \|f - f^{(k)}\|_{\mathbb{H}_f} \leq C \left( \frac{C_2 - C_1}{2\alpha + C_1 + C_2} \right)^k \quad (4.7.4) \]

which tends to zero as \( k \to \infty \).
3. Let $u_0$, $p_0$, $f_0$ be a solution of (3.6.2) when $\alpha = 0$. According to the theory of ill-posed problems ([151] and [153]) we have: $\|f_0 - f_\alpha\|_{H^1} \to 0$ as $\alpha \to 0^+$, where $(f_\alpha, u_\alpha, p_\alpha)$ is the solution of (3.6.2) for $\alpha > 0$. Hence

$$\|u_0 - u_\alpha\|_V + \|p_0 - p_\alpha\|_{W^1} \to 0, \text{ as } \alpha \to 0^+.$$ 

Owing to (4.7.3) this concludes our proof. $\square$

### 4.8 Numerical results for unsteady generalized Stokes problem

To test and extend our methodology we consider some test problems on the initial configuration already used in the previous test case.

Velocity values $u_{in}$ at the inflow are chosen in such a way that the Reynolds number $Re = \frac{\bar{u}D}{v}$ be of the order $10^3$, the mean Reynolds number is 1250, the maximum is 2500. The inlet Poiseuille velocity profile has a pulsatile nature over the period $T = 1$ s (heart beat) and the law considered was: $u_{in} = -0.475(y - 1)(y - 2)(1 - t)\hat{n}$, see Figure 4.8. Blood kinematic viscosity $\nu = \frac{\mu}{\rho}$ is equal to $4 \cdot 10^{-6} \text{ m}^2 \text{ s}^{-1}$, blood density $\rho = 1 \text{ g cm}^{-3}$ and dynamic viscosity $\mu = 4 \cdot 10^{-2} \text{ g cm}^{-1} \text{ s}^{-1}$; $\bar{u}$ is the mean inflow velocity $\bar{u} = \left( \frac{\int_{t_{in}}^{t_{out}} u_{in} \hat{n} d\Gamma dt}{\int_{t_{in}}^{t_{out}} d\Gamma dt} \right)$ which yields the desired Reynolds number, while $D$ is the arterial diameter (3.5 mm), see Quarteroni and Formaggia [126], Quarteroni, Tuveri and Veneziani [125].

Dealing with unsteady blood flow into the coronaries (from the Aorta artery) we have to face some difficulties and fix some hypotheses to model the pulsatile flow and the fluid mechanical phenomena in the endograft region (see for preliminary considerations Bertolotti et al. [18]). The first aspect deals with the fact that we consider a complete stenosis into the artery (the flow rate from the stenosed section is set to zero), and this yields a retrograde flow, see Kute and Vorp [74]. The second aspect deals with the way blood flow is modelled into the coronaries. Being embedded on the heart surface muscle, coronaries do not share the same behavior similar the other arteries during systolic and diastolic phases. In fact, during systolic phase, the left ventricle contracts and squeezes the coronaries. To take that into account, we have adopted the same pulsatile flow at the inlet of the bypass bridge as the one of the aortic vessel from the opening of the coronaries to the end of the period. The flow-rate function goes from a maximum quantity to zero. For more details we suggest to see Sankaranarayanan et al. [140]. Figures 4.6 and 4.7 show the flow-rate waveform for the Aorta and the left coronary, respectively; we can see that the flow-rate in the aorta goes from zero to a maximum and then to zero in the systolic phase and remains at zero during diastolic phase. To model blood flow into the bypass bridge we have adopted the aorta flow-rate waveform but from the time step in which the coronaries are open (and not squeezed): for this reason in the bypass we have a flow-rate waveform starting from a maximum and going to zero.
In this section we present numerical results using as cost functional $J(f, u) = J_1(f, u)$ (introduced in Equation (4.5.3)). This is equivalent to the $L^2$ norm of the vorticity on $(\Omega_{wd} \times (0, T))$ (restricted in the downfield zone of the new incoming branch of the bypass, where the observation is made). We have set $R_{obs,1} = -0.45 \cdot R(f_0)u$.

For the space approximation of the unsteady Stokes equations we use $P^1 - P^1$ (piecewise linear) finite elements and SUPG stabilization (see [127]). Time discretization is based on first order backward differentiation (which is unconditionally stable).

In figures 4.9-4.14 we report the numerical results obtained and show how the shape of the bypass is changed to reduce downfield vorticity. The shape is smoothed out at the upper corner and a slightly cuffed incoming branch is created. These results provide a shape which resembles the Taylor patch (see Cole et al. [26]). We have shown original bypass configuration (and horizontal velocity to show relevant fluid dynamics phenomena) at different time-steps ($t_1 = 0.1$ s, $t_2 = 0.4$ s, $t_3 = 0.7$ s) and then the configurations obtained at different iterations of our optimization process (i.e. at $N_1 = 5$, $N_2 = 11$ and $N_3 = 17$). These results can be regarded as an improvement of previous results that were obtained using a steady fluid flow model. The similarity between the present results and those obtained in [4] (and introduced in the previous chapter) can be ascribed to the fact that shape $f^{(k)}$ does not depend on time, moreover the shape variation at each optimization step $\delta f^{(k)}$ is given by the sum of the contributions from all time-steps in $(0, T)$ and the first contribution is the one that dominates.
(when considering a pulsatile flow). In these simulations we can find the typical flow reversal caused by the stenosed zone and the other related aspects described for instance in Lei et al. [79]. Figure 4.15 shows pointwise vorticity ([1/s]) in bypass configuration before and at the end of the control process. Vorticity is diminished near the upper corner (in the original configuration we have a concentrated vorticity at the singularity) but also in the bed of the artery in the down-field zone: this indicates that the flow is less disturbed and the main flow decreases its attitude to recirculate in the stenosed zone. These phenomena are due to the fact that the bypass section is smoothed and increased and the flow is guided more smoothly through the section. Figure 4.16 shows the variation of corrections \((x, \delta f^n(x)y/\beta)\) at the first iteration and after 14 iterations of the shape design process. These plots represent a measure of shape variation (first corrections) with respect to the problem. Note that \(\delta f^n\) is related with state and adjoint solutions. At the beginning of the process we can see (plot on the left) that the corner is the most sensible zone of the domain, after 14 iterations of the process (plot on the right) the shape variation is reduced (\(\text{max} \delta f^{(n)}\) is 10% of the one at the first step). In Figure 4.17 we report total vorticity \(([m^2/s])\) in down-field zone during the control process at different time steps. We underline that only for the curve at \(t = 1\) s we have a complete result dealing with vorticity reduction over a period \(T\), other curves represent partial results considering a fraction of \(T\) (optimization is carried out over a period \(T\) at each iteration).

The most important contribution to total vorticity is given by the flow behavior in the first part of the period considered. The total vorticity reduction is quite substantial. Figure 4.17 shows the total vorticity trend in time. At the end of the process we can see that the vorticity behavior in time has the same trend, however its value is reduced.

Figure 4.8: Unsteady (pulsatile) Stokes velocity profiles at the inflow \([m/s]\).
Figure 4.9: Horizontal velocity $[ms^{-1} \cdot 10^{-2}]$ at $t = 0.1s$ for initial test configuration (left) and after 5 iterations (right).

Figure 4.10: Horizontal velocity $[ms^{-1} \cdot 10^{-2}]$ after $t = 0.1s$ at 11 (left) and 17 (right) iterations.

Figure 4.11: Horizontal velocity $[ms^{-1} \cdot 10^{-2}]$ at $t = 0.4s$ for initial test configuration (left) and after 5 iterations (right).
Figure 4.12: Horizontal velocity \([m s^{-1} \cdot 10^{-2}]\) after \(t = 0.4s\) at 11 (left) and 17 (right) iterations.

Figure 4.13: Horizontal velocity \([m s^{-1} \cdot 10^{-2}]\) at \(t = 0.7s\) for initial test configuration (left) and after 5 iterations (right).

Figure 4.14: Horizontal velocity \([m s^{-1} \cdot 10^{-2}]\) after \(t = 0.7s\) at 11 (left) and 17 (right) iterations.
Figure 4.15: Distributed vorticity \([s^{-1} \cdot 10^{-2}]\) in original configuration (left) and at the end of the optimization process (right). Vorticity in the upper corner and in the bed of the artery is diminished.

Figure 4.16: Variation of corrections \((x, \delta f^n(x)g/\beta)\) at the first iteration (left) and after 14 iterations of the process (right) \([m \cdot 10^{-2}]\).

Figure 4.17: Total vorticity \([m^2 s^{-1}]\) reduction at different time steps during shape optimization.
4.9 Conclusions about Part I

In the first part of this thesis we have proposed several mathematical approaches based on optimal flow control techniques for shape optimization problem. They both use simplified fluid models to be verified in feedback and are based on local boundary variations and small perturbation theory. Using a vorticity cost functional we tried to unify different approaches about bypass configurations already available in literature in which cost functionals depend on velocity field, viscous dissipation and stresses: all these quantities can indeed be related to vorticity. In the last approach we have also modelled the unsteady flow rate in the coronary after bypass implanting procedures.

The approach proposed and discussed in this chapter and in the previous one, based on small perturbation in fluid mechanics and set into a reference domain, is more general than the one discussed in Chapter 2 and has permitted us to investigate the problem from a theoretical point of view and to generalize the approach dealing with transient flows and to use a unique computational mesh.

The results we have analyzed in this part of the work underline the importance of an accurate modelling of the incoming branch of the bypass, in particular with cuffed and/or patched shapes, which provide an enlarged section between the bypass and the host artery. Some of the results found using shape optimization and optimal flow control were already known in literature, but not validated yet by optimization techniques and not interpreted and unified using vorticity cost functional.

What ought be done to generalize and validate our results on the local shape of the incoming branch of the bypass is to study a complete bridge configuration depending on a certain number of parameters. This aspect will be considered in the second part of the thesis by reduced basis techniques where we will deal both with Stokes and Navier-Stokes equations.
Another issue is how to reduce computational costs in solving optimal control problems, like the ones proposed in Chapters 2–4, which are very expensive: reduced basis methods can be adopted also in this respect and, in the last part, it will be shown how to use reduced basis to solve optimal control problems.
Part II

Reduced Basis Techniques for Optimization
Chapter 5

Reduced Basis Methods for Stokes Equations

In this chapter we extend reduced basis techniques\(^3\) to Stokes equations in domains with affine (or approximately affine) parameter dependence to have a rapid and reliable prediction of linear-functional outputs. Particular attention is given to the approximation problem: (i) to the pressure treatment of incompressible Stokes problem; (ii) to find an equivalent inf-sup condition to guarantee stability of reduced basis solutions enriching the reduced basis velocity approximation space with the solutions of a supremizer problem; (iii) to provide algebraic stability of the problem by bounding the condition number of reduced basis matrices using an orthonormalization procedure applied to functions used as basis; (iv) to improve accuracy of outputs estimation by a dual residual method based on adjoint problem approach; (v) to reduce computational costs in order to have real time solution of parametrized problem; (vi) finally, by considering more complex (“non-compliant”) outputs.

The motivation in developing reduced basis techniques is the set-up of computational methods which allow huge computational savings to solve also the problems introduced in the first part of this thesis. In this chapter we will derive a methodology focusing on bypass optimization problem, however we will also address some issues and results which are interesting independently of the problem considered in the previous chapters.

5.1 Reduced basis for viscous flows

In this chapter we are going to introduce the reduced basis formulation for a pre-process optimization on a Cartesian geometry for a simplified bypass configuration. Preliminary applications of reduced basis techniques to incompressible viscous flows problems are given in Peterson [113], in Ito and Ravindran [64] and also [62] for optimal control problems; however in these papers the pressure approximation is not considered and stability of solutions is not discussed. More recent works dealing with free divergence velocity spaces and physical parameters and focused on a posteriori error estimates are the ones by Patera, Veroy et

\(^3\)For an introduction on reduced basis methods and related aspects see Appendix A.
This chapter (which generalizes the results of [111]) is organized as follows: in Section 5.2 we formulate the problem for parametrized Stokes equations. In Section 5.3 we introduce the Stokes reduced basis formulation and analyze the stability of the approximation. In Section 5.4 we study a procedure to control $N$ (the number of reduced basis functions) more tightly and to apply an adaptive procedure for the basis construction using a suitable error projection. In Section 5.5 we analyze the inf-sup condition which is necessary for the stability of reduced basis approximation. Then in the Section 5.6 we present some numerical results based on three different test cases. In Section 5.7 we deal with algebraic stability problem and we present possible solutions to achieve it by an orthonormalization procedure. In Section 5.8 we introduce some sensitivity results based on geometrical quantities of interest for the bypass configuration. In Section 5.9 we start dealing with more complex outputs related with fluid mechanics quantities and based on a dual residual approach.

A recent interesting approach to parametrized domains is provided by Maday and Ronquist [88], where a reduced basis element method based on a domain decomposition approach is proposed. The idea is to decompose the computational domain into a series of subdomains that are deformations of a few reference domains into geometrically similar parts ("elements"). Associated with each reference subdomain there are precomputed solutions corresponding to the same governing partial differential equation (the case deals with steady Stokes equations), but solved for different choices of parameters deforming the subdomains and mapped onto the reference shape. In our case we do not consider a domain decomposition approach and we deform the complete parametrized domain. Our approach is an extension of the one used to introduce the parametrization of physical quantities by affine (and then non-affine) maps.

### 5.2 The parametrized Stokes problem

We start considering parametrized Stokes equations in the same "T" domain introduced in Appendix A with one more parameter, $H$, the bypass bridge height as in Figure 5.1. Referring to the following steady Stokes problem in a domain $\Omega$

\[
\begin{align*}
\begin{cases}
-\nu \Delta \hat{u} + \nabla \hat{p} &= \hat{f} \text{ in } \hat{\Omega}, \\
\nabla \cdot \hat{u} &= 0 \text{ in } \hat{\Omega}, \\
\hat{u} &= 0 \text{ on } \Gamma_w; \quad \hat{u} = \hat{g}_m \text{ on } \Gamma_{in}, \\

\end{cases}
\end{align*}
\]

its weak formulation reads: find $\hat{u} \in Y = H^1_D(\Omega) \times H^1_D(\Omega)$, $\hat{p} \in Q = L^2(\Omega)$, $\hat{\Omega} \subset \mathbb{R}^2$ such that:

\[
\begin{align*}
\begin{cases}
\nu \int_{\hat{\Omega}} \nabla \hat{u} \cdot \nabla \hat{w} d\Omega - \int_{\hat{\Omega}} \hat{p} \nabla \cdot \hat{w} d\Omega = \int_{\hat{\Omega}} \hat{f} \cdot \hat{w} d\Omega + \langle \hat{F}^0, \hat{w} \rangle \forall \hat{w} \in Y, \\
\end{cases}
\end{align*}
\]

\[
\begin{align*}
\int_{\hat{\Omega}} \hat{q} \nabla \cdot \hat{u} d\Omega = \langle \hat{G}^0, \hat{q} \rangle \forall \hat{q} \in Q,
\end{align*}
\]

$\hat{F}^0, \hat{G}^0$ are terms due to non-homogeneous Dirichlet boundary condition ($\hat{u} = \hat{g}_m$) on $\Gamma_{in}$, $\hat{\Gamma}_D = \Gamma_{in} \cup \Gamma_w$, on $\Gamma_{out}$ we have a free stress Neumann condition. In our case the true
domain is made up of $R = 4$ subdomains: $\hat{\Omega} = \bigcup_{r=1}^{R} \hat{\Omega}^r$, so that the bilinear and linear forms of the problem weak statement read for $1 \leq i, j \leq d = 2$ and $\nu_{i,j} = \nu \delta_{i,j}$ as:

\[
\langle \hat{A} \hat{u}, \hat{w} \rangle = \sum_{r=1}^{R} \int_{\hat{\Omega}^r} \frac{\partial \hat{u}}{\partial x_i} \frac{\partial \hat{w}}{\partial x_j} d\hat{\Omega}, \tag{5.2.3}
\]

\[
\langle \hat{B} \hat{p}, \hat{w} \rangle = -\sum_{r=1}^{R} \int_{\hat{\Omega}^r} \hat{p} \nabla \cdot \hat{w} d\hat{\Omega}, \tag{5.2.4}
\]

\[
\langle \hat{F}, \hat{w} \rangle = \langle \hat{F}_s, \hat{w} \rangle + \langle \hat{F}^0, \hat{w} \rangle, \tag{5.2.5}
\]

and

\[
\langle \hat{F}_s, \hat{w} \rangle = \sum_{r=1}^{R} \int_{\Omega^r} \hat{f} \hat{w} d\hat{\Omega}, \langle \hat{F}^0, \hat{w} \rangle = -\langle \hat{A} \hat{g}_m, \hat{w} \rangle, \langle \hat{G}^0, \hat{q} \rangle = \langle \hat{B} \hat{q}, \hat{g}_m \rangle. \tag{5.2.6}
\]

Then we write:

\[
\left\{ \begin{array}{l}
\langle \hat{A} \hat{u}, \hat{w} \rangle + \langle \hat{B} \hat{p}, \hat{w} \rangle = \langle \hat{F}, \hat{w} \rangle \quad \forall \hat{w} \in Y, \\
-\langle \hat{B} \hat{q}, \hat{u} \rangle = \langle \hat{G}^0, \hat{q} \rangle \quad \forall \hat{q} \in Q. \end{array} \right. \tag{5.2.7}
\]

We want to build a system of $P^2 DEs$ (Parametrized Partial Differential Equations) depending on a set of geometrical coefficients $\mu$, that we call parameters. Then (5.2.7) is traced back to a reference domain (Fig.5.2) by an affine mapping of the subdomains $\hat{\Omega}^r$ into $\Omega^r$. For any $\hat{x} \in \hat{\Omega}^r$, $r = 1, \ldots, R$, its image $x \in \Omega^r$ is given by

\[
x = G^r(\mu; \hat{x}) = G^r(\mu)\hat{x} + g^r, \quad 1 \leq r \leq R; \tag{5.2.8}
\]
we thus write
\[ \frac{\partial}{\partial x_i} = \frac{\partial x_j}{\partial \bar{x_i}} \frac{\partial}{\partial x_j} = G_{ji}^r(\mu) \frac{\partial}{\partial x_j} \]  
(5.2.9)
and we get in the reference domain \( \Omega \):
\[ \langle A\mathbf{u}, \mathbf{w} \rangle = \sum_{r=1}^{R} \int_{\Omega^r} \frac{\partial \mathbf{u}}{\partial x_i} \left( G_{ir}^r(\mu) \nu_{ij}^rG_{jr}^r(\mu) \text{det}(G^r(\mu))^{-1} \right) \frac{\partial \mathbf{w}}{\partial x_j} \, d\Omega \quad \forall \mathbf{w} \in Y, \]  
(5.2.10)
\[ \langle Bp, \mathbf{w} \rangle = -\sum_{r=1}^{R} \int_{\Omega^r} p \left( G_{ij}^r(\mu) \text{det}(G^r(\mu))^{-1} \right) \frac{\partial \mathbf{w}}{\partial x_i} \, d\Omega \quad \forall \mathbf{w} \in Y, \]  
(5.2.11)
\[ \langle F, \mathbf{w} \rangle = \langle F_s, \mathbf{w} \rangle + \langle F^0, \mathbf{w} \rangle, \]  
(5.2.12)
where
\[ \langle F_s, \mathbf{w} \rangle = \sum_{r=1}^{R} \int_{\Omega^r} \left( \hat{f}^r \text{det}(G^r(\mu))^{-1} \right) \mathbf{w} \, d\Omega; \quad \langle F^0, \mathbf{w} \rangle = -\langle A_{\text{gin}}, \mathbf{w} \rangle; \quad \langle G^0, q \rangle = \langle Bq, g_{\text{in}} \rangle. \]  
(5.2.13)
We introduce a vector of parameters \( \mathbf{\mu} = \{ t, D, L, H, S, \theta \} \in \mathcal{D}^\mu \subset \mathbb{R}^P, \mathcal{D}^\mu \) is given by:
\[ [t_{\text{min}}, t_{\text{max}}] \times [D_{\text{min}}, D_{\text{max}}] \times [L_{\text{min}}, L_{\text{max}}] \times [S_{\text{min}}, S_{\text{max}}] \times [H_{\text{min}}, H_{\text{max}}] \times [\theta_{\text{min}}, \theta_{\text{max}}]. \]
The transformation tensors for bilinear viscous terms are defined as follows:
\[ \nu_{ij}^r(\mu) = G_{ir}^r(\mu) \nu_{ij}^r G_{jr}^r(\mu) \text{det}(G^r(\mu))^{-1}, 1 \leq i, j \leq 2, r = 1, ..., R, \]
then in our case:
\[ \nu^1 = \nu \begin{bmatrix} \frac{4}{P} & -\tan \theta \\ -\tan \theta & \frac{1}{1+\tan^2 \theta} \end{bmatrix}; \quad \nu^2 = \nu \begin{bmatrix} \frac{S}{P} & 0 \\ 0 & \frac{D}{S} \end{bmatrix}; \]  
(5.2.14)
\(\nu^3 = \nu \begin{bmatrix} \nabla & 0 \\ 0 & \frac{D}{T} \end{bmatrix} ; \nu^4 = \nu \begin{bmatrix} \nabla & 0 \\ 0 & \frac{D}{T} \end{bmatrix}. \) (5.2.15)

The tensors for pressure and divergence forms are:

\[ \chi_{ij}^r(\mu) = G_{ij}^r \text{det}(G^r(\mu))^{-1} \]

and are given by:

\[ \chi^1 = \begin{bmatrix} t & -H \tan \theta \\ 0 & H \end{bmatrix} ; \chi^2 = \begin{bmatrix} S & 0 \\ 0 & D \end{bmatrix} ; \]

\[ \chi^3 = \begin{bmatrix} t & 0 \\ 0 & D \end{bmatrix} ; \chi^4 = \begin{bmatrix} L & 0 \\ 0 & D \end{bmatrix}. \] (5.2.16)

Furthermore, we may define

\[ \Theta^{q(i,j,r)}(\mu) = \nu_{ij}^r(\mu), \quad \langle A^{q(i,j,r)}u, w \rangle = \int_{\Omega^r} \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_j} d\Omega, \]

\[ \Phi^{s(i,j,r)}(\mu) = \chi_{ij}^r(\mu), \quad \langle B^{s(i,j,r)}p, w \rangle = -\int_{\Omega^r} p \frac{\partial w}{\partial x_j} d\Omega, \] (5.2.18)

for \(1 \leq r \leq R, 1 \leq i, j \leq d = 2 \) (\(q\) and \(s\) are condensed indexes of \(i, j, r\) quantities) and we apply affine decomposition:

\[ A(\Theta(\mu), u, w) = \sum_{q=1}^{Q^a} \Theta^q(\mu)A(u, w)^q ; \]

\[ B(\Phi(\mu), p, w) = \sum_{s=1}^{Q^b} \Phi^s(\mu)B(p, w)^s ; \]

in our case \(Q^a = 20\) and \(Q^b = 9\), in reality: \(\max(Q^a) = d \times d \times d \times R = 32\), \(\max(Q^b) = d \times d \times R = 16\);

The Stokes problem rewritten on the reference domain \(\Omega\) reads: find \((u(\mu), p(\mu)) \in Y \times Q\) such that:

\[ \begin{cases} A(\mu; u(\mu), w) + B(\mu; p(\mu), w) = \langle F, w \rangle \forall w \in Y, \\ B(\mu; q, u(\mu)) = \langle G^0, q \rangle \forall q \in Q. \end{cases} \] (5.2.20)

A necessary condition for the well posedness of this problem is the so-called \textit{inf-sup} condition (LBB) [127]:

\[ \exists \beta_0 > 0 : \beta(\mu) = \inf_{q \in Q} \sup_{w \in Y} \frac{B(\mu, q, w)}{\|w\|_Y \|q\|_Q} \geq \beta_0, \forall \mu \in \mathcal{D}^\mu, \] (5.2.21)

where \(Y = H^1_0 \times H^1_0\). To verify it let us introduce a supremizer operator \(T^\mu: Q \to Y\) defined as follows:

\[ (T^\mu q, w)_Y = B(\mu; q, w), \forall w \in Y. \] (5.2.22)

It is readily shown that:

\[ T^\mu q = \arg \sup_{w \in Y} \frac{B(\mu; q, w)}{\|w\|_Y}. \] (5.2.23)
then
\[ \beta^2(\mu) = \inf_{q \in Q} \frac{(T^\mu q, T^\mu q)_Y}{\|q\|_Q^2}, \]
by definition (5.2.22):
\[ \|T^\mu q\|_Y^2 = B(\mu; q, T^\mu q), \]
and by applying (5.2.23) and (5.2.21)
\[ \inf_{q \in Q} \frac{(T^\mu q, T^\mu q)_Y}{\|q\|_Q^2} = \inf_{q \in Q} \frac{\|T^\mu q\|_Y^2}{\|q\|_Q^2} = \]
this proofs (5.2.24). In our numerical approximation the Stokes problem has been solved by Galerkin-Finite Element Method using Taylor-Hood \( P_2 - P_1 \) elements for velocity and pressure, respectively. See Girault and Raviart [43], Gresho and Sani [46], and Gunzburger [48].

### 5.3 The reduced basis formulation of the Stokes equations

In the reduced basis approximation we take some \( \{\mu^n\} \) samples \( S_N^\mu = \{\mu^1, \ldots, \mu^N\} \), where \( \mu^n \in D^\mu, n = 1, \ldots, N \).

The **reduced basis pressure space** is \( Q_N = \text{span} \{\xi_n, n = 1, \ldots, N\} \), where \( \xi_n = p(\mu^n) \).

We can build the **reduced basis velocity space** as \( Y_N^\mu = \text{span} \{\zeta_n, n = 1, \ldots, N; T^\mu \xi_n, n = 1, \ldots, N\} \), where \( \zeta_n = u(\mu^n) \). The reduced basis approximation problem reads:

find \( (u_N(\mu), p_N(\mu)) \in Y_N^\mu \times Q_N \) s.t.:

\[
\begin{aligned}
A(\mu; u_N(\mu), w) + B(\mu; p_N(\mu), w) &= \langle F, w \rangle \quad \forall \ w \in Y_N^\mu, \\
B(\mu; q, u_N(\mu)) &= \langle G, q \rangle \quad \forall \ q \in Q_N.
\end{aligned}
\]

Problem (5.3.1) is subject to an equivalent reduced basis inf-sup condition.

**Lemma 5.3.1** We define
\[ \beta_N(\mu) = \inf_{q \in Q_N} \sup_{w \in Y_N^\mu} \frac{B(\mu, q, w)}{\|w\|_Y \|q\|_Q}. \]
Then we can prove that
\[ \beta_N(\mu) \geq \beta(\mu) \geq \beta_0 > 0, \forall \mu \in D^\mu. \]

**Proof:**
\[
\beta(\mu) = \inf_{q \in Q} \sup_{w \in Y} \frac{B(\mu, q, w)}{\|w\|_Y \|q\|_Q} \leq \inf_{q \in Q_N} \sup_{w \in Y_N^\mu} \frac{B(\mu, q, w)}{\|w\|_Y \|q\|_Q} = \inf_{q \in Q_N} \frac{\|T^\mu q\|_Y \|q\|_Q}{\|T^\mu q\|_Y \|q\|_Q} \leq \]
\[
\leq \inf_{q \in Q_N} \sup_{w \in Y_N^\mu} \frac{B(\mu, q, w)}{\|w\|_Y \|q\|_Q} = \beta_N(\mu). \quad \square
\]
To demonstrate the Lemma above we have applied the fact that $Q_N \subset Q$, the definition of the supremizer and the fact that the velocity space is enriched by supremizers, respectively. For further elements dealing with supremizer operator and the reduced basis framework see Rovas [134] for the general non-coercive problem and Patera et al. [161] especially for Helmholtz and Burgers equations.

We rewrite for computationally convenience $Y^\mu_N$ using the affine dependence of $B(\mu, q, w)$ on the parameter and the linearity of $T^\mu$: $T^\mu \xi = \sum_{q=1}^{Q^b} \Phi^q(\mu)T^q \xi$ for any $\xi$ and $\mu$, which allows us to write:

$$Y^\mu_N = \text{span} \{ \sum_{k=1}^{Q^b} \Phi^k(\mu)\sigma_{kn}, \ n = 1, \ldots, 2N \},$$

where $\overline{Q^b} = Q^b + 1$, $\Phi\overline{Q^b} = 1$. For $n = 1, \ldots, N$:

$$\sigma_{kn} = 0, \text{ for } k = 1, \ldots, Q^b; \sigma_{\overline{Q^b}n} = \zeta_n = u(\mu^n).$$

For $n = N + 1, \ldots, 2N$:

$$(\sigma_{kn}, w)_Y = B(\xi_{n-N}, w)^k, \forall w \in Y, \text{ for } k = 1, \ldots, Q^b; \sigma_{\overline{Q^b}n} = 0.$$  \hspace{1cm} (5.3.2)

For a new “$\mu$” we want a solution given by a combination of previously computed stored solutions as basis functions:

$$u_N(\mu) = \sum_{j=1}^{2N} u_{Nj}(\mu) \left( \sum_{k=1}^{Q^b} \Phi^k(\mu)\sigma_{kj} \right),$$

$$p_N(\mu) = \sum_{l=1}^{N} p_{Nl}(\mu)\xi_l,$$

whose weights $u_{Nj}$ and $p_N$ are given by the following reduced basis linear system:

$$\begin{cases} \sum_{j=1}^{2N} A^\mu_{ij} u_{Nj}(\mu) + \sum_{l=1}^{N} B^\mu_{ij} p_{Nl}(\mu) = F^\mu_i, \ 1 \leq i \leq 2N, \\ \sum_{j=1}^{2N} B^\mu_{lj} u_{Nj}(\mu) = G^\mu_l, \ 1 \leq l \leq N, \end{cases} \hspace{1cm} (5.3.3)$$

where the sub-matrices $A$ and $B$ are given by:

$$A^\mu_{ij} = \sum_{k=1}^{Q^b} \sum_{k'=1}^{Q^b} \sum_{k''=1}^{Q^b} \Phi^k(\mu)\Phi^{k'}(\mu)\Phi^{k''}(\mu)A(\sigma_{k'i}, \sigma_{k''j})^k, \ 1 \leq i, j \leq 2N,$$

$$B^\mu_{il} = \sum_{k=1}^{Q^b} \sum_{k'=1}^{Q^b} \Phi^k(\mu)\Phi^{k'}(\mu)B(\sigma_{k'i}, \xi_l)^k, \ 1 \leq i \leq 2N, \ 1 \leq l \leq N,$$

and:

$$F^\mu_i = \sum_{k'=1}^{Q^b} \Phi^{k'}(\mu)\langle F, \sigma_{k'i} \rangle, \ 1 \leq i \leq 2N; \ G^\mu_l = \langle G^0, \xi_l \rangle, \ 1 \leq l \leq N.$$
Finally problem (5.7.10) can be written in compact form as

\[
\begin{pmatrix}
  A & B \\
  B^T & 0
\end{pmatrix}
\begin{pmatrix}
  u_N \\
  p_N
\end{pmatrix} =
\begin{pmatrix}
  F \\
  G
\end{pmatrix}.
\]

(5.3.4)

This linear system whose unknowns are the coefficients of the linear combination of previously computed off-line solutions has the same structure of a finite element Stokes problem. Using reduced basis we deals with matrix of smaller dimension (of order of \( N \)) and with full matrices (instead of sparse ones).

5.3.1 Outputs of interest

As outputs of interest we may consider, for example, the mean value of the velocity components:

\[
s_1(\mu) = \frac{\sum_{r=1}^{R} \int_{\Omega_r} u_1 d\Omega}{\sum_{r=1}^{R} \int_{\Gamma_r} d\Omega}; \quad s_2(\mu) = \frac{\sum_{r=1}^{R} \int_{\Omega_r} u_2 d\Omega}{\sum_{r=1}^{R} \int_{\Gamma_r} d\Omega}
\]

(5.3.5)

other outputs are derived from velocity such as vorticity:

\[
s_\nu(\mu) = \sum_{r=1}^{R} \int_{\Omega_r} \left( \frac{\partial u_2(\mu)}{\partial x_1} - \frac{\partial u_1(\mu)}{\partial x_2} \right) d\Omega,
\]

(5.3.6)

where we have set \( u = (u_1, u_2)^T \) and \( x = (x_1, x_2)^T \), or wall shear stress:

\[
s_\tau(\mu) = \int_{\Gamma_w} \nu \frac{\partial u(\mu)}{\partial n} \cdot \hat{t} d\Gamma,
\]

(5.3.7)

where \( \hat{t} \) is the tangential unit vector. In the last two examples we may introduce a dual residual correction based on an adjoint problem to improve output accuracy. See Section 5.9.

5.3.2 Reduced basis on-line complexity

We have the following computational costs to build (on-line) reduced basis matrix, given also the supremizer component in the velocity space: \( O(Q^a(Q^b)^2 4N^2) \) for sub-matrix \( A \), \( O((Q^b)^2 2N^2) \) for \( B \), \( O((Q^b) 2N) \) for \( F \) and \( O(9N^3) \) for the inversion of the full reduced basis matrix (5.3.4). Note that the approach presented to build reduced basis velocity space is only one of the possible solutions and the space \( Y_N^\mu \), in this case, is depending on the value of the on-line \( \mu \) parameter. Other options that avoid ill conditioning problems characterized by different computational costs are available and will be presented in Section 5.7.

5.4 Off-line optimized basis assembling: adaptive procedure

In this case an adaptive procedure based on \( H^1 \) max relative error projection for velocity \( E_{H^1} \), has been developed, optionally we can also considering and combining \( L^2 \) max relative error projection for pressure \( E_{L^2} \). We underline that, given the higher powers of \( N \) that appear
in our cost computing estimation, it is crucial (both as regards online and offline effort) to control N more tightly. To this end, we may gainfully apply our off line assembling procedure adaptively, see also Veroy et al. [161] and Prud’homme [121]. We first construct, offline, an approximation that, over most of the domain, exhibits an error \( \epsilon (E_{H^1} \text{ or } E_{L^2}) \) less than \( \epsilon_{d}^{\text{prior}} \): we begin with a first point \( \mu^1 \) \((S_{N'=1} = \{ \mu^1 \})\); we next (inexpensively) evaluate error \( \epsilon_{N'=1}(\mu) \) over a large test sample of parameter points in \( \mathcal{D}^\mu \), denoted with \( \Sigma_{\text{prior}} \); we then choose for \( \mu^2 \) (and hence \( S_{N'=2} = \{ \mu^1, \mu^2 \} \)) the maximizer of \( \epsilon_{N'=1}(\mu) \) over \( \Sigma_{\text{prior}} \). We repeat this process until the maximum of \( \epsilon_{N'=N_{\text{prior}}}(\mu) \) over \( \Sigma_{\text{prior}} \) is less than \( \epsilon_{d}^{\text{prior}} \). Then, online, given a new value of the parameter, \( \mu \), and an error tolerance \( \epsilon_{d}^{\text{post}}(\mu) \), we essentially repeat this adaptive process - but now our sample points are drawn from \( S_{N_{\text{prior}}} \), and the test sample is a singleton - \( \mu \). Typically we choose \( \epsilon_{d}^{\text{prior}} \ll \epsilon_{d}^{\text{post}}(\mu) \) since our test is not exhaustive; and therefore, typically, \( N_{\text{post}}(\mu) \ll N_{\text{prior}} \). With the adaptive process we get higher accuracy at lower \( N \): modest reductions in \( N \) can translate into measurable performance improvements. This procedures is very important not only to get a computationally cheaper and faster procedure but also to avoid ill-conditioning in matrix assembling procedures. Error projection procedure, described below, has permitted us to have an off-line adaptive (and optimized) assembling procedure which is very fast and inexpensive, without reduced basis matrix assembling procedures, but only the solution of a linear system.

### 5.4.1 \( L^2 \) pressure error projection

Given a new \( \mu \) and hence a new approximated \( p_h(\mu) \) pressure solution we solve the following linear system:

\[
\sum_{j=1}^{N} \int_{\Omega} \xi_i \xi_j c_j = \int_{\Omega} \xi_i p_h(\mu), \quad 1 \leq i \leq N, \tag{5.4.1}
\]

where \( \xi_i \) are \( N \) pressure solutions used as basis given by: \( \xi_i = \sum_{k=1}^{N_P} \eta_{ik} \Phi_k \), computed solving a finite element problem on \( N_P \) pressure nodes of the mesh; \( \Phi_k \) are finite element local functions for the pressure. Expanding (5.4.1) by finite element approximation we can write:

\[
\sum_{j=1}^{N} \sum_{l=1}^{N_p} \sum_{k=1}^{N_P} p_l(\int_{\Omega} \Phi_l \Phi_k) p_{jk} c_j = \sum_{m=1}^{N_P} \sum_{n=1}^{N_P} p_m(\int_{\Omega} \Phi_m \Phi_n) p_{h_n}(\mu), \quad 1 \leq i \leq N,
\]

\[
(p^T M^p p_j) c = \mu^T M^p p_h(\mu),
\]

\[
\bar{M}^p c = E^p.
\]

Once we have \( c \) we write the \( L^2 \) pressure error \( E_{L^2} \) as:

\[
E_{L^2}^2 = \| p_h(\mu) \|^2_{L^2} - \sum_{i=1}^{N} \sum_{j=1}^{N} c_i (\int_{\Omega} \xi_i \xi_j) c_j,
\]

where \( (\int_{\Omega} \xi_i \xi_j) \) is the pressure mass matrix \( \bar{M}^p_{ij} = \sum_{l=1}^{N_P} \sum_{k=1}^{N_P} p_l(\int_{\Omega} \Phi_l \Phi_k) p_{jk}, \quad 1 \leq i, j \leq N \), i.e. \( p^T M^p p_j \).
5.4.2 \( H^1 \) velocity error projection

Given a new \( \mu \) and hence a new approximated velocity solution \( u_h(\mu) \) we solve the following linear system

\[
\sum_{j=1}^{2N} \left( \int_{\Omega} \zeta_i \zeta_j + \int_{\Omega} \nabla \zeta_i \cdot \nabla \zeta_j \right) c_j = \int_{\Omega} \zeta_i u_h(\mu) + \int_{\Omega} \nabla \zeta_i \cdot \nabla u_h(\mu), \quad 1 \leq i \leq 2N,
\]

(5.4.2)

where \( \zeta_i \) are 2\( N \) velocity solutions used as basis, made up of finite element approximated velocities and supremizer solutions, written as: \( \zeta_i = \sum_{k=1}^{N} u_k \Psi_k \) on \( N \) velocity nodes by \( \Psi_k \) finite element velocity local functions. Expanding the system (5.4.2) by finite element approximation (on \( N \) nodes) we get, for \( 1 \leq i \leq 2N \):

\[
\sum_{j=1}^{2N} \sum_{l=1}^{N} \sum_{k=1}^{N} u_i \left( \int_{\Omega} \Psi_l \Psi_k u_j k \right) + \sum_{j=1}^{2N} \sum_{l=1}^{N} \sum_{k=1}^{N} u_i \left( \int_{\Omega} \nabla \Psi_l \cdot \nabla \Psi_k u_j k \right) c_j = \sum_{m=1}^{N} \sum_{n=1}^{N} u_{im} \left( \int_{\Omega} \Psi_m \Psi_n u_h n(k) \right) + \sum_{m=1}^{N} \sum_{n=1}^{N} u_{im} \left( \int_{\Omega} \nabla \Psi_m \cdot \nabla \Psi_n u_h n(k), \right)
\]

\[
(c_i^T (M + K) c_j) = u_i^T (M + K) u_j,
\]

Once we have \( c \) we write the \( H^1 \) velocity error \( E_{H^1} \) as:

\[
E_{H^1}^2 = \|u_h(\mu)\|_{H^1}^2 - \sum_{i=1}^{2N} \sum_{j=1}^{2N} c_i \left( \int_{\Omega} \zeta_i \zeta_j + \nabla \zeta_i \cdot \nabla \zeta_j \right) c_j,
\]

where \( \left( \int_{\Omega} \zeta_i \zeta_j \right) \) is the velocity mass matrix \( \sum_{l=1}^{N} \sum_{k=1}^{N} u_i \left( \int_{\Omega} \Psi_l \Psi_k \right) u_j k \), i.e. \( u_i^T M u_j \), and \( \left( \int_{\Omega} \nabla \zeta_i \cdot \nabla \zeta_j \right) = \sum_{l=1}^{N} \sum_{k=1}^{N} u_i \left( \int_{\Omega} \nabla \Psi_l \cdot \nabla \Psi_k \right) u_j k \) the velocity stiffness matrix, i.e. \( u_i^T K u_j \).

5.5 The computation of the constant \( \beta \) of the inf-sup condition

The calculation of the constant of the inf-sup condition for finite element (\( \beta \)) and reduced basis method (\( \beta_N \)) has been carried out as a test to guarantee approximation stability.

- In the case in which we want to calculate \( \beta \) we have a generalized eigenvalue problem on sparse matrices (see Gresho and Sani [46], Malkus [93], and Quarteroni and Valli [127]):

\[
(G^T K^{-1} G) x = \lambda M^P x
\]

(5.5.1)

\( K \) is the finite element (velocity) stiffness matrix \( \left( \int_{\Omega} \nabla \Psi_l \cdot \nabla \Psi_k \right) \), \( G \) is the velocity-divergence finite element matrix \( \left( \int_{\Omega} \Phi_l \nabla \cdot \Psi_k \right) \), \( M^P \) is the finite element pressure mass matrix \( \left( \int_{\Omega} \Phi_l \Phi_k \right) \), \( \lambda \) are eigenvalues and \( x \) eigenvectors. The \( \beta \) constant is given by

\[
\beta = \sqrt{\lambda_{min}}
\]

(5.5.2)
5.6. SOME PRELIMINARY NUMERICAL RESULTS

To calculate $\beta_N$ we have a generalized eigenvalues problem on reduced basis (full) matrices (previously defined):

$$(B^T A^{-1} B)z_N = S_N z_N = \lambda_N \nabla p z_N$$  \hspace{1cm} (5.5.3)

$$\beta_N = \sqrt{\lambda_{N_{\text{min}}}}$$  \hspace{1cm} (5.5.4)

We have tested that equivalent reduced basis inf-sup parameter condition: $\beta_N(\mu) \geq \beta(\mu) \geq \beta_0 > 0, \forall \mu \in D^e$ has been satisfied. Table 5.1 shows a test driven to calculate the constant $\beta$ of the inf-sup condition studying a one-parameter varying configuration (the diameter $D$, for example) at different $N$.

<table>
<thead>
<tr>
<th>$D$</th>
<th>$\beta$</th>
<th>$\beta_N, N = 1$</th>
<th>$\beta_N, N = 2$</th>
<th>$\beta_N, N = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.63</td>
<td>2.0249</td>
<td>3.9096</td>
<td>2.7895</td>
<td>2.531</td>
</tr>
<tr>
<td>0.68</td>
<td>2.0286</td>
<td>3.9316</td>
<td>2.8708</td>
<td>2.6625</td>
</tr>
<tr>
<td>0.73</td>
<td>2.0274</td>
<td>3.952</td>
<td>2.9456</td>
<td>2.7992</td>
</tr>
<tr>
<td>0.78</td>
<td>2.0257</td>
<td>3.971</td>
<td>3.0145</td>
<td>2.9386</td>
</tr>
<tr>
<td>0.83</td>
<td>2.0237</td>
<td>3.9888</td>
<td>3.0781</td>
<td>3.0729</td>
</tr>
<tr>
<td>0.88</td>
<td>2.0216</td>
<td>4.0054</td>
<td>3.137</td>
<td>3.0729</td>
</tr>
<tr>
<td>0.93</td>
<td>2.0192</td>
<td>4.0208</td>
<td>3.1918</td>
<td>3.186</td>
</tr>
<tr>
<td>0.98</td>
<td>2.0167</td>
<td>4.0353</td>
<td>3.243</td>
<td>3.264</td>
</tr>
<tr>
<td>1.03</td>
<td>2.0139</td>
<td>4.0487</td>
<td>3.291</td>
<td>3.3466</td>
</tr>
<tr>
<td>1.08</td>
<td>2.011</td>
<td>4.0612</td>
<td>3.3361</td>
<td>3.3759</td>
</tr>
</tbody>
</table>

Table 5.1: $\beta$ and $\beta_N$ in one parameter dependent configuration ($D$ for example). Other parameters are frozen.

5.6 Some preliminary numerical results

Several numerical tests were carried out to develop all the Stokes reduced basis toolbox and to study different related aspects: affine mapping transformation in reference domains, off-line optimization basis assembling procedure, approximation and algebraic stability (LBB inf-sup equivalent condition fulfillment and condition number control). Different meshes have been used (from coarse to fine). Using finer mesh CPU (Pentium IV, 2GHZ) operation time increases a lot during computing procedures such as assembling off-line adaptive calculations (i.e. reduced basis approximation spaces and matrix assembling, error estimation and, eventually, orthogonalization) and it is maintained reasonable during on-line calculation. A mesh adaptation procedure in some zones has been achieved. Taylor-Hood finite elements have been used: $\mathbb{P}^2-\mathbb{P}^1$ elements (for velocity and pressure respectively) [127]. Different solver for finite element and reduced basis systems have been used: both iterative (such as Bi-CGSTAB, Conjugate Gradient Method, GMRES) and direct, see [129]. The reduced basis solutions have been compared directly with true finite element solutions by computing $H^1$ relative error for velocity and $L^2$ relative error for pressure.
5.6.1 First Test: homogeneous Dirichlet boundary conditions, forced flow, 3(1) varying parameter(s)

The first test we introduce deals with a forced Stokes flow in the “T” domain with all zero Dirichlet boundary conditions (flow closed into a cavity), varying only three parameters available (depending by one adimensional quantity: \( \tau \)). Data and relationships for geometrical parameters used to solve the problem follow (see also Figure 5.1).

- (I) The bypass diameter is \( t = 3 - 2\tau \), the stenosis length \( S = \tau \) and the outflow length \( L = \tau \).
- (II) The viscosity is \( \nu = 0.04 \, m^2s^{-1} \), the force field is \( \mathbf{f} = (0, 10x)^T ms^{-2} \) in the true domain \( \Omega \).
- (III) The parameter \( \tau \) is ranging in \([0.1, 1.45]\). All other parameters are frozen: the bridge height \( H = 1 \), the arterial diameter \( D = 1 \) and the graft angle \( \theta = 0 \).

Figures 5.3 and 5.4 show examples of flow solution: velocity and pressure for two configuration with \( \tau_{min} \) and \( \tau_{max} \); Figure 5.5 shows the error reduction using the adaptive optimized procedure of Section 5.4 during basis assembling procedure, based on \( H^1 \) relative (projected) error minimization, on the left, and the parameters distribution during basis assembling procedure, on the right. Figure 5.6 shows true relative error reduction (\( H^1 \) for velocity and \( L^2 \) for pressure) considering a great number of different geometrical configurations. Tables 5.2-5.3 show error reduction (and its magnitude); Table 5.4 shows the proofs of approximation stability of reduced basis formulation computing \( \beta_N \), the LBB inf-sup equivalent constant and its comparison with \( \beta \) from Galerkin (FEM) approximation. Table 5.5 shows true error on a possible output of interest, for example \( s(\mu) = \int_{\Omega} \mathbf{f} \cdot \mathbf{u}(\mu)d\Omega \). Results on outputs are reported also in Figure 5.7.

![Figure 5.3: Velocity and pressure solution for \( \tau = 0.1 \).](image-url)
### §5.6. SOME PRELIMINARY NUMERICAL RESULTS

![Velocity and pressure solution for \( \tau = 1.45 \).](image)

**Figure 5.4:** Velocity and pressure solution for \( \tau = 1.45 \).

![Relative max projected \( H^1 \) error (see Section 5.4.2) on velocity at each iteration during adaptive basis assembling (left) and parameters distribution during off-line reduced basis assembling procedure (right).](image)

**Figure 5.5:** Relative max projected \( H^1 \) error (see Section 5.4.2) on velocity at each iteration during adaptive basis assembling (left) and parameters distribution during off-line reduced basis assembling procedure (right).

<table>
<thead>
<tr>
<th>( N )</th>
<th>Rel. Err. ( H^1 ) Max</th>
<th>Rel. Err. ( H^1 ) mean</th>
<th>Rel. Err. ( L^2 ) Max</th>
<th>Rel. Err. ( L^2 ) mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.8900e-001</td>
<td>5.9138e-001</td>
<td>2.5790e-001</td>
<td>1.0878e-001</td>
</tr>
<tr>
<td>2</td>
<td>2.1583e-001</td>
<td>7.8312e-002</td>
<td>2.0387e-002</td>
<td>1.1826e-002</td>
</tr>
<tr>
<td>3</td>
<td>2.3301e-002</td>
<td>1.0174e-002</td>
<td>5.9742e-003</td>
<td>1.4583e-003</td>
</tr>
<tr>
<td>4</td>
<td>3.7543e-003</td>
<td>8.7714e-004</td>
<td>4.4999e-004</td>
<td>5.5973e-005</td>
</tr>
<tr>
<td>5</td>
<td>3.2670e-003</td>
<td>5.0383e-004</td>
<td>3.2752e-004</td>
<td>3.3635e-005</td>
</tr>
<tr>
<td>6</td>
<td>1.4271e-004</td>
<td>3.6446e-005</td>
<td>2.8591e-005</td>
<td>1.7688e-006</td>
</tr>
<tr>
<td>7</td>
<td>8.4314e-005</td>
<td>1.8559e-005</td>
<td>2.7009e-005</td>
<td>1.5313e-006</td>
</tr>
</tbody>
</table>

**Table 5.2:** Table of \( H^1 \) and \( L^2 \) relative errors on velocity and pressure, respectively, 50 test configurations, 3(1) parameters, \( N < 7 \).
Figure 5.6: Left: relative true $H^1$ velocity errors, max and mean error over a large test sampling (50 configurations). Right: relative true $L^2$ pressure errors, max and mean error over a large test sampling (50 configurations); errors in log-log scale.

Table 5.3: Table of $H^1$ and $L^2$ relative errors on velocity and pressure, respectively, 50 test configurations, 3(1) parameters, $N > 7$.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>$4.3142e-005$</td>
<td>$7.4860e-006$</td>
<td>$7.0025e-007$</td>
<td>$1.6084e-007$</td>
</tr>
<tr>
<td>9</td>
<td>$2.4504e-006$</td>
<td>$5.2102e-007$</td>
<td>$2.2660e-007$</td>
<td>$4.1354e-008$</td>
</tr>
<tr>
<td>10</td>
<td>$1.6485e-006$</td>
<td>$2.5155e-007$</td>
<td>$1.3400e-007$</td>
<td>$2.8537e-008$</td>
</tr>
<tr>
<td>11</td>
<td>$1.8195e-007$</td>
<td>$1.7635e-008$</td>
<td>$8.8023e-008$</td>
<td>$4.5140e-009$</td>
</tr>
<tr>
<td>12</td>
<td>$5.3852e-008$</td>
<td>$6.9109e-009$</td>
<td>$8.6392e-010$</td>
<td>$1.5917e-010$</td>
</tr>
<tr>
<td>13</td>
<td>$2.4053e-008$</td>
<td>$2.9963e-009$</td>
<td>$1.5813e-010$</td>
<td>$4.1002e-011$</td>
</tr>
<tr>
<td>14</td>
<td>$1.2923e-009$</td>
<td>$2.5982e-010$</td>
<td>$1.0129e-010$</td>
<td>$1.4957e-011$</td>
</tr>
<tr>
<td>15</td>
<td>$1.0702e-009$</td>
<td>$1.6946e-010$</td>
<td>$8.7277e-011$</td>
<td>$1.1840e-011$</td>
</tr>
</tbody>
</table>

Table 5.4: $\beta_N$ equiv. LBB inf-sup constant. $\beta = 5.0164$ ($\tau = 0.76$).
Figure 5.7: Left: Functional output difference $\Delta s = (s - s_N); s = \int_\Omega f \cdot u$. Right: convergence of $s_N$ to $s$ (*) versus $N$ ($\tau = 0.356$).

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\Delta s$</th>
<th>$N$</th>
<th>$\Delta s$</th>
<th>$N$</th>
<th>$\Delta s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$9.4249e + 002$</td>
<td>6</td>
<td>$1.2142e - 002$</td>
<td>11</td>
<td>$2.4146e - 006$</td>
</tr>
<tr>
<td>2</td>
<td>$4.0905e + 000$</td>
<td>7</td>
<td>$9.6898e - 003$</td>
<td>12</td>
<td>$7.7663e - 007$</td>
</tr>
<tr>
<td>3</td>
<td>$7.0984e + 000$</td>
<td>8</td>
<td>$1.3454e - 003$</td>
<td>13</td>
<td>$8.4071e - 007$</td>
</tr>
<tr>
<td>4</td>
<td>$9.3965e - 001$</td>
<td>9</td>
<td>$4.4785e - 004$</td>
<td>14</td>
<td>$2.4822e - 007$</td>
</tr>
<tr>
<td>5</td>
<td>$9.5652e - 001$</td>
<td>10</td>
<td>$3.5133e - 004$</td>
<td>15</td>
<td>$2.0706e - 007$</td>
</tr>
</tbody>
</table>

Table 5.5: Output error: $\Delta s = s - s_N$ for $\tau = 0.356$. 
5.6.2 Second Test: mixed Dirichlet/Neuman boundary conditions, 5(2) varying parameters

The second test deals with a flow in the “T” shape bypass geometry with Neumann homogeneous boundary conditions on inflow $\Gamma_{in}$ and outflow $\Gamma_{out}$ and zero Dirichlet condition on the wall $\Gamma_w$. We deal with five varying parameters (bound by 3 relationships depending on two quantities $\tau$ and $\rho$). Data and values used follow.

- The bypass diameter is $t = 3 - 2\tau$, the stenosis length is $S = \tau$, the outflow length is $L = \tau$, the artery diameter is $D = \rho$, the bypass height is $H = 1 - \rho$.
- The viscosity is 0.04 $m^2 s^{-1}$, while the force field is $f = (0, 10)^T ms^{-2}$ in true domain $\Omega$.
- The parameters range is $\tau \in [0.1, 1.45]$ and $\rho \in [0.1, 1.9]$.

Figure 5.8 shows an example of flow solution (velocity) for a certain parameters combination; Figure 5.9 shows on the left the off-line parameters distribution during basis assembling procedures based on the reduction of the $H^1$ relative projected error using the adaptive procedure; on the right we have max and mean output error $\Delta s$ over a large number of different configurations, considering, for example, $s(\mu) = \int_{\Omega} f \cdot u(\mu) d\Omega$. Figure 5.10 shows the true relative error reduction ($H^1$ for velocity and $L^2$ for pressure) considering a great number of different geometrical configurations. Table 5.6 shows error reduction (and its magnitude) and Table 5.7 the true error over the output of interest.

![Figure 5.8: Example of flow solution (velocity $[ms^{-1} \cdot 10^{-2}]$) in the “T” shape parametrized domain.](image-url)
5.6. SOME PRELIMINARY NUMERICAL RESULTS

Figure 5.9: Basis assembling: parameters distribution during off-line reduced basis optimized assembling procedure (left). Functional output difference $\Delta s = (s - s_N)$; $s = \int_\Omega f \cdot u$: max and mean error over large test sample configurations (right).

Figure 5.10: Left: relative true $H^1$ velocity errors: max and mean error over a large test sampling (50 configurations). Right: relative true $L^2$ pressure errors: max and mean error over a large test sampling (50 configurations); errors in log-log scale.
Table 5.6: Table of $H^1$ relative errors on velocity and $L^2$ relative errors on pressure, 50 test configurations, 5(2) parameters, $N \leq 20$.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2.93e−001</td>
<td>7.86e−002</td>
<td>3.09e−002</td>
<td>6.14e−003</td>
</tr>
<tr>
<td>10</td>
<td>1.33e−002</td>
<td>4.47e−003</td>
<td>6.79e−003</td>
<td>6.96e−004</td>
</tr>
<tr>
<td>15</td>
<td>2.98e−003</td>
<td>6.65e−004</td>
<td>8.16e−004</td>
<td>9.021e−005</td>
</tr>
<tr>
<td>20</td>
<td>4.61e−004</td>
<td>8.01e−005</td>
<td>6.24e−004</td>
<td>1.60e−005</td>
</tr>
</tbody>
</table>

Table 5.7: $\Delta s = s - s_N$: max and mean error over output $s$ for 50 configurations.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\Delta s$ max</th>
<th>$N$</th>
<th>$\Delta s$ mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>9.62e−002</td>
<td>10</td>
<td>4.73e−002</td>
</tr>
<tr>
<td>10</td>
<td>2.49e−002</td>
<td>15</td>
<td>5.27e−003</td>
</tr>
<tr>
<td>15</td>
<td>2.09e−003</td>
<td>20</td>
<td>5.63e−004</td>
</tr>
<tr>
<td>20</td>
<td>7.68e−004</td>
<td></td>
<td>5.35e−005</td>
</tr>
</tbody>
</table>

5.6.3 Third Test: homogeneous Dirichlet boundary conditions, forced flow, 6(3) varying parameters

The third test we carried out considered a forced Stokes flow in the “T” domain with all zero Dirichlet boundary conditions, varying all the six parameters available (depending by two quantities: $\tau, \rho$), including the graft angle $\theta$. Data and relationship used in this test are reported below.

- The bypass diameter is $t = 3 - 2\tau$, the stenosis length is $S = \tau$, the outflow length is $L = \tau$, the artery diameter is $D = \rho$, the bypass height is $H = 2 - \rho$ and the graft angle is $\theta$.
- The viscosity is 0.04 $m^2s^{-1}$, the force field is $\mathbf{f} = (0, 10x)^T ms^{-2}$ in the true domain $\hat{\Omega}$.
- The parameters range are: $\tau \in [0.1, 1.45]$, $\rho \in [0.1, 1.9]$ and for the graft angle $\theta \in [0, \pi/3]$. 

Figure 5.11 shows the projected error reduction (as shown in Section 5.4.2) using the adaptive procedure in basis assembling, based on $H^1$ error minimization and parameters distribution during basis assembling procedure. Figure 5.12 shows true error reduction ($H^1$ for velocity and $L^2$ for pressure) considering a great number of different geometrical configurations. The plateau in pressure error plot is due to the fact that we are optimizing reduced basis velocity approximation space with adaptive procedures, but nothing is done with pressure. Table 5.8 shows error reduction (and its magnitude) and Table 5.9 the true error on an output of interest: $s(\mu) = \int_{\Omega} \mathbf{f} \cdot \mathbf{u}(\mu)d\Omega$. A second option is carried out using and adaptive off-line
procedure based on total projected error (velocity and pressure). Figure 5.13 shows total (projected) error reduction using the adaptive procedure during basis assembling. Figure 5.14 shows true error reduction ($H^1$ for velocity and $L^2$ for pressure) considering different geometrical configurations. In this case the plateau in pressure error behavior disappears because we are optimizing reduced basis velocity and pressure approximation spaces with adaptive procedures minimizing off-line error.

Figure 5.11: Basis assembling: total $H^1$ velocity (projected) error reduction during basis assembling (left) and parameters distribution during off-line reduced basis assembling (right).

Figure 5.12: Left: relative true $H^1$ velocity errors: max and mean error over a large test sampling (90 configurations.). Right: relative true $L^2$ pressure errors: Max and mean error over a large test sampling (90 configurations); errors in log-log scale.
Table 5.8: Table of velocity relative errors $H^1$ and pressure $L^2$, 90 test configurations, 6(3) parameters, $N \leq 25$.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>$1.60e-001$</td>
<td>$5.65e-002$</td>
<td>$9.55e-001$</td>
<td>$7.08e-002$</td>
</tr>
<tr>
<td>10</td>
<td>$1.01e-001$</td>
<td>$1.60e-002$</td>
<td>$9.48e-001$</td>
<td>$5.74e-002$</td>
</tr>
<tr>
<td>15</td>
<td>$1.348e-002$</td>
<td>$4.14e-003$</td>
<td>$7.02e-003$</td>
<td>$2.17e-004$</td>
</tr>
<tr>
<td>20</td>
<td>$7.82e-003$</td>
<td>$1.49e-003$</td>
<td>$5.64e-003$</td>
<td>$1.20e-004$</td>
</tr>
<tr>
<td>25</td>
<td>$4.79e-003$</td>
<td>$9.37e-004$</td>
<td>$3.63e-003$</td>
<td>$6.66e-005$</td>
</tr>
</tbody>
</table>

Table 5.9: Max and mean for 90 configurations Functional output difference $\Delta s = (s - s_N)$: Max and mean values for 90 configurations $s = \int_\Omega f \cdot u$.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\Delta s$ Max</th>
<th>$N$</th>
<th>$\Delta s$ mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>$1.34e-001$</td>
<td>5</td>
<td>$3.26e-002$</td>
</tr>
<tr>
<td>10</td>
<td>$1.26e-001$</td>
<td>10</td>
<td>$6.91e-003$</td>
</tr>
<tr>
<td>15</td>
<td>$4.064e-003$</td>
<td>15</td>
<td>$6.87e-004$</td>
</tr>
<tr>
<td>20</td>
<td>$2.50e-003$</td>
<td>20</td>
<td>$2.92e-004$</td>
</tr>
<tr>
<td>25</td>
<td>$1.38e-003$</td>
<td>25</td>
<td>$2.31e-004$</td>
</tr>
</tbody>
</table>

Figure 5.13: Basis assembling: total $H^1$ velocity and $L^2$ pressure (projected) errors reduction during basis assembling (left) and corresponding parameters distribution during off-line reduced basis assembling (right).
5.6. SOME PRELIMINARY NUMERICAL RESULTS

Figure 5.14: Left: relative true $H^1$ velocity errors: max and mean error over a large test sampling (90 configurations). Right: relative true $L^2$ pressure errors: Max and mean error over a large test sampling (90 configurations); errors in log-log scale.

5.6.4 On computational costs

In Table 5.10 we report some computational costs (CPU time) versus $N$. Finite element solution is obtained with a CPU time of 112.23, adopting a mesh with $O(10^3)$ elements, see Figure 5.2. Note the computational savings in the on-line step once we have assembled reduced basis matrix. In the same table we have inserted $H^1$ error indication to compare the computational costs and precision reached. The example refers to the case of a single parameter $D$, the arterial diameter.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$CPU_{time}$</th>
<th>$H^1$ error</th>
<th>%</th>
<th>$N$</th>
<th>$CPU_{time}$</th>
<th>$H^1$ error</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.597</td>
<td>6.24$E$ − 1</td>
<td>1.4</td>
<td>2</td>
<td>2.445</td>
<td>2.16$E$ − 1</td>
<td>2.2</td>
</tr>
<tr>
<td>3</td>
<td>3.225</td>
<td>6.58$E$ − 3</td>
<td>2.9</td>
<td>4</td>
<td>4.937</td>
<td>5.30$E$ − 4</td>
<td>4.4</td>
</tr>
<tr>
<td>5</td>
<td>5.908</td>
<td>1.92$E$ − 4</td>
<td>5.2</td>
<td>6</td>
<td>5.908</td>
<td>1.006$E$ − 4</td>
<td>5.2</td>
</tr>
<tr>
<td>7</td>
<td>6.68</td>
<td>5.27$E$ − 5</td>
<td>5.9</td>
<td>8</td>
<td>7.55</td>
<td>3.36$E$ − 5</td>
<td>6.7</td>
</tr>
<tr>
<td>9</td>
<td>8.442</td>
<td>6.83$E$ − 7</td>
<td>7.52</td>
<td>10</td>
<td>9.314</td>
<td>1.11$E$ − 7</td>
<td>8.3</td>
</tr>
<tr>
<td>11</td>
<td>10.615</td>
<td>2.66$E$ − 8</td>
<td>9.45</td>
<td>12</td>
<td>14.11</td>
<td>1.98$E$ − 8</td>
<td>12.57</td>
</tr>
<tr>
<td>13</td>
<td>14.38</td>
<td>1.25$E$ − 8</td>
<td>12.81</td>
<td>14</td>
<td>17.895</td>
<td>2.91$E$ − 12</td>
<td>15.9</td>
</tr>
<tr>
<td>15</td>
<td>20.601</td>
<td>1.38$E$ − 12</td>
<td>18.4</td>
<td>16</td>
<td>24.646</td>
<td>4.05$E$ − 13</td>
<td>21.9</td>
</tr>
<tr>
<td>17</td>
<td>24.625</td>
<td>2.20$E$ − 13</td>
<td>21.9</td>
<td>18</td>
<td>23.614</td>
<td>2.74$E$ − 14</td>
<td>21.1</td>
</tr>
</tbody>
</table>

Table 5.10: Computational costs (CPU time) varying $N$ (and $H^1$ velocity error) and comparison with computational cost of FEM solution (%).
5.7 On algebraic and approximation stability

To control the condition number of reduced basis matrix we have adopted an orthonormalization procedure applied to velocity and pressure basis functions. After orthonormalization (to achieve algebraic stability) we have to satisfy the approximation stability condition on \( N \), the equivalent reduced basis (RB) LBB \( \text{inf-sup} \) constant. But if we apply orthonormalization algorithm to reduced basis approximation spaces, assembled as proposed in Section 5.3, we may lose the validity of Lemma 5.3.1 to guarantee the stability of the approximation. For this reason we are going to propose other options in building the reduced basis velocity space.

5.7.1 Orthonormalization: Gram-Schmidt (GS) algorithm

We recall very briefly the main step of GS orthonormalization:

- Given: \( z_j, j = 1, ..., N \) a family vector of functions;
- we obtain \( q_j = \frac{P_j z_j}{||P_j z_j||} \), where \( P_j \) is an orthogonal projector onto the orthogonal complement of \( \langle q_1, q_2, ..., q_{j-1} \rangle \).
- Each \( q_j \) is orthogonal to \( q_1, q_2, ..., q_{j-1} \) and lies in \( \langle z_1, z_2, ..., z_j \rangle \).
- Each \( ||q_j|| = 1 \).
- \( P_j = I - Q_{j-1} Q_{j-1}^T; Q_{j-1} = \{ q_1, ..., q_{j-1} \}; \)
- \( P_j z_j = z_j - (q_1^T z_j) q_1 - (q_2^T z_j) q_2 - ... - (q_{j-1}^T z_j) q_{j-1}; \)
- \( \{ q_1, ..., q_N \} \) is an orthogonal basis of \( \langle z_1, z_2, ..., z_N \rangle \).

The norm \( ||\cdot|| \) used is the \( Y = (H^1(\Omega))^2 \) for velocity (and supremizers) and \( L^2(\Omega) \) for pressure. The scalar product \( q_i^T z_j \) is the one induced by the functional space and the norm we use. The orthonormalization procedure has been applied to reduced basis functions. For velocity and pressure the procedure is standard. For the supremizer we have to introduce some considerations. Referring to previous supremizer formulation of Equation 5.3.2 and to the compact notation already introduced, for \( n = N+1, \ldots, 2N \) in the reference domain we have:

\[
\langle \sigma_n, \mathbf{v} \rangle_{H^1} = \sum_{q=1}^{Q^b} \Phi^q(\mu) \langle \sigma_{qn}, \mathbf{v} \rangle_{H^1} = \sum_{q=1}^{Q^b} \Phi^q(\mu) B(\xi_{n-N}, \mathbf{v})^q \quad \forall \mathbf{v} \in (H^1_{1/2, 0}(\Omega))^2,
\]

where

\[
\langle \sigma_n, \mathbf{v} \rangle_{H^1} = B(\mu, \xi_{n-N}, \mathbf{v}), \quad (5.7.1)
\]

we recall that

\[
\langle \sigma_{qn}, \mathbf{v} \rangle_{H^1} = B(\xi_{n-N}, \mathbf{v})^q \quad (5.7.2)
\]

and

\[
B(\mu, \xi, \mathbf{v}) = \sum_{q=1}^{Q^b} \Phi^q(\mu) B(\xi, \mathbf{v})^q.
\]
At this point we have two possibilities (referring to \( n \)th supremizer \( \sigma_n, n = N + 1, \ldots, 2N \)) in applying orthonormalization:

i) an orthonormalization (GS) directly on \( \sigma_n \) done on-line (being \( \sigma_n \) dependent on \( \mu \)) to obtain \( \sigma_n^\perp \) as new element (basis function) to enrich RB velocity space:

\[
\sigma_n^\perp = \frac{P_n^\perp \sigma_n}{||P_n^\perp \sigma_n||}
\]

\[
\sigma_n^\perp = \frac{P_n^\perp (\sum_{q=1}^{Q} \Phi_q(\mu)\sigma_{q_n})}{||P_n^\perp (\sum_{q=1}^{Q} \Phi_q(\mu)\sigma_{q_n})||}
\]

\[P_n^\perp = I - Q_{n-1}Q_{n-1}^T, Q_i = \{\sigma_i^\perp, \ldots, \sigma_i^{\perp*}\};\]

ii) an orthonormalization (GS) on components \( \sigma_{q_n} \) made o-line (\( \sigma_{q_n} \) are not depending on \( \mu \)) to get \( \sigma_{q_n}^{\perp*} \):

\[
\sigma_{q_n}^{\perp*} = \frac{P_{q_n}^\perp \sigma_{q_n}}{||P_{q_n}^\perp \sigma_{q_n}||}
\]

\[P_{q_n}^\perp = I - Q_{q(n-1)}Q_{q(n-1)}^T, Q_i = \{\sigma_{q_1}^{\perp*}, \ldots, \sigma_{q_i}^{\perp*}\};\]

\[\sigma_{q_n}^{\perp*} = \sum_{k=1}^{n-1} \gamma_{q_n}^k \sigma_{q_k}^{\perp*} .\]

Figure 5.15 shows the reduction of the condition number of Stokes reduced basis linear system matrix by orthonormalizing reduced basis functions (using method (i) for suprimer, including also orthonormalization for velocity and pressure basis functions). A very interesting property (visible also in Figure 5.15) is the following:
Property 5.7.1 The condition number is limited and bounded after orthonormalization. The upper bound is given by the ratio between continuity and coercivity constant associated with the bilinear form.

Proof
We define the coercivity and continuity constant, associated with the bilinear form $A$, as:

$$
\alpha(\mu) \equiv \inf_{v \in Y} \frac{A(\mu, v, v)}{\|v\|^2_Y},
$$

(5.7.3)

$$
\gamma(\mu) \equiv \sup_{v \in Y} \frac{A(\mu, v, v)}{\|v\|^2_Y}.
$$

(5.7.4)

We now assume that our basis functions $\sigma_n, n = 1, \ldots, 2N$ are orthonormal, that is,

$$
(\sigma_i, \sigma_j)_Y = \delta_{i,j}, \ 1 \leq i, j \leq 2N,
$$

(5.7.5)

and we wish to bound the condition number of the reduced basis matrix $A^\mu$ given by

$$
A^\mu_{ij} = A(\mu, \sigma_i, \sigma_j).
$$

(5.7.6)

We note that for any $\phi \in \mathbb{R}^{2N}$,

$$
\phi^T A^\mu \phi = \sum_{i=1}^{2N} \sum_{j=1}^{2N} \phi_i \phi_j A(\mu, \sigma_i, \sigma_j)
$$

$$
\geq \alpha(\mu) \sum_{i=1}^{2N} \sum_{j=1}^{2N} \phi_i \phi_j (\sigma_i, \sigma_j)_Y = 
$$

$$
= \alpha(\mu) \sum_{i=1}^{2N} \phi_i^2.
$$

Similarly

$$
\phi^T A^\mu \phi \leq \gamma(\mu) \sum_{i=1}^{2N} \phi_i^2.
$$

(5.7.7)

It then follows that

$$
\alpha(\mu) \leq \frac{\phi^T A^\mu \phi}{\phi^T \phi} \leq \gamma(\mu) \forall \phi \in \mathbb{R}^{2N},
$$

(5.7.8)

and therefore the condition number of $A$ is bounded by $\gamma(\mu)/\beta(\mu)$. □
5.7.2 Approximation stability: other supremizer options

By orthonormalizing the supremizer solutions according to the approach (i) we could lose approximation stability (guaranteed by Lemma 5.3.1) in the attempt of preserving algebraic stability by reducing the condition number. We can orthonormalize just using method (i) pressure $\xi$ and velocity $\zeta$ basis functions and not the supremizer $\sigma_n$ solutions and use the approach (ii) to orthogonalize the supremizer on its component $\sigma_{kn}$ (before summation) to preserve Lemma 5.3.1.

To achieve this goal we may introduce two further different options in assembling the supremizer solutions for stabilization procedure by building in a different way the reduced basis velocity approximation space so that we may guarantee both approximation and algebraic stability.

**First option**

We have considered the following reduced basis spaces:

$$ Y_N = \text{span} \{ \sigma_n, n = 1, \ldots, N \mathcal{Q}^b \}, $$

where $\mathcal{Q}^b = Q^b + 1$. For $n = 1, \ldots, N$:

$$ \sigma_n = \zeta_n = u(\mu^n). $$

For $n = N + 1, \ldots, N \mathcal{Q}^b$, condensing index $m$ and $k$ in $n$, we have: $(\sigma_n, w)_Y = (\tilde{\sigma}_{mk}, w)_Y$, where

$$ (\tilde{\sigma}_{mk}, w)_Y = B(\zeta_m, w)^k, \forall w \in Y, k = 1, \ldots, Q^b, m = 1, \ldots, N; $$

$$ u_N(\mu) = \sum_{j=1}^{N \mathcal{Q}^b} u_{Nj}(\mu) \sigma_j, $$

$$ p_N(\mu) = \sum_{l=1}^{N} p_{Nl}(\mu) \xi_l; $$

the reduced basis system becomes:

$$ \begin{cases} 
\sum_{j=1}^{N \mathcal{Q}^b} A_{ij}^\mu u_{Nj}(\mu) + \sum_{l=1}^{N} B_{il}^\mu p_{Nl}(\mu) = F_i, & 1 \leq i \leq N \mathcal{Q}_{b}, \\
\sum_{j=1}^{N \mathcal{Q}^b} B_{ij}^\mu u_{Nj}(\mu) = G_l, & 1 \leq l \leq N, 
\end{cases} \tag{5.7.9} $$

where:

$$ A_{ij}^\mu = \sum_{k=1}^{Q^b} \Theta^k(\mu) A(\sigma_i, \sigma_j)^k, \quad 1 \leq i, j \leq N \mathcal{Q}_{b}, $$

$$ B_{il}^\mu = \sum_{k=1}^{Q^b} \Phi^k(\mu) B(\sigma_i, \xi_l)^k, \quad 1 \leq i \leq N \mathcal{Q}_{b}, \quad 1 \leq l \leq N, $$

$$ F_i = \langle F, \sigma_i \rangle, \quad 1 \leq i \leq N \mathcal{Q}_{b}, \quad G_l = \langle G^b, \xi_l \rangle, \quad 1 \leq l \leq N. $$
In this case the basis is no longer \( \mu \) (on line) dependent, the reduced basis velocity space, enriched by supremizers, has a bigger dimension \((NQ_b > 2N)\) than previously. The computational costs are as follows: \( O(Q^a(Q^b)^2N^2) \) for sub-matrix \( \mathbf{A} \), \( O((Q^b)^2N^2) \) for \( \mathbf{B} \), \( O(Q^bN) \) for \( F \), but the cost for inversion of the full reduced basis matrix (6.4.6) increases now to \( O((Q^b + 1)^3N^3) \). This approach has the big advantage to preserve Lemma 5.3.1, to let us apply orthonormalization (method (ii)) and to preserve the mentioned Lemma also after orthonormalization. This approach is the best if we want to be sure to preserve both algebraic and approximation stability after orthonormalization, if necessary. Figure 5.16 shows errors behavior by testing this first new supremizer option over a large test sampling (zero Dirichlet conditions and 3 geometrical parameters).

![Figure 5.16](image)

Figure 5.16: New supremizer first option, \( H^1 \) relative error (for velocity) and \( L^2 \) relative error (for pressure) using 50 configurations.

**Second option**

Another approach is based on the idea that supremizers are built upon summation using the same \( \mu^j \) values used to store velocity \( \zeta_j(\mu^j) \) and pressure solutions \( \xi_j(\mu^j) \) (also in this case the basis for velocity is not dependent on the on-line value of \( \mu \) and it is completely assembled off-line):

\[
Y_N = \text{span} \{ \sigma_n = \sum_{k=1}^{Q^b} \Phi^k(\mu^n)\sigma_{kn}, \ n = 1, \ldots, 2N \},
\]

where \( Q_b = Q^b + 1, \Phi(Q^b) = 1 \). For \( n = 1, \ldots, N \):

\[
\sigma_{kn} = 0, \text{ for } k = 1, \ldots, Q^b; \sigma_{Q^b n} = \zeta_n = u(\mu^n).
\]

For \( n = N + 1, \ldots, 2N \):

\[
(\sigma_{kn}, w)_Y = B(\xi_{n-N}, w)^k, \forall w \in Y, \text{ for } k = 1, \ldots, Q^b.
\]
\[ \sigma_{Q^n} = 0. \]

The reduced basis solution is given by

\[ u_N(\mu) = \sum_{j=1}^{2N} u_{Nj}(\mu) \left( \sum_{k=1}^{Q^b} \Phi^k(\mu^j) \sigma_{kj} \right), \]

\[ p_N(\mu) = \sum_{l=1}^{N} p_{Nl}(\mu) \xi_l, \]

by solving the system:

\[
\begin{cases}
\sum_{j=1}^{2N} A_{ij}^\mu u_{Nj}(\mu) + \sum_{l=1}^{N} B_{il}^\mu p_{Nl}(\mu) = F_i, & 1 \leq i \leq 2N, \\
\sum_{j=1}^{2N} B_{ij}^\mu u_{Nj}(\mu) = G_l, & 1 \leq l \leq N;
\end{cases}
\]

where:

\[ A_{ij}^\mu = \sum_{k=1}^{Q^a} \Theta^k(\mu) A(\sigma_i, \sigma_j)^k = \]

\[ = \sum_{k=1}^{Q^a} \sum_{k'=1}^{Q^b} \sum_{k''=1}^{Q^b} \Theta^k(\mu) \Phi^{k'}(\mu^i) \Phi^{k''}(\mu^j) A(\sigma_{k'i}, \sigma_{k''j})^k, & 1 \leq i, j \leq 2N; \]

\[ B_{il}^\mu = \sum_{k=1}^{Q^b} \Phi^k(\mu) B(\sigma_i, \xi_l)^k = \]

\[ = \sum_{k=1}^{Q^b} \sum_{k'=1}^{Q^b} \Phi^k(\mu) \Phi^{k'}(\mu^i) B(\sigma_{k'i}, \xi_l)^k, & 1 \leq i \leq 2N, 1 \leq l \leq N; \]

\[ F_i = \langle F, \sigma_i \rangle = \sum_{k'=1}^{Q^b} \Phi^{k'}(\mu^i) \langle F, \sigma_{k'i} \rangle, & 1 \leq i \leq 2N, G_l = \langle G^0, \xi_l \rangle, & 1 \leq i, l \leq N. \]

This option is also competitive concerning computational costs dealing with \(3N \times 3N\) reduced basis matrices (5.3.4) instead of \((Q^b + 1)N \times (Q^b + 1)N\) matrix (usually \((Q^b + 1) \gg 3\)). We have the following computational costs to build reduced basis matrices, given also the supremizer components in the velocity space: \(O(Q^a 4N^2)\) for sub-matrix \(A\), \(O(Q^b 2N^2)\) for \(B\), \(O(N)\) for \(F\) and \(O(9N^3)\) for the inversion of the full reduced basis matrix (5.3.4).

Using this option we cannot demonstrate that Lemma 5.3.1 is preserved (even without orthonormalization). We have tested numerically this option and we can argue that also this approximation is reasonably stable. Numerical results are shown in Figure 5.17 where we have reported errors behavior always testing a large number of configurations (Dirichlet conditions and 3 varying parameters). Convergence is very fast in this case. Figure 5.18 shows condition number reduction of reduced basis linear system and a comparison of condition number of the matrix without orthonormalization.
Figure 5.17: New supremizer second option, $H^1$ relative error (velocity) and $L^2$ relative error (pressure) over 50 configurations.

Figure 5.18: Condition number (max and mean) for the Stokes reduced basis system, with the new supremizer second option. On the left results with orthonormalization, on the right without it.
5.8 Some preliminary results on bypass configurations

As preliminary reported in [136] and [137] we start introducing some results on the complete bypass configurations using reduced basis techniques. We refer to the configuration reported in Figure 5.19. The idea has been the use of reduced basis (with also different approximation spaces) to test a huge number of different bypass configurations (i.e. a great number of parameters combinations).

As measure of blood flow perturbation we have considered for example the mean blood velocity:

$$s(\mu) = s_1(\mu) + s_2(\mu) = \sum_{r=1}^{R} \int_{\Omega_r} u_1 d\Omega + \sum_{r=1}^{R} \int_{\Omega_r} u_2 d\Omega,$$

(5.8.1)

We underline that, as seen in Chapter 2, also the velocity field is an interesting quantity to be studied in parametrized configuration, above all if we consider inflow and outflow of the downfield zone of the bypass configuration (being related with vorticity by Green’s Theorem). In our case the inflow is the intersection between the new bridge and the host artery which is the zone studied in the local shape optimization approach.

With great computational costs savings we can provide in real time useful clinical indication dealing with a great number (i.e. hundreds) of bypass configurations and to understand the role of each geometrical parameter and their reciprocal influence. Numerical results indicate a very good convergence behavior and a tight control on the maximum $N$. Numerical tests on the bypass configuration (Figure 5.19) have been carried out imposing a mean Reynolds number of $10^3$, a blood kinematic viscosity $\nu = 4 \cdot 10^{-6} m^2 s^{-1}$ and a force field: $f = (0, 9.8)^T ms^{-2}$.

Solutions used as basis functions are always obtained by Galerkin-Finite Element method with Taylor-Hood elements ($P^2$ and $P^4$ for velocity and pressure, respectively). Figure 5.20 shows good convergence of the relative mean errors ($H^1$ for velocity and $L^2$ for pressure) testing a great number of configurations. We have carried out three different tests on parametrized families of bypass configurations by our input-output methodology. Figure 5.21 shows the first case of study where we have investigated the bypass graft angle perturbation (other parameters are frozen) measuring the increase of our output of interest (5.8.1): varying $\theta$ form $0$ to $\pi/6$ the increase of the mean blood flow is very high in the range $[0, \pi/6]$ and smoothed in the range $[\pi/6, \pi/3]$. Results are shown for different $N$ to underline the fidelity of approximation with a few basis functions. Experimental results reported in [146] and other numerical simulations described in [152], [140], [141] and [60] provide the same interpretation of the role played by graft angle into an end-to-side anastomosis concerning flow separation and perturbation. Figure 5.22 shows the flow perturbations with respect to the quantity
$\frac{S}{D}$: the ratio between the stenosis length and the arterial diameter (best performances when $\frac{S}{D} \geq 1$) and the quantity $\frac{D}{t}$ (improving performances when the ratio is less than unity, i.e. bypass diameter smaller than arterial diameter). These results can be linked with the ones obtained during shape optimization where we find that it is better to get a flow which is diminishing its velocity when it is approaching the host artery (section is becoming larger, i.e. $D > t$, but also $S > D$). These results are a consequence of the continuity equation and of the fact that we are dealing with incompressible fluids.

![Figure 5.19: Schematic bypass configuration and reference domain.](image)

![Figure 5.20: Reduced basis convergence results: mean error on velocity and pressure.](image)

![Figure 5.21: Output $s [ms^{-1} \cdot 10^{-2}]$ versus the parameter $\theta$ at different $N$.](image)
5.9 Non-compliant outputs: dual residual approach

Finally, as it is typical in engineering practice, we assume that we are not interested in calculating only the solution of parametrized problem (5.3.1). Rather we are interested in obtaining performance measures that characterize the particular configuration $\mu \in \mathcal{D}$ and have physical importance like, for example, flow-rate, vorticity, shear stress. Given the solution $[u(\mu), p(\mu)]$ to (5.3.1), the output of interest can be written as:

$$s(\mu) = \ell^O([u(\mu), p(\mu)]; \mu) = \ell^O_u(u(\mu); \mu) + \ell^O_p(p(\mu); \mu);$$

(5.9.1)

with $\ell^O_u(\cdot; \mu) \in Y'$, $\forall \mu \in \mathcal{D}$ and $\ell^O_p(\cdot; \mu) \in Q'$, $\forall \mu \in \mathcal{D}$, which implies that $\ell^O(\cdot; \mu) \in (Y \times Q)'$, $\forall \mu \in \mathcal{D}$ is a bounded linear functional. We also require in the following a dual, or adjoint, problem associated with $\ell^O(\cdot; \mu)$: find $[\psi(\mu), \lambda(\mu)] \in Y \times Q$ such that

$$\left\{ \begin{array}{l}
\mathcal{A}(v, \psi(\mu); \mu) + \mathcal{B}(v, \lambda(\mu); \mu) = -\ell^O_u(v; \mu) \quad \forall v \in Y, \\
\mathcal{B}(\psi(\mu), q; \mu) = -\ell^O_p(q; \mu) \quad \forall q \in Q.
\end{array} \right.$$  

(5.9.2)

We note that $\mathcal{A}$ and $\mathcal{B}$ are the continuous bilinear forms introduced in (5.2.10) and (5.2.11), respectively. The adjoint problem (5.9.2) is subject to the same inf-sup condition (5.2.21) of the state problem.

Finally, we make the assumption of affine parameter dependence also for the linear form representing the output:

$$\ell^O([w, q]; \mu) = \sum_{q=1}^{M_O} \pi^q(\mu) \ell^O_q([w, q]),$$

(5.9.3)

$\forall w \in Y$, $\forall q \in Q$, where $M_O$ is an integer that depend on the problem (and output) in consideration. For a general overview on the adjoint problem and its reduced basis formulation see Patera et al. [118] and Rovas [134].
5.10 Reduced basis approximation for adjoint problem

We next define our primal and dual reduced-basis approximation spaces. To wit, for the primal and dual problem we choose \( N \) and \( M \) points: \( \mu^i, \ i = 1, \ldots, N \) and \( \mu^i, \ i = 1, \ldots, M \), respectively, not necessarily the same as for the primal problem in our parameter set \( \mathcal{D} \). We denote the collection of the samples as:

\[
\mathcal{S}_N^{pr} = \{ \mu^1, \ldots, \mu^N \}, \quad \mathcal{S}_M^{du} = \{ \mu^1, \ldots, \mu^M \}.
\]

We then compute \([u(\mu^i), p(\mu^i)] \in Y \times Q \) and \([\psi(\mu^i), \lambda(\mu^i)] \in Y \times Q \), the solutions of (5.3.1) for the primal problem \((pr)\) and (5.9.2) for the dual one \((du)\), respectively, for all \( \mu^i \in \mathcal{S}_N^{pr} \) and \( \mu^i \in \mathcal{S}_M^{du} \). We have to calculate the quantities \( \sigma_{kn}^{pr} \in Y \), and \( \sigma_{kn}^{du} \in Y \) \( q = 1, \ldots, Q^b \), and \( n = N + 1, \ldots, 2N \) or \( n = M + 1, \ldots, 2M \), respectively, which satisfy the supremizer problem:

\[
(\sigma_{kn}^{pr}, v)_Y = B^k(v, p(\mu^{n-N})), \quad \forall v \in Y, \ k = 1, \ldots, Q^b, \ n = N + 1, \ldots, 2N \tag{5.10.1}
\]

\[
(\sigma_{kn}^{du}, v)_Y = B^k(v, \lambda(\mu^{n-M})), \quad \forall v \in Y, \ k = 1, \ldots, Q^b, \ n = M + 1, \ldots, 2M. \tag{5.10.2}
\]

We then define the primal and dual “pressure” approximation spaces \( Q_N^{pr} \) and \( Q_M^{du} \):\n
\[
Q_N^{pr} = \text{span} \{ p(\mu^i), \ i = 1, \ldots, N \} \equiv \text{span} \{ \xi_i^{pr}, \ i = 1, \ldots, N \}, \tag{5.10.3}
\]

\[
Q_M^{du} = \text{span} \{ \lambda(\mu^i), \ i = 1, \ldots, M \} \equiv \text{span} \{ \xi_i^{du}, \ i = 1, \ldots, M \}; \tag{5.10.4}
\]

and the “velocity” approximation spaces \( Y_N^{pr}(\mu) \) and \( Y_M^{du}(\mu) \) (referring to the first supremizer option with \( N \) or \( M \) pressure functions and \( 2N \) or \( 2M \) velocity functions for primal and dual problem, respectively)\n
\[
Y_N^{pr}(\mu) = \text{span} \left\{ u(\mu^i), \sum_{q=1}^{Q^b} \Phi^q(\mu) \sigma_{q(N+i)}^{pr}, \ i = 1, \ldots, N \right\} \equiv \text{span} \{ \sigma_i^{pr}, \ i = 1, \ldots, 2N \}, \tag{5.10.5}
\]

\[
Y_M^{du}(\mu) = \text{span} \left\{ \psi(\mu^i), \sum_{q=1}^{Q^b} \Phi^q(\mu) \sigma_{q(M+i)}^{du}, \ i = 1, \ldots, M \right\} \equiv \text{span} \{ \sigma_i^{du}, \ i = 1, \ldots, 2M \}; \tag{5.10.6}
\]

with dimensions \( \dim Q_N^{pr} = N, \dim Y_N^{pr}(\mu) = 2N, \dim Q_M^{du} = M, \) and \( \dim Y_M^{du}(\mu) = 2M. \)

In the construction of the reduced-basis spaces, we do not necessarily need to choose an equal number of pressure and velocity modes for primal and dual problem. We can also choose \( N_u^{pr} \) velocity basis functions for \( Y_N^{pr}(\mu) \), and \( N_p^{pr} \) basis functions for the for \( Q_N^{pr} \) and the same for the dual problem. Another remark deals with the existence of two different kinds of approach for the primal and dual reduced basis approximation. Here we have introduced the so-called “non-integrated” (building different spaces for primal and dual problem) approach, as used in [134], another method is introduced in [118] where the “integrated” approach is used: in this case the approximation space for primal and dual problem is the same and made up of
both $\zeta_i^{pr}$ and $\zeta_i^{du}$. The approach used here (the “non-integrated”) has significant advantages concerning computational costs and conditioning aspects.

Using the definitions of reduced basis approximation spaces we can define the reduced-basis problems for primal and dual problem. We look for $[u_N(\mu), p_N(\mu)] \in Y_N^{pr}(\mu) \times Q_N^{pr}$ and $[\psi_M(\mu), \lambda_M(\mu)] \in Y_M^{du}(\mu) \times Q_M^{du}$, such that:

$$
\begin{align*}
\mathcal{A}(u_N(\mu); \mu) + B(v, p_N(\mu); \mu) &= \ell(v; \mu) \quad \forall v \in Y_N^{pr}(\mu), \\
B(u_N(\mu), q; \mu) &= 0 \quad \forall q \in Q_N^{pr},
\end{align*}
$$

(5.10.7)

and,

$$
\begin{align*}
\mathcal{A}(v, \psi_M(\mu); \mu) + B(v, \lambda_M(\mu); \mu) &= -\ell^O(v; \mu) \quad \forall v \in Y_M^{du}(\mu), \\
B(\psi_M(\mu), q; \mu) &= -\ell^O_p(q; \mu), \quad \forall q \in Q_M^{du},
\end{align*}
$$

(5.10.8)

respectively. The solution of the dual problem is carried out as already seen for the state problem with the splitting of the procedure into an off-line (performed once) and online part (performed many times). If $[u_N, p_N] \in Y_N^{pr}(\mu) \times Q_N^{pr}$ and, $[e^{pr}, e^{pr}_p](\mu) \equiv [u - u_N, p - p_M](\mu)$ is the error, the residual $R^{pr}_u(\cdot; [u_N, p_N]; \mu) \in Y'$ is defined

$$
R^{pr}_u(v; [u_N, p_N]; \mu) = \ell(v; \mu) - \mathcal{A}(u_N(\mu), v; \mu) - B(v, p_N(\mu); \mu) =
$$

$$
= \mathcal{A}(e^{pr}_u(\mu), v; \mu) + B(v, e^{pr}_p(\mu)) \quad \forall v \in Y,
$$

(5.10.9)

where the second line follows from equation (5.3.1). Similarly the residual related to the incompressibility constraint $R^{pr}_p(\cdot; [u_N, p_N]; \mu) \in Q'$ is

$$
R^{pr}_p(q; [u_N, p_N]; \mu) = -B(u_N(\mu), q; \mu) = B(e^{pr}_u(\mu), q; \mu) \quad \forall q \in Q.
$$

(5.10.10)

We can the define the primal residual $R^{pr}(\cdot; [w_N, p_N]; \mu) \in (Y \times Q)'$, from

$$
R^{pr}([w, q]; [u_N, p_N]; \mu) = R^{pr}_u(w; [u_N, p_N]; \mu) + R^{pr}_p(q; [u_N, p_N]; \mu) =
$$

$$
= \mathcal{A}(e^{pr}_u(\mu), w; \mu) + B(w, e^{pr}_p(\mu)) + B(e^{pr}_u(\mu), q; \mu) \quad \forall [w, q] \in Y \times Q.
$$

(5.10.11)

For the dual problem, if $[\psi_M, \lambda_M] \in Y_M^{du}(\mu) \times Q_M^{du}$ and, $[e^{du}_u, e^{du}_p](\mu) \equiv [\psi - \psi_M, \lambda - \lambda_M](\mu)$ is the error, we define in a similar way the residuals $R^{du}_u(\cdot; [\psi_M, \lambda_M]; \mu) \in Y'$ and $R^{du}_p(\cdot; [\psi_M, \lambda_M]; \mu) \in Q'$:

$$
R^{du}_u(v; [\psi_M, \lambda_M]; \mu) = -\ell^O_u(v; \mu) - \mathcal{A}(v, \psi_M(\mu); \mu) - B(v, \lambda_M(\mu); \mu) =
$$

$$
= \mathcal{A}(e^{du}_u(\mu), v; \mu) + B(v, e^{du}_p(\mu)) \quad \forall v \in Y,
$$

(5.10.12)

and

$$
R^{du}_p(q; [\psi_M, \lambda_M]; \mu) = -\ell^O_p(q; \mu) - B(\psi_M(\mu), q; \mu) =
$$

$$
= B(e^{du}_u(\mu), q; \mu) \quad \forall q \in Q.
$$

(5.10.13)

The dual residual is then $R^{du}(\cdot; [w_M, q_M]; \mu) \in (Y \times Q)'$ is then

$$
R^{du}([w, q]; [\psi_M, \lambda_M]; \mu) = R^{du}_u(w; [\psi_M, \lambda_M]; \mu) + R^{du}_p(q; [\psi_M, \lambda_M]; \mu)
$$

$$
= \mathcal{A}(w, e^{du}_u(\mu)) + B(w, e^{du}_p(\mu)) + b(e^{du}_u(\mu), q; \mu) \quad \forall [w, q] \in Y \times Q.
$$

(5.10.14)
The output approximation is then obtained from
\[ s_N(\mu) = \ell_0^O(\{u_N, p_N\}(\mu); \mu) - R_{\text{pr}}^p([\psi_M, \lambda_M](\mu); [u_N, p_N](\mu); \mu); \]  
(5.10.15)
where the adjoint correction helps improving the accuracy of the approximation. In Patera and Renquist [110] a detailed analysis concerning the output correction is provided. For a more general framework we suggest to see also Giles and Pierce [114]. For the computation of the primal and dual basis functions, a total of \( N + M \) Stokes problems need to be solved. In addition, \((N + M)Q_b\) \(Y\)-solves are required for the calculation of supremizers \(\sigma_{kn}\). Finally, a number of matrix-vector and inner products are required for the formation of a number of auxiliary quantities. The important thing to note it that once the expensive and memory intensive off-line part is completed, a database with \(O((N^2 + M^2)Q_a Q_b^2)\) quantities, is created. In the on-line part, for each new \(\mu \in D\), and using this database: first, \(O((N^2 + M^2)Q_a Q_b^2)\) operations are required to form the reduced-basis problems; and second \(O(N^3 + M^3)\) operations are required to invert the resulting linear systems and compute the output approximation. Note that during on-line calculation no explicit reference is made to the continuous or, in practice, finite-element problem and so computational costs are not depending on the costs of finite-element solution. As \(N\) and \(M\) will typically be small, significant computational savings are expected.

5.10.1 Some elements for output error estimation

Let \([u, p](\mu) \in Y \times Q\) be the exact solution for the primal problem (5.3.1), and \([u_N, p_N](\mu) \in Y_N \times Q_N\) the reduced-basis approximation obtained by solving (5.10.7). Subtracting (5.3.1) and (5.10.7), the error \([e_{u^p}^p, e_{p^p}^p](\mu) \equiv [u - u_N, p - p_N] \in Y \times Q\) to the primal problem satisfies the following equation:
\[ \begin{aligned}
&\mathcal{A}(e_{u^p}^p(\mu), \nu; \mu) + \mathcal{B}(\nu, e_{p^p}^p(\mu)) = R_{u^p}^p(\nu; [u_N, p_N](\mu); \mu), \quad \forall \nu \in Y, \\
&\mathcal{B}(e_{u^p}^p(\mu), q; \mu) = R_{p^p}^p(q; [u_N, p_N](\mu); \mu), \quad \forall q \in Q,
\end{aligned} \]
(5.10.16)
with similar equation valid for the dual error \([e_{u^d}^d, e_{p^d}^d](\mu) \equiv [\psi - \psi_M, \lambda - \lambda_M](\mu) \in Y \times Q\).

We can get a posteriori error estimator for the output. From (5.9.1) and (5.10.15) the error in the output is given by
\[ s(\mu) - s_N(\mu) = \ell_0^O(u(\mu); \mu) + \ell_0^p(p(\mu); \mu) - \ell_0^O(u_N(\mu); \mu) - \ell_0^p(p_N(\mu); \mu) - R_{\text{pr}}^p([\psi_M, \lambda_M](\mu); [u_N, p_N](\mu); \mu) = \ell_0^O(e_{u^p}^p(\mu); \mu) + \ell_0^p(e_{p^p}^p(\mu); \mu) + R_{\text{pr}}^p([\psi_M, \lambda_M](\mu); [u_N, p_N](\mu); \mu) \]
which from the definition of the adjoint problem (5.9.2) and the primal residual (5.10.11) can be written as
\[ s(\mu) - s_N(\mu) = -\mathcal{A}(e_{u^p}^p(\mu), \psi(\mu); \mu) - \mathcal{B}(e_{p^p}^p(\mu), \lambda(\mu); \mu) - \mathcal{B}(\psi(\mu), e_{p^p}^p(\mu); \mu) + \mathcal{A}(e_{u^d}^d(\mu), \psi_M(\mu); \mu) + \mathcal{B}(e_{p^d}^d(\mu), \lambda_M(\mu); \mu) + \mathcal{B}(\psi_M(\mu), e_{u^d}^d(\mu); \mu) = \ell_0^O(e_{u^p}^p(\mu); \mu) - e_{u^d}^d(\mu; \mu) - e_{p^d}^d(\mu; \mu) - B(e_{u^d}^d(\mu), e_{u^d}^d(\mu); \mu) = -R_{u^d}(\psi_M, \lambda_M)(\mu; \mu) - R_{p^d}(\psi_M, \lambda_M)(\mu; \mu); \]
here the definitions for the primal and dual residuals (5.10.11) and (5.10.14), respectively, have been used. We then have:

\[ |s(\mu) - s_N(\mu)| \leq \sup_{v \in V} \frac{R^{du}(\psi_{M,M}|(\mu)\mu)}{||v||_Y} (e^p_{u}(\mu) + \sup_{q \in Q} \frac{R^{du}(\psi_{M,M}|(\mu)\mu)}{||q||_Q} (e^p_{u}(\mu) Q )

\]

obtaining a bound depending on the dual norms for the dual residuals multiplied for the primal error. We can conclude that if we consider corrected outputs with dual-residual we have a faster convergence and a greater accuracy with respect to the non-dual corrected version because in the dual corrected case the output error is bounded by a product of dual residuals and primal errors. There is a trade-off between accuracy and computational cost: for example, for exponentially convergent primal and dual approximation (the worst case concerning the advantage of the dual approach), the primal-dual combined approach gives comparable accuracy at 1/4 the computational cost even if the problem complexity has increased in solving an adjoint problem and assembling dual basis.

### 5.10.2 Results using non-compliant outputs

In Figure 5.23 we present a comparison considering the mean error on output using the dual residual correction with respect to the non-corrected one (considering only Stokes primal problem). The improvement in accuracy achieved is shown in the worst case, where the primal space has already some approximation properties for the dual problem. The test is concerned with a 4-parameters bypass configuration (using several parameters combinations) and the “non-compliant” output considered was the distributed vorticity:

\[ s = \int_{\Omega} (\frac{\partial u_1}{\partial x_2} - \frac{\partial u_2}{\partial x_1}) d\Omega. \]

In Figure 5.24 we show the convergence for primal and dual error ($H^1$ velocity relative errors) and a mean relative error over the non-compliant vorticity output considering a great number of different 3-parameters bypass configuration. In this case we can see the “square effect” (the output error is bounded by the product between the primal error with the dual residual as shown in 5.10.17) and we have an accuracy in the dual corrected case for $N = 10$ which is the same we get for $N = 20$ in the non-corrected case and with computational costs which are 1/8 (based on LU scaling) with respect to the non-corrected approach. In Figure 5.25 we show some results dealing with the dual corrected output error, the relative $H^1$ velocity error for primal problem, the error for the non-corrected output ($\tilde{s}_N(\mu)$) and a theoretical upper bound for the non-corrected output. The latter is given by

\[ s(\mu) - \tilde{s}_N(\mu) = \ell_{u}^{O}(\mu(\mu), \mu) + \ell_{p}^{O}(\mu, \mu)
- \ell_{u}^{O}(\mu_N(\mu), \mu) - \ell_{p}^{O}(\mu_N(\mu), \mu) =
\ell_{u}^{O}(e_{u}^{pr}(\mu), \mu) + \ell_{p}^{O}(e_{p}^{pr}(\mu), \mu) =
\ell_{u}^{O}(e_{u}^{pr}(\mu), \mu) \|e_{u}^{pr}(\mu)\|_Y + \ell_{p}^{O}(e_{p}^{pr}(\mu), \mu) \|e_{p}^{pr}(\mu)\|_Q
\leq \sup_{v \in V} \frac{\ell_{u}^{O}(\psi_{M,M}|(\mu)\mu)}{||v||_Y} (e_{u}^{pr}(\mu) + \sup_{q \in Q} \frac{\ell_{p}^{O}(\psi_{M,M}|(\mu)\mu)}{||q||_Q} (e_{p}^{pr}(\mu) Q ),
\]
where the upper bound is the dual norm of the linear functional times the primal error. Also
in this case we can see how the dual correction improves the output accuracy.

Now we can turn our attention on our main application and study bypass parametrized
configuration with more complex outputs. We report some summary results dealing with 4
varying parameters (i.e. \( t, D, S, \theta \)). Results are an extension of the ones obtained considering
the flow perturbation due to the increase of mean velocity. In Figures 5.26 and 5.27 we have
summarized our results by reporting only few spots (each representing a different configura-
tion) to keep the plot quite clear, which show the distributed vorticity \( m^2 \cdot s^{-1} \) cost functional
with respect to \( \frac{S}{\theta} \) (reported in \( x_1 \) axis) and \( \frac{D}{\theta} \) (represented by the colored scale) and \( \theta \) graft
angle (reported in \( x_2 \) axis and represented by the dimension of the colored spots). We can
conclude that quantities \( \frac{S}{\theta} \) and \( \frac{D}{\theta} \) are more important and only freezing them the graft angle
has a crucial role (to have a look on some numerical results dealing with blood recirculation
at different graft angle see [141], where other geometrical parameters are frozen). Vorticity
is reduced when the ratio \( \frac{S}{\theta} \) is bigger than unity and when \( \frac{D}{\theta} \) is smaller than unity. These

Figure 5.23: Max output error (comparison with and without correction).

Figure 5.24: \( H^1 \) velocity relative error for dual and primal problems and output relative
errors.
results are a further explanation of fluid mechanics phenomena related with geometry for incompressible flows. We want also to underline how the output seems to be sensible with the changes in bypass configuration.

Figure 5.25: $H^1$ error, upper bound for output error (non-corrected), adjoint-corrected and non-corrected output error.

Figure 5.26: Output $s_N [m^2 s^{-1} \cdot 10^{-2}]$ versus $\theta$ parameter (spots dimension) and the ratios $\frac{\tau}{\pi}$ (colorbar) and $\frac{\nu}{\pi}$. Real-time “clinical” indications by a great number (i.e. hundreds) of bypass configurations (Re= $10^3$).
5.11 Conclusions

We have provided some elements to solve rapidly and efficiently Stokes equations in parametrized domains and to extract information on some fluid mechanics outputs and a sensitivity analysis on some geometrical parameters, describing a complete configuration for a bypass. Reduced basis techniques allow us huge computational savings and real-time results and they guarantee at the same time the stability of the approximation. To complete this study we have to extend the application of reduced basis to Navier-Stokes equations in parametrized domains and in problems involving non-affine mapping dependence (i.e. shape design problem with curved walls). Some recent works are [159] for Navier-Stokes reduced basis formulation and references [145] and [12] dealing with non-affine parameter dependence. The latter especially introduces an efficient reduced basis discretization procedure replacing non-affine coefficient functions with a collateral reduced basis expansion which permits an offline-online computational decomposition by a stable and inexpensive empirical interpolation procedure. These issues will make the subject of the following chapter.
Chapter 6

Reduced Basis Methods for Stokes Equations in Domains with Non-Affine Parametric Dependence

In this chapter we extend reduced basis techniques applied to Stokes equations in domains parametrized by affine and non-affine maps with respect to a reference domain. The proposed method is ideally suited also for shape optimization problem with more realistic configuration. An “empirical”, stable and inexpensive interpolation method has permitted to replace non-affine coefficient functions with an expansion which leads to a computational decomposition between the offline–parameters independent–stage for reduced basis generation and the online–parameters dependent–approximation stage based on Galerkin projection, used to find a new solution for a new set of parameters by a combination of previously computed stored solutions. As in the affine case this computational decomposition leads us to preserve reduced basis properties: rapid and accurate convergence and computational economies. The applications and results are based on parametrized geometries describing domains with curved walls, a preliminary test based on a stenosed channel and then some tests on the bypass configuration. This method is well suited to treat also problems in fixed domains with non-affine parameters dependence expressing varying physical coefficients.

6.1 An extension on the use of reduced basis

This chapter is an extension of applications and results collected in [138] dealing with the application of reduced basis techniques for Stokes equations in curved parametrized domains. After this introduction, in Section 6.2 the empirical interpolation procedure proposed by Maday et al. [12] and applied to non-affine transformation terms, mapping the real domain into a reference one, is briefly described. In this chapter we focus on the use of this interpolation procedure for geometrical non-affine transformation terms, the coupling with the affine ones and the introduction of parametrized complex geometries in the reduced basis problems. In Section 6.3 we recall the parametrized Stokes equations framework coupling affine and
non-affine parametric dependence. In Section 6.4 we introduce Stokes reduced basis formulation extended to the non-affine case. Then in Section 6.5 we introduce some numerical results based on two different geometries. Finally, in Section 6.6 we give a preview on the development guidelines for viscous flow and shape optimization problems.

6.2 Empirical interpolation for (coefficient) functions approximation

To start with we recall the empirical interpolation procedure, proposed in [12]. This procedure has been applied also in Grepl [44] for parabolic equations. We consider a (coefficient) function \( g(x, \mu) \), depending on spatial coordinates and on a set of parameters \( \mu \in \mathcal{D} \subset \mathbb{R}^P \) (for some \( P \geq 1 \)). The function \( g(x, \mu) \) represents, for example, a coefficient for a linear or bilinear form, that shows up when a non-affine mapping transformation of a physical domain into a reference one indicated with \( \Omega \) is applied. We assume that \( g(x, \mu) \in L^\infty(\Omega) \) for all choice of \( \mu \). Our goal is to rewrite this function as an expansion given by products between parameters dependent coefficients and “shape functions” depending only on spatial coordinates. We introduce \( g \) as a suitably fine parameter sample over \( \mathcal{D} \) and the related quantities

\[
\mu_M^q = \arg\max_{\mu \in \Xi^q} \inf_{z \in W_{M-1}^q} \|g(\cdot, \mu) - z\|_{L^\infty(\Omega)}
\]

to build the sets \( S_M^q = \{\mu_m^q, 1 \leq m \leq M\}, 1 \leq M \leq M_{max}, \) \( S_M^q = S_{M-1}^q \cup \mu_M^q \) and the following approximation spaces

\[
W_M^q = \text{span}\{g_m^q(x), 1 \leq m \leq M\}, 1 \leq M \leq M_{max}.
\]

The quantity \( \mu_M^q \) is chosen “a priori” so that \( \gamma_1(x) \neq 0 \). We need to construct nested sets of interpolation points:

\[
T_M = \{t_1, \ldots, t_M\}, 1 \leq M \leq M_{max},
\]

by the following algorithm. Starting from \( M = 1 \) we store

\[
\gamma_1(x) = g(x, \mu_1^q), t_1 = \arg\sup_{x \in \Omega} |\gamma_1(x)|, \quad q_1(x) = \gamma_1(x)/\gamma_1(t_1),
\]

for \( M \geq 2 \) we have to solve a linear system to get \( \sigma_{M-1} \) and then \( r_M(x) \) in assembling \( T_M \):

\[
M = 2, \ldots, M_{max} : \Sigma_{j=1}^{M-1} \sigma_j^{M-1} q_j(t_i) = \gamma_M(t_i), 1 \leq i \leq M-1;
\]

\[
r_M(x) = \gamma_M(x) - \Sigma_{j=1}^{M-1} \sigma_j^{M-1} q_j(x), t_M = \arg\sup_{x \in \Omega} |r_M(x)|,
q_M(x) = r_M(x)/r_M(t_M).
\]

At the end of the algorithm our function is approximated by \( g_M(x, \mu) \) and split into two parts (decoupled) each of them depending only on the \( \mu \) parameter (\( \lambda_m(\mu) \)) or on \( x \) coordinates (\( q_m(x) \)), i.e:

\[
g_M(x, \mu) = \Sigma_{m=1}^{M} \lambda_m(\mu) q_m(x),
\]
where \( \lambda_j \) are given by the solution of the following linear system

\[
\Sigma_{j=1}^{M} q_j(t_i) \lambda_j(\mu) = g(t_i, \mu), \quad 1 \leq i \leq M.
\]

We choose \( M_{\text{max}} \) as being the minimum \( M \) s.t. the maximum interpolation error:

\[
\varepsilon_M(\mu) = \|g(\cdot, \mu) - g_M(\cdot, \mu)\|_{L^\infty(\Omega)} \quad (6.2.1)
\]

satisfies \( \varepsilon_M \leq \varepsilon_{\text{max}} \) (for a prescribed tolerance \( \varepsilon_{\text{max}} \)).

This interpolation process can be justified. A priori we can introduce a Lebesgue constant

\[
M = \sup_{x \in \Omega} \sum_{m=1}^{M} |V^M_m(x)|,
\]

where \( V^M_m(\cdot) = g_M(\cdot, \mu) \). The bound on the Lebesgue constant is pessimistic and compensated (as in polynomial approximation) by the rapid convergence of the term \( \varepsilon_M^*(\mu) \).

The problem of locally non-affine dependence of \( g(x, \mu) \) on parameter \( \mu \) is also studied in Solodukhov [145].

### 6.3 The Stokes problem with non-affine parametric dependence

The parametrized Stokes equations have been formulated in Section 5.2 in a domain \( \hat{\Omega} \in \mathbb{R}^2 \) whose shape is depending affinely on a set of geometrical parameters in each subdomain \( \hat{\Omega}^r \) such that \( \hat{\Omega} = \bigcup_{r=1}^{R} \hat{\Omega}^r \).

In this section we build a system of \( P^2DEs \) (Parametrized Partial Differential Equations) depending on a set of geometrical parameters (\( \mu \)) as coefficients whose dependence from parameters is both affine and non-affine (in different subdomains). As already seen, problem (5.2.7) can be traced back to a reference domain, in this case by an affine mapping on some subdomains \( \hat{\Omega}^G_r \) into \( \hat{\Omega}^G \) and by a non-affine mapping on the remaining subdomains \( \hat{\Omega}^T_r \) into \( \hat{\Omega}^T \). In this case \( R = R_G + R_T \) and \( \Omega = \bigcup_{r=1}^{R_G} \Omega^G_r \cup \bigcup_{r=1}^{R_T} \Omega^T_r \). For any \( \hat{x} \in \hat{\Omega}^G_r, r = 1, \ldots, R_G \), its image \( x \in \Omega^G_r \) is given by

\[
x = G^G_r(\mu; \hat{x}) = G^G_r(\mu) \hat{x} + g^G_r, \quad 1 \leq r \leq R_G;
\]

we thus write on \( \Omega^G_r \)

\[
\frac{\partial}{\partial \hat{x}_i} \frac{\partial}{\partial x_j} = G^G_r(\mu) \frac{\partial}{\partial x_j}, \quad 1 \leq r \leq R_G.
\]

For any \( \hat{x} \in \hat{\Omega}^T_r, r = 1, \ldots, R_T \), its image \( x \in \Omega^T_r \) is given by

\[
x = T^T_r(\mu; \hat{x}), \quad 1 \leq r \leq R_T;
\]
we thus write on $\Omega_T$:

$$\frac{\partial}{\partial x_i} = \frac{\partial x_j}{\partial x_i} \frac{\partial}{\partial x_j} = T_{ij}(\mu, x) \frac{\partial}{\partial x_j},$$  \hspace{1cm} (6.3.4)

Referring to problem (5.2.7) and to formulation from (5.2.10) to (5.2.13), we can write with
the current parametrization in the reference domain $\Omega$:

$$\langle Au, w \rangle = \sum_{r=1}^{R_G} \int_{\Omega_T} \frac{\partial u}{\partial x_i} (G_{ir}^r(\mu) \nu_{ijr}^r G_{jir}^r(\mu) \det(G^r(\mu))^{-1}) \frac{\partial w}{\partial x_j} \, d\Omega +$$

$$\sum_{r=1}^{R_T} \int_{\Omega_T} \frac{\partial u}{\partial x_i} (T_{ir}^r(\mu, x) \nu_{ijr}^r T_{jir}^r(\mu, x) \det(T^r(\mu, x))^{-1}) \frac{\partial w}{\partial x_j} \, d\Omega \forall w \in Y,$$

$$\langle Bp, w \rangle = - \sum_{r=1}^{R_G} \int_{\Omega_T} p (G_{ijr}(\mu) \det(G^r(\mu))^{-1}) \frac{\partial w}{\partial x_i} \, d\Omega +$$

$$- \sum_{r=1}^{R_T} \int_{\Omega_T} p (T_{ijr}(\mu, x) \det(T^r(\mu, x))^{-1}) \frac{\partial w}{\partial x_i} \, d\Omega \forall w \in Y,$$

$$\langle F, w \rangle = \langle F_s, w \rangle + \langle F^0, w \rangle,$$

where

$$\langle F_s, w \rangle = \sum_{r=1}^{R_G} \int_{\Omega_T} \left( \mathbf{f}^r \det(G^r(\mu))^{-1} \right) w \, d\Omega + \sum_{r=1}^{R_T} \int_{\Omega_T} \left( \mathbf{f}^r \det(T^r(\mu, x))^{-1} \right) w \, d\Omega;$$

$$\langle F^0, w \rangle = - \langle A g_m, w \rangle; \hspace{0.5cm} \langle G^0, q \rangle = \langle B q, g_{im} \rangle.$$

In Section 6.5 we introduce two different parametrizations based on different test cases and we show their explicit forms. The transformation tensors for bilinear forms with affine and non-affine mappings are defined, respectively, as follows:

$$\nu_{G_{ijr}}^r(\mu) = G_{ir}^r(\mu) \nu_{ijr} G_{jir}^r(\mu) \det(G^r(\mu))^{-1}, 1 \leq i, j \leq 2, r = 1, ..., R_G,$$

$$\nu_{T_{ijr}}^r(\mu, x) = T_{ir}^r(\mu, x) \nu_{ijr} T_{jir}^r(\mu, x) \det(T^r(\mu, x))^{-1}, 1 \leq i, j \leq 2, r = 1, ..., R_T,$$

The tensors for pressure and divergence forms are defined, respectively, for affine and non-affine mappings as:

$$\chi_{G_{ijr}}^r(\mu) = G_{ijr}^r \det(G^r(\mu))^{-1},$$

$$\chi_{T_{ijr}}^r(\mu, x) = T_{ijr}^r(\mu, x) \det(T^r(\mu, x))^{-1},$$

For the non-affine parts we apply the empirical interpolation procedure of Section 6.2 to expand non-affine mapping terms and decouple the parameters dependent contribution from the one depending only on spacial coordinates. We write:

$$\nu_{T_{ijr}}^r(\mu, x) = \sum_{m=1}^{M_{ijr}} \beta_{ijm}^r(\mu) \chi_{ijm}^r(x),$$

$$\chi_{T_{ijr}}^r(\mu, x) = \sum_{m=1}^{M_{ijr}} \alpha_{ijm}^r(\mu) \omega_{ijm}^r(x).$$
Furthermore, we may define

\[ \Theta^{(i,j,r)}(\mu) = \nu_{G,i,j}^r(\mu), \quad A_G^{(i,j,r)}(u, w) = \int_{\Omega_G} \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_j} d\Omega, \]  
\[ \Phi^{(i,j,r)}(\mu) = \chi_{G,i,j}^r(\mu), \quad B_G^{(i,j,r)}(p, w) = -\int_{\Omega_G} p \frac{\partial w}{\partial x_j} d\Omega, \]  
for \( 1 \leq r \leq R_G, 1 \leq i, j \leq d = 2 \) (\( q \) and \( s \) are condensed indexes of \( i, j, r \) quantities) and

\[ \Psi^{(i,j,r,m)}(\mu) = \beta_{ij,m}^r(\mu), \quad A_T^{(i,j,r,m)}(\gamma(x), u, w) = \int_{\Omega_T} \gamma_{ij,m}^r(x) \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_j} d\Omega, \]  
\[ \Psi^{(i,j,r,m)}(\mu) = \alpha_{ij,m}^r(\mu), \quad B_T^{(i,j,r,m)}(\omega(x), p, w) = -\int_{\Omega_T} \omega_{ij,m}^r(x) p \frac{\partial w}{\partial x_j} d\Omega, \]  
for \( 1 \leq r \leq R_T, 1 \leq i, j \leq d = 2, 1 \leq m \leq \max(M^a_{ijr}, M^b_{ijr}) \) (\( t \) and \( p \) are condensed indexes of \( i, j, r, m \) quantities). We apply an effectively affine decomposition:

\[ A(\mu, u, w) = \sum_{q=1}^{Q_G^d} \Theta^{q}(\mu) A_G^{q}(u, w) + \sum_{t=1}^{Q_T^d} \Psi^{t}(\mu) A_T^{t}(\gamma(x), u, w); \]
\[ B(\mu, p, w) = \sum_{s=1}^{Q_G^b} \Phi^{s}(\mu) B_G^{s}(p, w) + \sum_{p=1}^{Q_T^b} \Psi^{p}(\mu) B_T^{p}(\omega(x), p, w); \]
in general \( \max(Q_G^d) = d \times d \times d \times R_G, \max(Q_G^b) = d \times d \times R_G, Q_G^d = \Sigma_{j=1}^d \Sigma_{i=1}^d \Sigma_{r=1}^{R_G} M^a_{ijr} \) and \( Q_T^d = \Sigma_{j=1}^d \Sigma_{r=1}^{R_T} M^a_{ijr} \) and \( Q_T^b = \Sigma_{j=1}^d \Sigma_{r=1}^{R_T} M^b_{ijr} \). The Stokes problem rewritten on the reference domain \( \Omega \) reads:

\begin{align*}
\left\{ \begin{array}{ll}
A(\mu; u(\mu), w) + B(\mu; p(\mu), w) = (F, w) & \forall \, w \in Y, \\
B(\mu; q, u(\mu)) = (G^0, q) & \forall \, q \in Q.
\end{array} \right.
\end{align*}

We recall that this problem has an inf-sup condition (LBB) [127] to be guaranteed:

\[ \beta(\mu) = \inf_{q \in Q} \sup_{w \in Y} \frac{B(\mu; q, w)}{\|w\|_Y \|q\|_Q} \geq \beta_0 > 0, \forall \, \mu \in D; \]

here \( Y = H^1_0 \times H^1_0 \).

We re-introduce the supremizer operator \( T^\mu: Q \to Y \) so that

\[ (T^\mu q, w)_Y = B(\mu; q, w), \quad \forall \, w \in Y \]

and

\[ T^\mu q = \arg\sup_{w \in Y} \frac{B(\mu; q, w)}{\|w\|_Y}. \]
Then
\[ \beta^2(\mu) = \inf_{q \in Q} \frac{(T^\mu q, T^\mu q)_Y}{\|q\|_Q^2}. \]

For the proofs see the previous chapter. At this step we can solve the Stokes problem by Galerkin-Finite Element Method using Taylor-Hood $\mathbb{P}^2-\mathbb{P}^1$ elements for velocity and pressure, respectively. See Girault and Raviart [43], Gresho and Sani [46], and Gunzburger [48].

### 6.4 The Stokes reduced basis formulation with non-affine parameters

As seen, in the reduced basis approximation we construct a set of \( \mu \) parameters samples
\[ S_N^\mu = \{ \mu^1, \ldots, \mu^N \}, \]
where \( \mu^n \in \mathcal{D}, \ n = 1, \ldots, N. \)

Correspondingly, we define a set of couples \( (u(\mu^n), p(\mu^n)) \) which are approximate solutions of the Stokes problem. Then the reduced basis pressure space is \( Q_N = \text{span}\{\xi_n, n = 1, \ldots, N\} \), while the reduced basis velocity space is \( Y_N = \text{span}\{T^\mu \xi_n, n = 1, \ldots, N\} \).

The problem in reduced basis approximation reads: find \( (u_N(\mu), p_N(\mu)) \in Y_N \times Q_N \) s.t.

\[
\begin{aligned}
\mathcal{A}(\mu; u_N(\mu), w) + \mathcal{B}(\mu; p_N(\mu), w) &= (F, w) \ \forall \ w \in Y_N, \\
\mathcal{B}(\mu; q, u_N(\mu)) &= (G^0, q) \ \forall \ q \in Q_N.
\end{aligned}
\]

This problem is well posed if it does admit an inf-sup property. We introduce

\[ \beta_N(\mu) = \inf_{q \in Q_N} \sup_{w \in Y_N} \frac{\mathcal{B}(\mu, q, w)}{\|w\|_Y \|q\|_Q}, \]

where \( \beta(\mu) \) is the inf-sup constant related to Galerkin-Finite Element Method. We rewrite for computational convenience \( Y_N \) using the effectively affine dependence of \( \mathcal{B}(\mu, q, w) \) on the parameter and the linearity of \( T^\mu \):

\[ T^\mu \xi = \sum_{q=1}^{Q^G} \Phi^q(\mu) T^G_q \xi + \sum_{p=1}^{Q^T} \Psi^p(\mu) T^T_p \xi \]

(6.4.2)

for any \( \xi \) and \( \mu \), where:

\[ (T^G_q \xi, w)_Y = \mathcal{B}^G_q(q, w), \ \forall \ w \in Y, \]

\[ (T^T_p \xi, w)_Y = \mathcal{B}^T_p(\omega, q, w), \ \forall \ w \in Y, \]

which allows us to write:

\[ Y_N = \text{span}\{\sum_{k=1}^{Q^G} \Phi^k(\mu^n) \sigma_{kn} + \sum_{k'=1}^{Q^T} \Psi^{k'}(\mu^n) \sigma_{k'n}, n = 1, \ldots, 2N\}, \]

where \( Q^G = Q^G + 1, \Phi^{Q^G} = 1, \mu^{N+j} = \mu^j, 1 \leq j \leq N. \)

For \( n = 1, \ldots, N: \)

\[ \sigma_{kn} = 0, \text{ for } k = 1, \ldots, Q^G; \]
For $n = N + 1, \ldots, 2N$:

\[
(\sigma_{kn}, w)_Y = B^k_G(\xi_{n-N}, w), \forall w \in Y, \text{for } k = 1, \ldots, Q^b_G;
\]

\[
\sigma_{Q^b_n} = u(\mu^n).
\]

\[
(\tilde{\sigma}_{kn}, w)_Y = B^k_T(\omega, \xi_{n-N}, w), \forall w \in Y, \text{for } k = 1, \ldots, Q^b_T.
\]

For a new "$\mu$" we want a solution given by a combination of previously computed stored solutions as basis functions:

\[
u_N(\mu) = \sum_{j=1}^{2N} u_{Nj}(\mu)\left(\sum_{k=1}^{Q^b_G} \Phi^k(\mu^i)\sigma_{kj} + \sum_{k'=1}^{Q^b_T} \Upsilon^{k'}(\mu^j)\tilde{\sigma}_{k'j}\right),
\]

\[
p_N(\mu) = \sum_{l=1}^{N} p_{Nl}(\mu)\xi_l,
\]

whose weights $u_{Nj}$ and $p_{Nl}$ are given by the following reduced basis linear system:

\[
\begin{cases}
\sum_{j=1}^{2N} A^\mu_{ij} u_{Nj}(\mu) + \sum_{i=1}^{N} B^\mu_{il} p_{Nl}(\mu) = F_i, & 1 \leq i \leq 2N, \\
\sum_{j=1}^{2N} B^\mu_{lj} u_{Nj}(\mu) = G_l, & 1 \leq l \leq N;
\end{cases}
\]

(6.4.5)

where the sub-matrices $A$ and $B$ are built by online-offline computational decoupled procedure:

\[
A^\mu_{ij} = \sum_{k=1}^{Q^b_G} \Theta^k(\mu)A^k_G\left(\sum_{k'=1}^{Q^b_T} \Phi^{k'}(\mu^j)\sigma_{kj} + \sum_{k''=1}^{Q^b_T} \Phi^{k''}(\mu^j)\tilde{\sigma}_{k''j}\right) +
\]

\[
+ \sum_{k=1}^{Q^b_T} \Phi^k(\mu)A^k_T(\gamma, \sum_{k'=1}^{Q^b_T} \Upsilon^{k'}(\mu^j)\tilde{\sigma}_{k'j}, \sum_{k''=1}^{Q^b_T} \Upsilon^{k''}(\mu^j)\tilde{\sigma}_{k''j}), & 1 \leq i, j \leq 2N;
\]

\[
B^\mu_{il} = \sum_{k=1}^{Q^b_G} \Phi^k(\mu)B^k_G\left(\sum_{k'=1}^{Q^b_T} \Phi^{k'}(\mu^i)\sigma_{k'i} + \sum_{k''=1}^{Q^b_T} \Phi^{k''}(\mu^i)\tilde{\sigma}_{k''i}\right) +
\]

\[
+ \sum_{k=1}^{Q^b_T} \Upsilon^k(\mu)B^k_T(\omega, \sum_{k'=1}^{Q^b_T} \Upsilon^{k'}(\mu^i)\tilde{\sigma}_{k'j}, \xi_l), & 1 \leq i \leq 2N, 1 \leq l \leq N;
\]

where $A^k_T, A^k_G, B^k_T$ and $B^k_G$ are assembled offline, and:

\[
F_i = \langle F, \sum_{k=1}^{Q^b_T} \Phi^{k'}(\mu^i)\tilde{\sigma}_{k'j}\rangle + \langle F, \sum_{k'=1}^{Q^b_T} \Upsilon^{k'}(\mu^j)\tilde{\sigma}_{k'j}\rangle, & 1 \leq i \leq 2N;
\]

\[
G_l = \langle G^d, \xi_l\rangle, & 1 \leq l \leq N.
\]
System (6.4.5) can therefore be written as:

\[
\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u}_N \\ p_N \end{pmatrix} = \begin{pmatrix} \mathbf{F} \\ \mathbf{G} \end{pmatrix} \quad (6.4.6)
\]

Remark 6.4.1 We recall that more options dealing with the supremizer calculation and basis assembling procedures are available. At this step we have adopted, for simplicity, the second alternative option introduced in the previous chapter (and in [111]), where all the velocity reduced basis functions are \( \mu \) (on-line) independent (they depend only on samples \( \mu \) used to store basis). By using this option we cannot demonstrate the validity of the stability condition \( \beta_N(\mu) \geq \beta(\mu) \geq \beta_0 > 0, \forall \mu \in \mathcal{D}^\mu \), even if results guarantee the stability of approximation. This option is competitive concerning computational costs dealing with \( 3N \times 3N \) reduced basis matrices (6.4.6) instead of \( (Q^b_{G}+1)N \times (Q^b_{T}+1)N \) matrix (usually \( (Q^b_{G}+1) \gg 3, Q^b_{G} = Q^b_{G} + Q^b_{T} \)), introduced as the first option to build reduced basis velocity space without the use of (6.4.2) and to apply orthonormalization, if needed. Also in this case first option is better, even if more expensive.

Remark 6.4.2 We have the following computational costs to build reduced basis matrices, given also the supremizer components in the velocity space: \( O(Q^a 4N^2) \) for sub-matrix \( A \), \( O(Q^2 2N^2) \) for \( B \), \( O(N) \) for \( F \) and \( O(9N^3) \) for the inversion of the full reduced basis matrix (6.4.6), where \( Q^a = Q^a + Q^a_{G} \), \( Q^b_{G} = Q^b_{G} + Q^b_{T} \). Note that the quantities \( Q^a_{G} \) and \( Q^b_{G} \) are depending only on the number of subdomains with affine mappings (\( R_{G} \)), while quantities \( Q^b_{T} \) and \( Q^b_{T} \) are depending also on the number of “shape functions” (\( \gamma(x) \) and \( \omega(x) \)) related with interpolation error (\( \varepsilon_{\text{max}} \)) and the number of subdomains with non-affine mappings (\( R_{T} \)).

6.5 Numerical results

Numerical tests were carried out to develop Stokes reduced basis techniques with affine/non-affine mappings. Taylor-Hood finite elements have been used to store approximation basis functions: \( P^2 - P^1 \) elements for velocity (with supremizer) and pressure, respectively [127]. The reduced basis solutions have been compared directly with approximate finite element solutions by computing the \( H^1 \) relative error for velocity and the \( L^2 \) relative error for pressure.

6.5.1 First test: curved upper wall

In this first test we apply empirical interpolation to describe a channel with a curved and parametrized upper wall. This simple geometry (see Figure 6.1) can be considered for example in the study of blood flow through an artery interested by a stenosis, i.e. featuring a reduced section (see Quarteroni and Formaggia [126]). A similar geometry can be used also in periodical series to set up and study an oxigenator for haemodinamic applications. In our application we have made the following assumptions:
To solve the parametrized Stokes problem in the domain outlined in Figure 6.1 we have imposed zero Dirichlet conditions on the boundary \( \Gamma_D \), Neumann non-homogeneous conditions on the inflow \( \Gamma_{N_1} \) (\( \tau_{\hat{n}} = 1, \tau_{\hat{t}} = 0 \), where \( \tau = (\nu \frac{\partial u}{\partial n} - p\hat{n}) \), with \( \hat{n} \) and \( \hat{t} \) normal and tangential directions, respectively) and Neumann homogeneous conditions on outflow \( \Gamma_{N_0} \) (\( \tau_{\hat{n}} = 0, \tau_{\hat{t}} = 0 \)).

We consider one parameter \( \mu \) ranging in the interval \([-0.8, 0.8]\) to describe the upper arterial wall in the physical domain, through \( \hat{x}_2 = f(\hat{x}_1, \mu) = 1 + \mu \sin(2\pi \hat{x}_1) \) (we have a single domain subject to an unique non-affine mapping). Referring to Section 6.3 we have \( \Omega_G = \emptyset \) and \( \Omega_T = \Omega \), so \( R_T = 1 \).

The coordinate transformation is \( T : \hat{\Omega} \rightarrow \Omega \) such as \( x = T(\hat{x}) \) and:

\[
(x_1, x_2) = T(\hat{x}_1, \hat{x}_2) = (\hat{x}_1, \frac{1}{f(\hat{x}_1, \mu)} \hat{x}_2)
\]  

(6.5.1)

in \( \Omega \). Then,

\[
d\hat{x}_1 d\hat{x}_2 = f(x_1, \mu) dx_1 dx_2
\]

and the following relations hold:

\[
\begin{aligned}
\frac{\partial \hat{u}_1}{\partial x_1} &= \frac{1}{f(x_1, \mu) \frac{\partial x_1}{\partial \hat{x}_1}}, \\
\frac{\partial \hat{u}_1}{\partial x_2} &= \frac{\partial \phi}{\partial x_1} - x_2 \frac{f_x(\hat{x}_1, \mu)}{f(x_1, \mu)} \frac{\partial u_1}{\partial x_2} \quad \text{(with } f_x := \frac{\partial f}{\partial x_1}),
\end{aligned}
\]

(6.5.2)

\[
\nabla \cdot \hat{\mathbf{u}} = \frac{\partial u_1}{\partial x_1} - x_2 \frac{f_x(\hat{x}_1, \mu)}{f(x_1, \mu)} \frac{\partial u_1}{\partial x_2} + \frac{1}{f(x_1, \mu)} \frac{\partial u_2}{\partial x_2}.
\]

(6.5.3)

Using the compact notation of Section 6.3 (6.3.10 and 6.3.12) and transformation (6.5.1) we get the following tensor for diffusion and divergence forms, respectively:

\[
\nu_T = \nu \begin{bmatrix}
 f(x_1, \mu) & -f_x^2(\hat{x}_1, \mu) x_2 \\ -f_x(\hat{x}_1, \mu) x_2 & 1 + \frac{f_x^2(\hat{x}_1, \mu) x_2^2}{f(x_1, \mu)}
\end{bmatrix};
\]

(6.5.4)

\[
\chi_T = \begin{bmatrix}
 f(x_1, \mu) & -f_x(\hat{x}_1, \mu) x_2 \\ 0 & 1
\end{bmatrix};
\]

(6.5.5)

where \( \nu = 0.04 N s m^{-2} \) is the viscosity. Referring to notation of Section 6.2 we get 5 different coefficients functions \( g^j_M(x, \mu) \) to expand.

We impose empirical interpolation (6.3.13 and 6.3.14) to the tensors (6.5.4 and 6.5.5) and we impose a maximum interpolation error \( \varepsilon_{max} \), thus considering different \( M_{max} \) “shape functions” for each \( g^j_M(x, \mu) \). Each \( g^j_M(x, \mu) \) refers to a different coefficient of a bilinear form of our Stokes problem (\( j = 5 \) in this test case). Owing to empirical interpolation we expand each tensor component to apply the effectively affine decomposition:

\[
\nu_T = \nu \begin{bmatrix}
 \Sigma_{m=1}^{M_1} \beta_{11m}(\mu) \gamma_{11m}(x) & \Sigma_{m=1}^{M_2} \beta_{12m}(\mu) \gamma_{12m}(x) \\
\Sigma_{m=1}^{M_1} \beta_{21m}(\mu) \gamma_{21m}(x) & \Sigma_{m=1}^{M_2} \beta_{22m}(\mu) \gamma_{22m}(x)
\end{bmatrix}
\]
Note that this tensor is symmetric. Moreover,
\[
\chi_T = \begin{bmatrix}
\sum_{m=1}^{M_1} \alpha_{11m}(\mu) \omega_{11m}(x) & \sum_{m=1}^{M_2} \alpha_{12m}(\mu) \omega_{12m}(x) \\
0 & 1
\end{bmatrix}.
\]
The index \( r \) referring to every different subdomain is omitted.

- At this step, having defined all the abstract formulation of the previous sections, we have applied the reduced basis method and assembled the approximation spaces as described in Section 6.4.

- Tables 6.1 and 6.2 (and Figures 6.3 and 6.4) show numerical results (mean \( H^1 \) relative errors on velocity, testing a large number of configurations) at different \( N \) and different max interpolation error \( \varepsilon_{\text{max}} \). At the end of the test we have carried out a comparison between magic points interpolation (using \( g^j_M(x, \mu) \)) and true functions (\( g^j(x, \mu) \)). We can see that for \( \varepsilon_{\text{max}} \leq 10^{-4} \) we have a good convergence and results are not dominated or influenced by interpolation error. When the interpolation error is dominating the reduced basis error is characterized by a constant “plateau” and is not diminished by increasing \( N \) (see for example the case in which \( \varepsilon_{\text{max}} \geq 10^{-2} \)).

![Geometrical scheme for the stenosis test problem (left) and velocity absolute value [ms^{-1}] for \( \mu = 0.5 \) (right).](image)

Figure 6.1: Geometrical scheme for the stenosis test problem (left) and velocity absolute value [ms^{-1}] for \( \mu = 0.5 \) (right).

In Figure 6.5 we represent a possible configuration of interest for biomedical applications, for example an oxygenator. It is possible to understand how the geometry can be described using a periodic criterium by assembling a more complex configuration starting from a simple subdomain. The framework should be completed addressing our attention to periodic boundary conditions. This configuration has been studied also by Wei et al. in [163] considering a wavy-walled channel with a poroelastic layer to investigate physiological flows in capillaries and venules.


\section*{\textit{\textsection 6.5. NUMERICAL RESULTS}}

Figure 6.2: Horizontal and vertical velocity [m/s] for $\mu = 0.5$.

Figure 6.3: $H^1$ relative mean error on velocity for different $\varepsilon_{max}$, imposed on all $g^T_M(x, \mu)$, considering different $\mu$ test values.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\varepsilon_{max} = 0.1$</th>
<th>$\varepsilon_{max} = 0.01$</th>
<th>$\varepsilon_{max} = 10^{-3}$</th>
<th>$\varepsilon_{max} = 10^{-4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$2.2549e+000$</td>
<td>$4.2557e-001$</td>
<td>$4.2656e-001$</td>
<td>$4.2653e-001$</td>
</tr>
<tr>
<td>2</td>
<td>$2.1926e+000$</td>
<td>$1.3614e-001$</td>
<td>$1.3708e-001$</td>
<td>$1.3709e-001$</td>
</tr>
<tr>
<td>3</td>
<td>$2.1361e+000$</td>
<td>$2.1768e-002$</td>
<td>$2.1880e-002$</td>
<td>$2.1886e-002$</td>
</tr>
<tr>
<td>4</td>
<td>$2.1267e+000$</td>
<td>$1.5939e-002$</td>
<td>$1.5963e-002$</td>
<td>$1.5964e-002$</td>
</tr>
<tr>
<td>5</td>
<td>$2.1226e+000$</td>
<td>$4.9955e-003$</td>
<td>$4.9358e-003$</td>
<td>$4.9388e-003$</td>
</tr>
<tr>
<td>6</td>
<td>$2.1172e+000$</td>
<td>$3.5905e-004$</td>
<td>$1.9177e-004$</td>
<td>$1.9183e-004$</td>
</tr>
<tr>
<td>7</td>
<td>$2.1131e+000$</td>
<td>$2.4352e-004$</td>
<td>$6.7685e-005$</td>
<td>$6.7584e-005$</td>
</tr>
<tr>
<td>8</td>
<td>$2.1080e+000$</td>
<td>$1.8644e-004$</td>
<td>$8.2645e-006$</td>
<td>$8.1374e-006$</td>
</tr>
<tr>
<td>9</td>
<td>$1.7357e+000$</td>
<td>$1.8016e-004$</td>
<td>$1.9044e-006$</td>
<td>$1.7880e-006$</td>
</tr>
<tr>
<td>10</td>
<td>$1.7043e+000$</td>
<td>$1.7929e-004$</td>
<td>$9.0032e-007$</td>
<td>$7.7022e-007$</td>
</tr>
<tr>
<td>11</td>
<td>$8.8174e-001$</td>
<td>$1.7868e-004$</td>
<td>$1.5746e-007$</td>
<td>$2.6138e-008$</td>
</tr>
<tr>
<td>12</td>
<td>$3.8732e-001$</td>
<td>$1.7743e-004$</td>
<td>$1.5181e-007$</td>
<td>$1.8676e-008$</td>
</tr>
</tbody>
</table>

Table 6.1: Mean $H^1$ velocity relative errors for different $\varepsilon_{max}$. 
Figure 6.4: $H^1$ mean relative error on velocity for different $\varepsilon_{\text{max}}$, imposed on all $g^j_{\text{max}}(x, \mu)$ (left) and comparison with the use of “true” $g^j(x, \mu)$ functions, without empirical interpolation (right).

<table>
<thead>
<tr>
<th>N</th>
<th>$\varepsilon_{\text{max}} = 10^{-5}$</th>
<th>$\varepsilon_{\text{max}} = 10^{-6}$</th>
<th>exact $g^j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.2654e – 001</td>
<td>4.2654e – 001</td>
<td>4.2654e – 001</td>
</tr>
<tr>
<td>2</td>
<td>1.3709e – 001</td>
<td>1.3709e – 001</td>
<td>1.3709e – 001</td>
</tr>
<tr>
<td>3</td>
<td>2.1885e – 002</td>
<td>2.1884e – 002</td>
<td>2.1884e – 002</td>
</tr>
<tr>
<td>4</td>
<td>1.5963e – 002</td>
<td>1.5963e – 002</td>
<td>1.5963e – 002</td>
</tr>
<tr>
<td>5</td>
<td>4.9380e – 003</td>
<td>4.9380e – 003</td>
<td>4.9380e – 003</td>
</tr>
<tr>
<td>6</td>
<td>1.9180e – 004</td>
<td>1.9181e – 004</td>
<td>1.9181e – 004</td>
</tr>
<tr>
<td>7</td>
<td>6.7581e – 005</td>
<td>6.7582e – 005</td>
<td>6.7582e – 005</td>
</tr>
<tr>
<td>8</td>
<td>8.1348e – 006</td>
<td>8.1356e – 006</td>
<td>8.1356e – 006</td>
</tr>
<tr>
<td>9</td>
<td>1.7812e – 006</td>
<td>1.7813e – 006</td>
<td>1.7813e – 006</td>
</tr>
<tr>
<td>10</td>
<td>7.6544e – 007</td>
<td>7.6542e – 007</td>
<td>7.6542e – 007</td>
</tr>
<tr>
<td>11</td>
<td>2.2070e – 008</td>
<td>2.1981e – 008</td>
<td>2.1980e – 008</td>
</tr>
<tr>
<td>12</td>
<td>1.4503e – 008</td>
<td>1.4421e – 008</td>
<td>1.4421e – 008</td>
</tr>
</tbody>
</table>

Table 6.2: Mean $H^1$ velocity relative errors for different $\varepsilon_{\text{max}}$ and comparison with “true” $g^j.$
6.5.2 Second test: bypass with curved incoming branch

We introduce a vector of parameters \( \mu = \{t, D, L, S, H, \theta, v\} \in \mathcal{D}^\mu \subseteq \mathbb{R}^7 \), \( \mathcal{D}^\mu \) is given by:

\[
[t_{\text{min}}, t_{\text{max}}] \times [D_{\text{min}}, D_{\text{max}}] \times [L_{\text{min}}, L_{\text{max}}] \times [S_{\text{min}}, S_{\text{max}}] \times [H_{\text{min}}, H_{\text{max}}] \times [\theta_{\text{min}}, \theta_{\text{max}}] \times [v_{\text{min}}, v_{\text{max}}].
\]

For a schematic view of the problem see Figure 6.6. The new element introduced in the study of the complete configuration is the curvature of the artery and of the bypass wall to deal with a more realistic configuration. In this problem we have affine and non-affine parameters dependence in different subdomains. The aim of the test is to combine the study of affine and the non-affine terms in the same problem by varying different geometrical parameters and then to test reduced basis convergence. Referring to notation in Section 6.3 we have \( R_G = 3 \) (number of subdomains with affine dependence, i.e. \( \Omega^2, \Omega^3, \Omega^4 \)), \( R_T = 1 \) (number of subdomains with non-affine dependence, i.e \( \Omega^1 \)). For more info on this applications dealing only with affine parameters dependence we suggest to see [137]. The coordinate transformation in \( \Omega^1 \) with non-affine parameter dependence is given by:

\[
\begin{align*}
    x_1 &= \frac{1}{\pi} \tilde{x}_1 \\
    x_2 &= \frac{1}{\tau} (\tilde{x}_2 - (vH^2x_1(x_1 - 1) + Hx_1 \tan(\theta)) )
\end{align*}
\]

(6.5.6)
The role of parameters \( t \) and \( H \) is to stretch subdomain \( \Omega^1 \) (as \( L, S, D \) stretch the remaining subdomains), the parameter \( v \) introduces a curvature in the walls of the incoming branch of the bypass and \( \theta \) is responsible for a rigid rotation varying the graft angle. The tensors for *viscous terms* are given by:

\[
\nu_T^1 = \nu \left[ \begin{array}{c} \frac{t}{\pi} \\ -\tan(\theta + 2tHx_1 - vH) \\ \frac{1}{(1+\tan(\theta + 2tHx_1 - vH)^2)H} \end{array} \right]; \quad (6.5.7)
\]

\[
\nu_G^1 = \nu \left[ \begin{array}{c} S \\ 0 \\ D \end{array} \right]; \quad \nu_G^2 = \nu \left[ \begin{array}{c} L \\ 0 \\ D \end{array} \right]; \quad \nu_G^3 = \nu \left[ \begin{array}{c} L \\ 0 \\ D \end{array} \right]. \quad (6.5.8)
\]

The tensors for *pressure and divergence forms* are given by:

\[
\chi_T^1 = \left[ \begin{array}{cc} t & -H(\tan(\theta + 2tHx_1 - vH)) \\ 0 & H \end{array} \right]; \quad \chi_G^2 = \left[ \begin{array}{c} S \\ 0 \\ D \end{array} \right]; \quad (6.5.9)
\]

\[
\chi_G^3 = \left[ \begin{array}{c} t \\ 0 \\ D \end{array} \right]; \quad \chi_G^4 = \left[ \begin{array}{c} L \\ 0 \\ D \end{array} \right]. \quad (6.5.10)
\]

We have applied empirical interpolation expansion to the components of tensor \( \nu_T^1 \) and \( \chi_T^1 \), and built reduced basis approximation.

**Figure 6.6:** Geometrical scheme for the bypass test problem (physical domain and reference one).

We have carried out four different tests, based on the same geometry but with different parameters:

- **(a)** we consider only \( v \) parameter to create a curvature in the wall, we deal only with non-affine mapping in one subdomain where we apply empirical interpolation;

- **(b)** then, we consider \( v \) and also parameters \( L \) and \( S \), each of them is operating into different subdomains and we combine affine and non-affine transformations;

- **(c)** in the third test we consider parameters \( v, L, S \) and \( t \) so that we have more parameters in the same subdomain subject to non-affine parametric dependence;
Figure 6.7: Haemodinamic flow: absolute value of velocity \([\text{ms}^{-1}, 10^{-2}]\) in bypass subject to curved wall and graft angle \((v = 0.41 \text{ and } \theta = \pi/3)\).

**Figure 6.7:** Haemodinamic flow: absolute value of velocity \([\text{ms}^{-1}, 10^{-2}]\) in bypass subject to curved wall and graft angle \((v = 0.41 \text{ and } \theta = \pi/3)\).

(d) in the last term we introduce also graft angle \(\theta\) so that we have rigid rotation, stretching (due to \(t\) or/and \(H\)) and curvature (due to \(v\)) in the same subdomain. Figure 6.7 shows an example of haemodinamic flow in our curved geometry.

In Tables 6.3-6.6 we report numerical results (\(H^1\) errors on velocity and \(L^2\) errors for pressure) considering about 50 configurations at different \(N\) for different test cases (a)-(d). The maximum interpolation error considered has been \(\varepsilon_{\text{max}} = 10^{-5}\) not to have interpolation error dominating our approximation and to avoid a constant “plateau” in error plots. In Figure 6.8 we report errors in log-log scale dealing with (d) test case. Figure 6.9 shows the offline set of parameters used to store the basis functions by the adaptive optimized assembling procedure we used, introduced in Veroy et al. [161] and Prud’homme [121].

**Figure 6.8:** \(L^2\) and \(H^1\) relative errors for pressure and velocity, respectively (max and mean).
<table>
<thead>
<tr>
<th>N</th>
<th>$H^1$ mean (a)</th>
<th>$H^1$ max (a)</th>
<th>$H^1$ mean (b)</th>
<th>$H^1$ max (b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.048058118</td>
<td>0.136051645</td>
<td>0.066384779</td>
<td>0.10111214</td>
</tr>
<tr>
<td>2</td>
<td>0.001552862</td>
<td>0.002320168</td>
<td>0.039400421</td>
<td>0.098294259</td>
</tr>
<tr>
<td>3</td>
<td>7.1431E-06</td>
<td>1.03781E-05</td>
<td>0.003468971</td>
<td>0.004683999</td>
</tr>
<tr>
<td>4</td>
<td>1.62586E-07</td>
<td>8.24285E-07</td>
<td>0.003163783</td>
<td>0.003657563</td>
</tr>
<tr>
<td>5</td>
<td>6.101E-10</td>
<td>1.41057E-09</td>
<td>0.000858429</td>
<td>0.00231849</td>
</tr>
<tr>
<td>6</td>
<td>7.01E-12</td>
<td>3.212E-11</td>
<td>0.000129567</td>
<td>0.000272338</td>
</tr>
<tr>
<td>7</td>
<td>--</td>
<td>--</td>
<td>5.5189E-05</td>
<td>4.23899E-05</td>
</tr>
<tr>
<td>8</td>
<td>--</td>
<td>--</td>
<td>3.18091E-05</td>
<td>3.83384E-05</td>
</tr>
<tr>
<td>9</td>
<td>--</td>
<td>--</td>
<td>1.23528E-05</td>
<td>2.2659E-05</td>
</tr>
<tr>
<td>10</td>
<td>--</td>
<td>--</td>
<td>7.29802E-06</td>
<td>9.86723E-06</td>
</tr>
<tr>
<td>11</td>
<td>--</td>
<td>--</td>
<td>1.51548E-06</td>
<td>2.30932E-06</td>
</tr>
<tr>
<td>12</td>
<td>--</td>
<td>--</td>
<td>2.34584E-07</td>
<td>5.31751E-07</td>
</tr>
<tr>
<td>13</td>
<td>--</td>
<td>--</td>
<td>1.2129E-07</td>
<td>2.40584E-07</td>
</tr>
<tr>
<td>14</td>
<td>--</td>
<td>--</td>
<td>7.51433E-08</td>
<td>1.18963E-07</td>
</tr>
<tr>
<td>15</td>
<td>--</td>
<td>--</td>
<td>3.23837E-08</td>
<td>9.11491E-08</td>
</tr>
</tbody>
</table>

Table 6.3: $H^1$ velocity relative errors for different number of varying parameters and $\varepsilon_{\text{max}} = 10^{-5}$: (a) only $v$ (non-affine) parameter varying, (b) $v, L, S$ (affine and non-affine) parameters (in different subdomains) varying.

Figure 6.9: Parameters distribution during off-line optimized basis assembling procedures.
### Table 6.4: \( H^1 \) velocity relative errors for different number of varying parameters and \( \varepsilon_{\text{max}} = 10^{-5} \): (c) \( v, L, S, t \) parameters varying, (d) \( v, t, L, S, \theta \) parameters varying.

<table>
<thead>
<tr>
<th>( N )</th>
<th>( H^1 ) mean (c)</th>
<th>( H^1 ) max (c)</th>
<th>( H^1 ) mean (d)</th>
<th>( H^1 ) max (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.563465887</td>
<td>1.106596438</td>
<td>0.487915824</td>
<td>3.904000929</td>
</tr>
<tr>
<td>2</td>
<td>0.155181008</td>
<td>0.426078474</td>
<td>0.302592595</td>
<td>1.385050105</td>
</tr>
<tr>
<td>3</td>
<td>0.023388203</td>
<td>0.062893850</td>
<td>0.243772546</td>
<td>0.511847739</td>
</tr>
<tr>
<td>4</td>
<td>0.069668648</td>
<td>0.013543197</td>
<td>0.032855726</td>
<td>0.067523767</td>
</tr>
<tr>
<td>5</td>
<td>0.00241075</td>
<td>0.007549560</td>
<td>0.018203568</td>
<td>0.048221219</td>
</tr>
<tr>
<td>6</td>
<td>0.0074955</td>
<td>0.009246898</td>
<td>0.035139624</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.000442849</td>
<td>0.001727769</td>
<td>0.007747967</td>
<td>0.022901482</td>
</tr>
<tr>
<td>8</td>
<td>0.00224509</td>
<td>0.00074055</td>
<td>0.001924017</td>
<td>0.000784976</td>
</tr>
<tr>
<td>9</td>
<td>0.57733E-05</td>
<td>0.00360971</td>
<td>0.001338957</td>
<td>0.012582395</td>
</tr>
<tr>
<td>10</td>
<td>0.122527E-05</td>
<td>0.000534446</td>
<td>0.001979371</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.46346E-06</td>
<td>0.000321415</td>
<td>0.001391688</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1.011E-06</td>
<td>0.000207283</td>
<td>0.000936665</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1.91343E-07</td>
<td>0.000123885</td>
<td>0.000930468</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>2.38061E-08</td>
<td>0.000340849</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>—</td>
<td>—</td>
<td>3.74193E-05</td>
<td>0.000260742</td>
</tr>
<tr>
<td>16</td>
<td>—</td>
<td>—</td>
<td>1.41356E-05</td>
<td>8.64251E-05</td>
</tr>
<tr>
<td>17</td>
<td>—</td>
<td>—</td>
<td>8.12437E-06</td>
<td>4.87599E-05</td>
</tr>
<tr>
<td>18</td>
<td>—</td>
<td>—</td>
<td>5.44325E-06</td>
<td>3.614E-05</td>
</tr>
</tbody>
</table>
Table 6.5: $L^2$ pressure relative errors for different number of varying parameters and $\varepsilon_{max} = 10^{-5}$: (a) only $v$ (non-affine) parameter varying, (b) $v, L, S$ (affine and non-affine) parameters (in different subdomains) varying.
### Table 6.6: \(L^2\) pressure relative errors for different number of varying parameters and \(\varepsilon_{\text{max}} = 10^{-5}\): (c) \(v, L, S, t\) parameters varying, (d) \(v, t, L, S, \theta\) parameters varying.

<table>
<thead>
<tr>
<th>(N)</th>
<th>(L^2) mean (c) (\times 10^{-6})</th>
<th>(L^2) max (c) (\times 10^{-6})</th>
<th>(L^2) mean (d) (\times 10^{-6})</th>
<th>(L^2) max (d) (\times 10^{-6})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.395384579</td>
<td>0.710666218</td>
<td>0.145944345</td>
<td>0.240938029</td>
</tr>
<tr>
<td>2</td>
<td>0.006901236</td>
<td>0.013288413</td>
<td>0.051283742</td>
<td>0.141755515</td>
</tr>
<tr>
<td>3</td>
<td>0.004135993</td>
<td>0.00882405</td>
<td>0.022163703</td>
<td>0.075274908</td>
</tr>
<tr>
<td>4</td>
<td>0.00917813</td>
<td>0.00212774</td>
<td>0.00232357</td>
<td>0.006473042</td>
</tr>
<tr>
<td>5</td>
<td>0.000217614</td>
<td>0.000709398</td>
<td>0.00227062</td>
<td>0.004974737</td>
</tr>
<tr>
<td>6</td>
<td>0.000102036</td>
<td>0.000478661</td>
<td>0.000609347</td>
<td>0.001651962</td>
</tr>
<tr>
<td>7</td>
<td>4.40703E-05</td>
<td>0.000189533</td>
<td>0.000577145</td>
<td>0.001413289</td>
</tr>
<tr>
<td>8</td>
<td>2.55188E-05</td>
<td>0.000105753</td>
<td>0.000239322</td>
<td>0.001273091</td>
</tr>
<tr>
<td>9</td>
<td>4.03863E-06</td>
<td>2.22016E-05</td>
<td>0.000122951</td>
<td>0.000623509</td>
</tr>
<tr>
<td>10</td>
<td>5.14632E-07</td>
<td>3.87984E-06</td>
<td>7.2085E-05</td>
<td>0.000511114</td>
</tr>
<tr>
<td>11</td>
<td>3.35225E-07</td>
<td>2.72714E-06</td>
<td>2.59278E-05</td>
<td>9.45921E-05</td>
</tr>
<tr>
<td>12</td>
<td>3.06439E-07</td>
<td>2.68244E-06</td>
<td>2.04926E-05</td>
<td>8.20901E-05</td>
</tr>
<tr>
<td>13</td>
<td>2.11621E-08</td>
<td>1.12451E-07</td>
<td>8.85109E-06</td>
<td>6.95492E-05</td>
</tr>
<tr>
<td>14</td>
<td>3.04801E-09</td>
<td>3.04801E-08</td>
<td>6.46217E-06</td>
<td>5.3356E-05</td>
</tr>
<tr>
<td>15</td>
<td>–</td>
<td>–</td>
<td>2.0293E-06</td>
<td>2.76806E-05</td>
</tr>
<tr>
<td>16</td>
<td>–</td>
<td>–</td>
<td>8.98324E-07</td>
<td>1.00727E-05</td>
</tr>
<tr>
<td>17</td>
<td>–</td>
<td>–</td>
<td>2.92535E-07</td>
<td>1.1717E-06</td>
</tr>
<tr>
<td>18</td>
<td>–</td>
<td>–</td>
<td>1.99739E-07</td>
<td>1.0067E-06</td>
</tr>
</tbody>
</table>
6.6 Conclusions

In this chapter we have shown how to expand and apply efficient, accurate and real-time reduced basis techniques to bio-mechanics problems using more realistic geometries such as in biomedical devices optimization problem and extended previous studies on bypass configuration (with curved walls). Interesting applications of the method presented in this chapter are shape optimization problems in which we may treat both affine and non-affine parametric dependence (see for example Cholaseuk [23], and Lim and Choi [81]): usually the affine parametric dependence could describe complete (macro and coarse) geometrical configuration, while the non-affine parametric dependence could describe local configuration. At this stage we may combine reduced basis techniques with shape design and optimal control related problem such as the ones considered in Chapter 3 and in [4], where shape optimization in reference domain and non-affine mapping have already been considered.

Results have shown that it is possible to consider more complex geometries in optimization problems using reduced basis techniques and introducing empirical, stable and inexpensive interpolation for non-affine mapping terms preserving reduced basis methods properties, first of all rapid and accurate convergence and computational online-offline decomposition. In the next chapter our attention will be devoted to the use of reduced basis for Navier-Stokes equations in parametrized domain. The following chapter is going to conclude the framework on the use of reduced basis for a (pre-process) macro-geometrical optimization on bypass and at the same time it allows us to extend methodology.
Chapter 7

Reduced Basis for Navier-Stokes Equations

In this chapter we introduce an extension of reduced basis methods for non-linear equations in parametrized domains. We deal with Navier-Stokes equations. The main difference with respect to the work described in the previous chapters, dealing with Stokes problem, is the introduction of a non-linear transport term and the extension of reduced basis machinery to this term, including the possibility to consider also non-affine parametric dependence by an empirical interpolation method. The extension to Navier-Stokes equations allows us to have a feedback on the results from pre-process optimization based on complete configuration and on the results of local shape optimization carried out by optimal control techniques using low order (and low fidelity) methods on the proposed application.

7.1 Introduction

Until now, the use of reduced basis in this thesis and its related application was aimed at providing real time information on solutions and fluid mechanics related outputs, computed in parametrized domains by using low fidelity (simplified) model, such as the Stokes equations, used to approximate low Reynolds blood flows and provide an optimized configuration. In this chapter, with the extension of the reduced basis framework to the non-linear steady Navier-Stokes problem, we are going to complete our pre-process optimization toolbox, providing also some comparison to validate our results achieved using low order (and low cost) methods. But we are also going to provide a different perspective: reduced basis methods provide real time solutions for the design and the optimization of systems and can be used not only as a pre-process but as a complete optimization tool in itself. This extension has been possible thanks to the introduction of treatment for non-linearities and non-affine mappings. Another aspect will be considered in the next chapter which provides some perspectives on the use of reduced basis to solve optimal control (and parametrized) problems.

The present chapter is organized as follows: after this introduction, in Section 7.2 we report an excursus of references and aspects dealing with reduced basis for Navier-Stokes, then
Section 7.3 deals with the formulation of the steady Navier-Stokes problem and its related aspects. From Section 7.4 to Section 7.6 we build the reduced basis formulation for Navier-Stokes equations in parametrized domains with (i) affine, (ii) non-affine and (iii) affine/non-affine combined parametric dependence. In these sections aspects dealing with algebraic and approximation stability are recalled and extended. In Sections 7.7, 7.8 and 7.9 numerical results and computational costs and savings are reported. Some conclusions follow.

7.2 The state of the art

The application of reduced basis methods to non-linear problems in fluid mechanics is becoming a problem with growing interest. Previous works were carried out in the 1990s by Peterson [113] and by Ito and Ravindran in [64]. The former work was focused on stream function and vorticity formulation for various Reynolds number and using Taylor and Lagrangian type basis. In the latter the authors built a reduced basis approximation for a Navier-Stokes problem (both steady and unsteady) in a cavity using different techniques to build reduced basis approximation spaces (Hermite and/or Taylor spaces, incorporating the derivatives of velocity with respect to the parameter, the Reynolds number). The reduced basis formulation does not consider the pressure in the reduced-order model. Then Patera and Veroy since 2003 carried out intensive research on reduced basis in non-linear fluid mechanics using Lagrangian basis (i.e. global approximation functions) studying natural convection problem, parametrized by physical quantities (e.g. Grashof number [104], viscosity [160]) considering free-divergence spaces. In 2004 they developed rigorous a posteriori error estimation and bounds for real time computation based on the Brezzi-Rappaz-Raviart theory, see [159]. A review is provided also in the recent work [45]. In our work we have followed the guidelines provided by Patera et al. in their recent works and focused our attention on the following aspects: i) the efficient treatment of non-linear term; ii) the domain parametrization; iii) the incorporation of a stable approximation for pressure (useful in haemodynamic applications); iv) the use of non-affine transformation maps by an empirical interpolation method to consider more complex geometries, such as curved walls. Related aspects deals with the use of different options to build the reduced basis velocity spaces and the efficient offline selection of basis function.

More recently, researchers dealing with flow optimization and using sensitivity analysis have addressed their attention to the reduced basis method, see for example the recent works of Burkardt [21] and Gunzburger [50].
7.3 Steady Navier-Stokes equations

We consider the steady Navier-Stokes equations in a domain $\Omega$ with proper boundary conditions on $\Gamma = \Gamma_{in} \cup \Gamma_{out} \cup \Gamma_w$:

$$
\begin{align*}
-\nu \Delta u + (u \cdot \nabla)u + \nabla p &= f \quad \text{in } \Omega, \\
\nabla \cdot u &= 0 \quad \text{in } \Omega, \\
\mathbf{u} &= \mathbf{0} \quad \text{on } \Gamma_w; \quad \mathbf{u} = \mathbf{g}_{in} \quad \text{on } \Gamma_{in}, \quad \left( \nu \frac{\partial u}{\partial n} - p \mathbf{n} \right) = \mathbf{0} \quad \text{on } \Gamma_{out}.
\end{align*}
$$

(7.3.1)

For the mathematical theory of the Navier-Stokes equations see e.g. Galdi [41] and for aspects dealing with their numerical solution see Temam [150], Girault and Raviart [43], Brezzi and Fortin [19].

Before proceeding we introduce the weak formulation of problem (7.3.1) which reads: find $u \in Y = H^1_D(\Omega) \times H^1_D(\Omega)$, $p \in Q = L^2(\Omega)$, $\Omega \subset \mathbb{R}^2$:

$$
\begin{align*}
\nu \int_{\Omega} \nabla u \cdot \nabla w d\Omega - \int_{\Omega} p \nabla \cdot w d\Omega + \int_{\Omega} (u \cdot \nabla)u \cdot w d\Omega &= \int_{\Omega} f \cdot w d\Omega + \langle F^0, w \rangle \quad \forall w \in Y, \\
\int_{\Omega} q \nabla \cdot u d\Omega &= \langle G^0, q \rangle \quad \forall q \in Q,
\end{align*}
$$

(7.3.2)

$F^0, G^0$ are terms due to non-homogeneous Dirichlet boundary condition ($u = g_{in}$) on $\Gamma_{in}$, $\Gamma_D = \Gamma_{in} \cup \Gamma_w$, on $\Gamma_{out}$ we impose a free stress Neumann condition. Then we discretize problem (7.3.2) by a stable approximation using finite element method (Taylor-Hood $P_2 - P_1$ elements for velocity and pressure, respectively) on a fine mesh triangulation; see, for example, Quarteroni and Valli [127]. The numerical methods used to solve the system of non-linear equations (7.3.2) is the iterative Newton method and it involves the linearization of the advection term $(u \cdot \nabla)u$ considering its derivative in the Frechet sense. The linearized version of the discretized problem (7.3.2) at each iteration reads: given $u_h^{(k)}$, find $u_h^{(k+1)} \in Y_h$, $p_h^{(k+1)} \in Q_h$, $\Omega \subset \mathbb{R}^2$, such that

$$
\begin{align*}
\nu \int_{\Omega} \nabla u_h^{(k+1)} \cdot \nabla w_h d\Omega - \int_{\Omega} p_h^{(k+1)} \nabla \cdot w_h d\Omega + \int_{\Omega} (u_h^{(k)} \cdot \nabla)u_h^{(k+1)} \cdot w_h d\Omega + \\
+ \int_{\Omega} (u_h^{(k+1)} \cdot \nabla)u_h^{(k)} \cdot w_h d\Omega - \int_{\Omega} (u_h^{(k)} \cdot \nabla)u_h^{(k)} \cdot w_h d\Omega &= \int_{\Omega} f \cdot w_h d\Omega + \langle F^0, w_h \rangle \quad \forall w_h \in Y_h, \\
\int_{\Omega} q_h \nabla \cdot u_h^{(k+1)} d\Omega &= \langle G^0, q_h \rangle \quad \forall q_h \in Q_h,
\end{align*}
$$

(7.3.3)

till the convergence is reached. We can adopt a condition based on the difference between two estimates

$$
||u_h^{(k+1)} - u_h^{(k)}|| \leq \epsilon
$$

where $\epsilon$ is a small tolerance. We recall that the convergence for Newton method is quadratic [129]. To solve the linearized Navier-Stokes system (7.3.3) at each Newton step we can use the so-called Pressure Matrix Method to decouple the calculation of pressure from the velocity field.

In this case we solve our problem in three steps: in the first one we get an auxiliary
velocity $\hat{u}_h$, in the second step we can calculate the pressure $p_h$ and then the corrected velocity $u_h$. See for example Quarteroni and Valli [127].

### 7.4 Parametrized formulation: affine parametric dependence

We suppose that the domain we are considering is composed of $R$ subdomains: $\hat{\Omega} = \bigcup_{r=1}^{R} \hat{\Omega}^r$, so that we rewrite (7.3.3) as follows, introducing the “hat” notation to indicate equations in real domain without parametrization and dropping the subscript $h$ for simplicity of notation:

\[
\begin{align*}
\hat{A}(\hat{u}^{(k+1)}, \hat{w}) + \hat{B}(\hat{p}^{(k+1)}, \hat{w}) + \hat{C}(\hat{u}^{(k+1)}, \hat{u}^{(k)}, \hat{w}) + \hat{C}(\hat{u}^{(k)}, \hat{u}^{(k+1)}, \hat{w}) &= \langle \hat{F}, \hat{w} \rangle + \hat{C}(\hat{u}^{(k)}, \hat{u}^{(k)}, \hat{w}) \quad \forall \hat{w} \in Y, \\
-\hat{B}(\hat{q}, \hat{u}^{(k+1)}) &= \langle \hat{G}^0, \hat{q} \rangle \quad \forall \hat{q} \in Q,
\end{align*}
\]

where for $1 \leq i, j \leq d = 2$ and $\hat{\nu}_{i,j} = \nu \delta_{i,j}$:

\[
\hat{A}(\hat{u}, \hat{w}) = \sum_{r=1}^{R} \int_{\Omega^r} \frac{\partial \hat{u}}{\partial \hat{x}_i} \hat{\nu}_{i,j} \frac{\partial \hat{w}}{\partial \hat{x}_j} d\hat{\Omega}, \\
\hat{B}(\hat{p}, \hat{w}) = -\sum_{r=1}^{R} \int_{\Omega^r} \hat{p} \nabla \cdot \hat{w} d\hat{\Omega}, \\
\hat{C}(\hat{u}, \hat{v}, \hat{w}) = \sum_{r=1}^{R} \int_{\Omega^r} (\hat{u} \cdot \nabla) \hat{v} \cdot \hat{w} d\hat{\Omega}, \\
\langle \hat{F}, \hat{w} \rangle = \langle \hat{F}_s, \hat{w} \rangle + \langle \hat{F}^0, \hat{w} \rangle,
\]

and

\[
\langle \hat{F}_s, \hat{w} \rangle = \sum_{r=1}^{R} \int_{\Omega^r} \hat{f} \hat{w} d\hat{\Omega}, \quad \langle \hat{F}^0, \hat{w} \rangle = -\langle \hat{A} \hat{g}_m, \hat{w} \rangle, \quad \langle \hat{G}^0, \hat{q} \rangle = \langle \hat{B} \hat{q}, \hat{g}_m \rangle.
\]

Now we want to build a system of $P^2$DEs (Parametrized Partial Differential Equations) affinely depending on a set of geometrical parameters ($\mu$). As already seen, problem (7.4.1) is traced back to a reference domain by an affine mapping on some subdomains $\hat{\Omega}^r_G$ into $\hat{\Omega}^r$. For any $\hat{x} \in \hat{\Omega}^r$, $r = 1, \ldots, R$, its image $x \in \Omega^r$ is given by

\[
x = G^r(\mu; \hat{x}) = G^r(\mu) \hat{x} + g^r, \quad 1 \leq r \leq R;
\]

we thus write on $\Omega^r$:

\[
\frac{\partial}{\partial \hat{x}_i} = \frac{\partial x_j}{\partial \hat{x}_i} \frac{\partial}{\partial x_j} = G^r_j(\mu) \frac{\partial}{\partial x_j}.
\]

We write the $P^2$DEs system depending on $\mu$ on a reference domain $\Omega \subset \mathbb{R}^2$. The problem reads: given $u^{(k)}$, find $u^{(k+1)} \in Y = H^1_D(\Omega) \times H^1_D(\Omega)$, $p^{(k+1)} \in Q = L^2(\Omega)$:

\[
\begin{align*}
\mathcal{A}(\mu; u^{(k+1)}, w) + \mathcal{B}(\mu; p^{(k+1)}, w) + \mathcal{C}(\mu; u^{(k+1)}, u^{(k)}, w) + \mathcal{C}(\mu; u^{(k)}, u^{(k+1)}, w) &= \langle F_s + F^0, w \rangle + \mathcal{C}(\mu; u^{(k)}, u^{(k)}, w) \quad \forall w \in Y, \\
\mathcal{B}(\mu; q; u^{(k+1)}) &= \langle G^0, q \rangle \quad \forall q \in Q,
\end{align*}
\]

where $\mathcal{A}(\mu; u^{(k+1)}, w)$ is the bilinear form associated to the momentum equation, $\mathcal{B}(\mu; p^{(k+1)}, w)$ is the bilinear form associated to the continuity equation, and $\mathcal{C}(\mu; u^{(k)}, u^{(k+1)}, w)$ is the trilinear form associated to the energy equation.
where:

\[
A(\mu, \mathbf{u}, \mathbf{w}) = \sum_{r=1}^{R} \int_{\Omega_r} \frac{\partial \mathbf{u}}{\partial x_i} \left( G_{ij}^r(\mu) \nu_{ij}^r(\mu) \det(G^r(\mu))^{-1} \right) \frac{\partial \mathbf{w}}{\partial x_j} d\Omega \forall \mathbf{w} \in Y,
\]

\[
B(\mu, p, \mathbf{w}) = - \sum_{r=1}^{R} \int_{\Omega_r} p \left( G_{ij}^r(\mu) \det(G^r(\mu))^{-1} \right) \frac{\partial w_i}{\partial x_j} d\Omega \forall \mathbf{w} \in Y,
\]

\[
C(\mu, \mathbf{u}, \mathbf{v}, \mathbf{w}) = - \sum_{r=1}^{R} \int_{\Omega_r} u_i \left( G_{ij}^r(\mu) \det(G^r(\mu))^{-1} \right) \frac{\partial v_j}{\partial x_i} d\Omega \forall \mathbf{w} \in Y,
\]

\[
\langle F, \mathbf{w} \rangle = \sum_{r=1}^{R} \int_{\Omega_r} \left( \hat{f}^r \det(G^r(\mu))^{-1} \right) \mathbf{w} d\Omega; \quad \langle F^0, \mathbf{w} \rangle = - \langle A g_{in}, \mathbf{w} \rangle;
\]

\[
\langle G^0, q \rangle = \langle B q, g_{in} \rangle;
\]

the general formulation of the tensors for **viscous terms** is

\[
\nu_{ij}^r(\mu) = G_{ij}^r(\mu) \nu_{ij}^r(\mu) \det(G^r(\mu))^{-1}, 1 \leq i, j \leq 2, r = 1, \ldots, R,
\]

(7.4.10)

the one of the tensors for **pressure, divergence and convective terms** is:

\[
\chi_{ij}^r(\mu) = \pi_{ij}^r(\mu) = G_{ij}^r \det(G^r(\mu))^{-1}, 1 \leq i, j \leq 2, r = 1, \ldots, R.
\]

(7.4.11)

Using the affine decomposition property to exploit and decouple the online computational stage (many queries) and the offline one (computed once) we introduce the following elements:

\[
\Theta^{q(i,j,r)}(\mu) = \nu_{ij}^r(\mu), \quad \mathcal{A}^{q(i,j,r)}(\mathbf{u}, \mathbf{w}) = \int_{\Omega_r} \frac{\partial \mathbf{u}}{\partial x_i} \frac{\partial \mathbf{w}}{\partial x_j} d\Omega,
\]

\[
\Phi^{s(i,j,r)}(\mu) = \chi_{ij}^r(\mu), \quad \mathcal{B}^{s(i,j,r)}(p, \mathbf{w}) = - \int_{\Omega_r} p \frac{\partial w_i}{\partial x_j} d\Omega,
\]

\[
\Upsilon^{s(i,j,r)}(\mu) = \pi_{ij}^r(\mu), \quad \mathcal{C}^{s(i,j,r)}(\mathbf{u}, \mathbf{w}) = \int_{\Omega_r} u_i \frac{\partial \mathbf{w}}{\partial x_j} d\Omega,
\]

\[
\mathcal{A}(\Theta(\mu), \mathbf{u}, \mathbf{w}) = \sum_{q=1}^{Q^a} \Theta^{q}(\mu) A(\mathbf{u}, \mathbf{w})^q;
\]

\[
\mathcal{B}(\Phi(\mu), p, \mathbf{w}) = \sum_{s=1}^{Q^b} \Phi^{s}(\mu) B(p, \mathbf{w})^s;
\]

\[
\mathcal{C}(\Upsilon(\mu), \mathbf{u}, \mathbf{w}) = \sum_{s=1}^{Q^c} \Upsilon^{s}(\mu) C(\mathbf{u}, \mathbf{w})^s;
\]

where \(s\) and \(q\) are condensed indexes for \(i, j, r\), \(\max(Q^a) = d \times d \times d \times R\), \(\max(Q^b) = d \times d \times R\) and \(\max(Q^c) = d \times d \times d \times R\). The reduced basis approximation for the Navier-Stokes system...
(7.4.9) reads: find \((u^{(k+1)}_N(\mu), p^{(k+1)}_N(\mu)) \in Y_N \times Q_N\), given \(u^{(k)}_N\) as guess solution for the non-linear solver (taken from the previous iteration or given as \(u^{(0)}_N\)), such that

\[
\begin{aligned}
& A(\mu; u^{(k+1)}_N(\mu), w) + B(\mu; p^{(k+1)}_N(\mu), w) + C(\mu; u^{(k+1)}_N(\mu), u^{(k)}_N(\mu), w) + \\
& C(\mu; u^{(k)}_N(\mu), u^{(k+1)}_N(\mu), w) = \langle F, w \rangle + C(\mu; u^{(k)}_N(\mu), u^{(k)}_N(\mu), w) \forall w \in Y_N, \\
& B(\mu; q, u^{(k+1)}_N(\mu)) = \langle G, q \rangle \forall q \in Q_N.
\end{aligned}
\]

As already seen we properly take some “\(\mu\)” samples \(S^\mu_N = \{\mu^1, \ldots, \mu^N\}\) (snapshots), where \(\mu^n \in D^\mu\), \(n = 1, \ldots, N\).

The reduced basis pressure space is:

\[
Q_N = \text{span} \{\xi_n, n = 1, \ldots, N\},
\]

where \(\xi_n = p(\mu^n)\). We use the supremizer solutions to enrich \(Y_N\), the reduced basis velocity space is given by:

\[
Y_N = \text{span} \{\sigma_n, n = 1, \ldots, 2N\} = \\
= \text{span} \{\zeta_n, T^\mu_n \xi_n, n = 1, \ldots, N\},
\]

where \(\zeta_n = u(\mu^n)\) and \(T^\mu_n: Q \to Y\) is the supremizer operator

\[
(T^\mu_n q, w)_Y = B(\mu; q, w), \quad \forall w \in Y.
\]

We recall that we can write \(T^\mu_n \xi = \sum_{q=1}^Q \Phi^q(\mu^n)T^q \xi\) for any \(\xi\) and \(\mu\).

**Remark 7.4.1** Note that in this case the reduced basis velocity space is not depending on the “online” value of the \(\mu\) parameter because in the supremizer affine assembling procedure we use the values of \(\mu^n\) corresponding at each solution \(\xi_n = p(\mu^n)\), considered as basis function. This option allows us to simplify notation and to reduce complexity of the problem in this first case.

Other options are available to get a different space \(Y_N\) for velocity, for example: i) a space which is \(\mu\) independent, using only \(T^q \xi\) components to enrich velocity space. This option is useful if we want to apply an orthonormalization procedure to restore algebraic stability; or ii) a space \(\mu\) dependent, using the online value of the parameter in \(\Phi^q\). For more elements on the choice of supremum space we recall what introduced in Section 5.7.2 dealing with Stokes flows and in [111]. We remind briefly also the reduced basis equivalent LBB inf-sup constant \(\beta_N(\mu)\) with respect to \(\beta(\mu)\) associated with the Galerkin method, and the LBB inf-sup condition which, if satisfied as in (ii), guarantees the stability of the approximation (Lemma 5.3.1):

\[
\beta_N(\mu) \geq \beta(\mu) \geq \beta_0 > 0, \forall \mu \in D^\mu
\]

where

\[
\beta_N(\mu) = \inf_{q \in Q_N} \sup_{w \in Y_N} \frac{B(\mu, q, w)}{\|w\|_Y \|q\|_Q}.
\]
and
\[ \beta(\mu) = \inf_{q \in Q} \sup_{w \in Y} \frac{B(\mu, q, w)}{\|w\|Y \|q\|Q}. \]
The reduced basis solution for a new sample \( \mu \) is built as:
\[ u_N(\mu) = \sum_{j=1}^{2N} u_{Nj}(\mu) \sigma_j, \quad p_N(\mu) = \sum_{l=1}^{N} p_{Nl}(\mu) \xi_l; \]
and the reduced basis non-linear system for \( 1 \leq i, j \leq 2N, \ 1 \leq l \leq N \) is:
\[
\begin{aligned}
\sum_{j=1}^{2N} A_{ij}^{(k+1)}(\mu) + \sum_{l=1}^{N} B_{il}^{(k+1)}(\mu) &+ \sum_{h=1}^{N} \sum_{j=1}^{2N} u_{Njh}^{(k)}(\mu) C_{ijh}^{(k+1)}(\mu) + \\
\sum_{h=1}^{N} \sum_{j=1}^{2N} u_{Njh}^{(k+1)}(\mu) C_{ijh}^{(k+1)}(\mu) &+ \sum_{h=1}^{N} \sum_{j=1}^{2N} u_{Njh}^{(k+1)}(\mu) C_{ijh}^{(k+1)}(\mu) + \\
&= F_i + \sum_{h=1}^{N} \sum_{j=1}^{2N} u_{Njh}^{(k)}(\mu) C_{ijh}^{(k)}(\mu) + \sum_{h=1}^{N} \sum_{j=1}^{2N} u_{Njh}^{(k+1)}(\mu) C_{ijh}^{(k+1)}(\mu)
\end{aligned}
\]
\[
\sum_{j=1}^{2N} B_{ij}^{(k+1)}(\mu) = G_i, \quad 1 \leq l \leq N;
\]
(7.4.16)
\[
A_{ij}^{\mu} = \sum_{m=1}^{Q} \Theta^{m}(\mu) A(\sigma_i, \sigma_j)^{m}, \quad B_{ij}^{\mu} = \sum_{m=1}^{Q} \Phi^{m}(\mu) B(\sigma_i, \xi_l)^{m};
\]
\[
C_{ijh}^{\mu} = \sum_{m=1}^{Q} \Theta^{m}(\mu) C(\sigma_h, \sigma_j, \sigma_i)^{m};
\]
\[
F_i^{\mu} = \langle F, \sigma_i \rangle, \quad G_i^{\mu} = \langle G, \xi_l \rangle.
\]
We recall that the \( k \) index refers to the Newton iteration. The form \( C^{\mu} \) is depending on 3 different indexes and is assembled online given the value of \( \mu \). Then for every \( k \) we have to incorporate in \( C^{\mu} \) the solution at the previous iteration to update the system matrix and the right-hand-side. We can adopt the following stopping criterium
\[
\|u_N^{(k+1)}(\mu) - u_N^{(k)}(\mu)\| \leq \epsilon_N
\]
where \( \epsilon_N \) is a prescribed tolerance. When the Stokes solution is available, as in our case, it may be used as initial guess \( u_N^{(0)} \). If we compare the assembling and computational costs of Navier-Stokes reduced basis problem with respect to Stokes problem we have to increase \( O(9KN^3) \) operations for the solution of the non-linear system (\( K \) is the number of Newton iteration). In reality we are going to use more efficient and faster solver with splitting procedures. The other computational cost comes from the assembling procedure of \( C \) matrices (advection terms), whose theoretical assembling costs are \( O(Q^68N^3) \). We recall that the assembling costs for \( A, B \) and \( F \) are, respectively, \( O(Q^64N^2) \), \( O(Q^62N^2) \) and \( O(N) \). The structure of the system (7.4.16) is
\[
\begin{pmatrix}
A + C^{(k+1)} & B \\
B^T & 0
\end{pmatrix}
\begin{pmatrix}
u_N^{(k+1)} \\
\bar{p}_N^{(k+1)}
\end{pmatrix}
= 
\begin{pmatrix}
F^{(k)} \\
G
\end{pmatrix};
\]
(7.4.17)
Numerical results dealing only with geometrical affine parametric dependence will be reported in Section 7.7.
In this section we extend the formulation given in the previous one dealing only with non-affine geometrical parametric dependence in the reduced basis framework for Navier-Stokes equations. We consider a non-affine mapping from the true subdomains $\Omega^r$ into the reference ones denoted with $\Omega^\ast$. For any $\tilde{x} \in \Omega^\ast$, $r = 1, \ldots, R$, its image $x \in \Omega^r$ is given by

$$x = T^r(\mu; \tilde{x}), \quad 1 \leq r \leq R;$$

we thus write on $\Omega^r$

$$\partial \frac{\partial x_j}{\partial x_i} = \partial x_j = T^r_{ji}(\mu, x) \frac{\partial}{\partial x_j}.$$  

(7.5.2)

Referring to the problem (7.4.9) each elements in the system has the following form for $1 \leq i, j \leq 2$:

$$A(\mu; u, w) = \sum_{r=1}^{R} \int_{\Omega^r} \frac{\partial u}{\partial x_i} \left( T^r_{ii}(\mu, x) T^r_{jji}(\mu, x) \det(T^r(\mu, x))^{-1} \right) \frac{\partial w_j}{\partial x_i} \, d\Omega \forall w \in Y,$$

$$B(\mu; p, w) = -\sum_{r=1}^{R} \int_{\Omega^r} p \left( T^r_{ij}(\mu, x) \det(T^r(\mu, x))^{-1} \right) \frac{\partial w_j}{\partial x_i} \, d\Omega \forall w \in Y,$$

$$C(\mu; u, v, w) = -\sum_{r=1}^{R} \int_{\Omega^r} u_i \left( T^r_{ij}(\mu, x) \det(T^r(\mu, x))^{-1} \right) \frac{\partial v_j}{\partial x_i} \, w \, d\Omega \forall w \in Y,$$

$$\langle F_s, w \rangle = \sum_{r=1}^{R} \int_{\Omega^r} \left( \tilde{f}^r \det(T^r(\mu, x))^{-1} \right) w \, d\Omega.$$

In Section 7.8 we report a test case and we show the explicit forms of the previous terms. The transformation tensors for diffusion bilinear forms with non-affine mappings are defined as follows:

$$\nu^r_{Tij}(\mu, x) = T^r_{ii}(\mu, x) T^r_{jji}(\mu, x) \det(T^r(\mu, x))^{-1}, 1 \leq i, j \leq 2, r = 1, \ldots, R.$$  

(7.5.3)

The tensors for pressure, divergence and advection forms are defined for non-affine mappings as:

$$\chi^r_{Tij}(\mu, x) = \pi^r_{Tij}(\mu, x) = T^r_{ij}(\mu, x) \det(T^r(\mu, x))^{-1}.$$  

(7.5.4)

To decouple the non-affine contributions we apply the empirical interpolation procedure of Section 6.2 to expand mapping terms and split the parameters dependent contribution from the one depending only on spacial coordinates. We can write:

$$\nu^r_{Tij}(\mu, x) = \sum_{m=1}^{M^r_{ij}} \beta^r_{ijm}(\mu) \gamma^r_{ijm}(x),$$  

(7.5.5)

$$\chi^r_{Tij}(\mu, x) = \pi^r_{Tij}(\mu, x) = \sum_{m=1}^{M^r_{ij}} \alpha^r_{ijm}(\mu) \omega^r_{ijm}(x),$$  

(7.5.6)
The goal is to develop:

functions (shape functions) used as basis.

Reminder 7.5.1 We recall very briefly the algorithm based on the empirical interpolation method proposed by Maday et al. in [12]:

g(x, \mu), is the non-affine mapping term (i.e. a shape).

The goal is to develop:

\[ g_M(x, \mu) = \sum_{m=1}^{M} \beta_m(\mu) q_m(x), \]

as a sum of products decomposed in two parts: \( \beta_m(\mu) \) parameters dependent weights; \( q_m(x) \) shape functions without a parametric dependence.

The main elements are: test shape functions and service interpolation points, respectively:

\[ W_M^g = \{ \gamma_m = g(., \mu_m^g), 1 \leq m \leq M \}, \mu_m^g properly chosen, \]

\[ T_M = \{ t_1, \ldots t_M \}, 1 \leq M \leq M_{\text{max}}, \text{sets of interpolation points}. \]

The interpolation algorithm is:

For \( M = 1, t_1 = \text{argsup}_{x \in \Omega} |\gamma_1(x)|, q_1 = \gamma_1(x)/\gamma_1(t_1), (\text{aff - line}) \)

then, for \( M = 2, \ldots, M_{\text{max}}: \sum_{j=1}^{M-1} \sigma_j^{M-1} q_j(t_i) = \gamma_M(t_i), 1 \leq i \leq M - 1, (\text{aff - line}) \)

\[ r_M(x) = \gamma_M(x) - \sum_{j=1}^{M-1} \sigma_j^{M-1} q_j(x), t_M = \text{argsup}_{x \in \Omega} |r_M(x)|, (\text{aff - line}) \]

\[ q_M(x) = r_M(x)/r_M(t_M); g_M(x, \mu) = \sum_{m=1}^{M} \beta_m(\mu) q_m(x), (\text{aff - line}) \]

\[ \sum_{j=1}^{M} q_j(t_i) \beta_j(\mu) = g(t_i, \mu), 1 \leq i \leq M, (\text{on - line}) \]

To stop the procedure we impose \( \| g(., \mu) - g_M(., \mu) \|_{L^{\infty}(\Omega)} \leq \epsilon_{\text{max}} \) where \( \epsilon_{\text{max}} \) is an interpolation error. \( \square \)

We go back to our problem and, in order to build an effectively affine decomposition, we define:

\[ \Psi^{(i,j,r,m)}(\mu) = \beta_{ijm}^r(\mu), \quad A^{(i,j,r,m)}(\gamma(x), u, w) = \int_{\Omega^r} \gamma_{ijm}^r(x) \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_j} d\Omega, \quad (7.5.7) \]

\[ \Upsilon^{(i,j,r,m)}(\mu) = \alpha_{ijm}^r(\mu), \quad B^{(i,j,r,m)}(\omega(x), p, w) = - \int_{\Omega^r} \omega_{ijm}^r(x) p \frac{\partial u_i}{\partial x_j} d\Omega, \quad (7.5.8) \]

\[ \Upsilon^{(i,j,r,m)}(\mu) = \alpha_{ijm}^r(\mu), \quad C^{(i,j,r,m)}(\omega(x), u, v, w) = - \int_{\Omega^r} \omega_{ijm}^r(x) u_j \frac{\partial v_i}{\partial x_j} d\Omega, \quad (7.5.9) \]

for \( 1 \leq r \leq R, 1 \leq i, j \leq d = 2, 1 \leq m \leq \max(M_{ijr}^a, M_{ijr}^b) \) \( t \) and \( p \) are condensed indexes of \( i, j, r, m \) quantities. We rewrite our terms as:

\[ A(\mu, u, w) = \sum_{l=1}^{Q^a} \Psi^a(\mu) A^l(\gamma(x), u, w); \]
in this case $Q^a = \sum_{j=1}^d \sum_{i=1}^d \sum_{r=1}^R \Sigma_{ijr}^a$; $Q^b = \sum_{j=1}^d \sum_{i=1}^d \sum_{r=1}^R \Sigma_{ijr}^b$. We can now solve the Navier-Stokes problem written in (7.4.9) on the reference domain $\Omega$ for some values of parameter $\mu$ with the scope of building the reduced basis spaces. The formulation in this case is the same as the one in the previous section. We use the same supremizer option (the one with the $\mu^0$ offline value) so that we have a reduced basis velocity space $Y_N$ (7.4.14) which is $\mu$ independent. The only difference is the use of the effectively affine dependence of $B(\mu; q, w)$ on the parameter which reads:

$$T^{\mu_0} \xi = \sum_{p=1}^{Q^b} T^p(\mu_0) T^p \xi$$ (7.5.10)

for any $\xi$ and $\mu$, where:

$$(T^p \xi, w)_Y = B^p(\omega, q, w), \forall w \in Y.$$ 

Referring to the reduced basis non-linear system (7.4.16) we have the following sub-matrices $A$, $B$ and $C$, respectively:

$$A^\mu_{ij} = \sum_{z=1}^{Q^a} \Psi^z(\mu) A^z(\gamma, \sigma_i, \sigma_j), \quad 1 \leq i, j \leq 2N;$$

$$B^\mu_{il} = \sum_{z=1}^{Q^b} \Psi^z(\mu) B^z(\omega, \sigma_i, \xi_l), \quad 1 \leq i \leq 2N, \quad 1 \leq l \leq N;$$

and:

$$C^\mu_{ijh} = \sum_{z=1}^{Q^c} \Psi^z(\mu) C^z(\omega, \sigma_i, \sigma_j, \sigma_h), \quad 1 \leq i, j, h \leq 2N.$$ 

Numerical results dealing with a test case (a furrowed channel) will be reported in Section 7.8.

7.6 Parametrized formulation: affine and non-affine parametric dependence

In this section we are going to combine the formulation introduced in the two previous sections allowing us to prepare the framework for the application of reduced basis for Navier-Stokes equations to our problem of interest: the study of a bypass configuration. We deal with parametrized domains with both affine and non-affine parametric dependence (in different subdomains). In this case we are going to use a different supremizer option, building reduced
basis approximation space which is depending on the value of the “online” parameter \( \mu \).
We consider again a \( P^2 \) DES (Parametrized Partial Differential Equations) system, whose formulation on the reference domain \( \hat{\Omega} \) is the following: find \( (u(\mu), p(\mu)) \in Y \times Q \):

\[
\begin{aligned}
A(\mu; u(\mu), w) + B(\mu; p(\mu), w) + C(\mu; u(\mu), u(\mu), w) &= \langle F, w \rangle \forall w \in Y, \\
B(\mu; q, u(\mu)) &= \langle G^0, q \rangle \forall q \in Q.
\end{aligned}
\] (7.6.1)

The problem has been traced back to a reference domain by an affine mapping on some subdomains \( \Omega_{G}^r \) into \( \hat{\Omega}_{G}^r \) and by a non-affine mapping on the remaining subdomains \( \Omega_{T}^r \) into \( \hat{\Omega}_{T}^r \). More precisely, the physical domain \( \hat{\Omega} \) is the union of two (finite) families of subdomains \( \{\hat{\Omega}_{G}^r, r = 1, \ldots, R_G\} \) and \( \{\hat{\Omega}_{T}^r, r = 1, \ldots, R_T\} \). Then \( \hat{\Omega} = \bigcup_{r=1}^{R_G} \hat{\Omega}_{G}^r \cup \bigcup_{r=1}^{R_T} \hat{\Omega}_{T}^r \). For any \( \hat{x} \in \hat{\Omega}_{G}^r, r = 1, \ldots, R_G \), its image \( x \in \Omega_{G}^r \) is given by

\[
x = G^r(\mu; \hat{x}) = G^r(\mu)\hat{x} + g^r, \quad 1 \leq r \leq R_G,
\] (7.6.2)

therefore for \( 1 \leq i, j \leq 2 \):

\[
\frac{\partial}{\partial x_i} = \frac{\partial x_j}{\partial \hat{x}_i} \frac{\partial}{\partial x_j} = G_{ij}^r(\mu) \frac{\partial}{\partial x_j}.
\] (7.6.3)

On the other hand, for any \( \hat{x} \in \hat{\Omega}_{T}^r, r = 1, \ldots, R_T \), its image \( x \in \Omega_{T}^r \) is given by

\[
x = T^r(\mu; \hat{x}), \quad 1 \leq r \leq R_T,
\] (7.6.4)

thus

\[
\frac{\partial}{\partial x_i} = \frac{\partial x_j}{\partial \hat{x}_i} \frac{\partial}{\partial x_j} = T_{ij}^r(\mu, x) \frac{\partial}{\partial x_j}.
\] (7.6.5)

In the reference domain \( \hat{\Omega} \) we have:

\[
A(\mu; u, w) = \sum_{r=1}^{R_G} \int_{\hat{\Omega}_{G}^r} \frac{\partial u}{\partial x_i} \left( G_{i\nu}^r(\mu) \hat{\nu}_{i\nu}^r G_{j\mu}^r(\mu) \det(G^r(\mu))^{-1} \right) \frac{\partial w}{\partial x_j} d\Omega + \\
\sum_{r=1}^{R_T} \int_{\hat{\Omega}_{T}^r} \frac{\partial u}{\partial x_i} \left( T_{i\nu}^r(\mu, x) \hat{\nu}_{i\nu}^r T_{j\mu}^r(\mu, x) \det(T^r(\mu, x))^{-1} \right) \frac{\partial w}{\partial x_j} d\Omega \forall w \in Y,
\] (7.6.6)

\[
B(\mu; p, w) = -\sum_{r=1}^{R_G} \int_{\hat{\Omega}_{G}^r} p \left( G_{i\nu}^r(\mu) \det(G^r(\mu))^{-1} \right) \frac{\partial w_i}{\partial x_i} d\Omega + \\
-\sum_{r=1}^{R_T} \int_{\hat{\Omega}_{T}^r} p \left( T_{i\nu}^r(\mu, x) \det(T^r(\mu, x))^{-1} \right) \frac{\partial w_i}{\partial x_i} d\Omega \forall w \in Y,
\] (7.6.7)

\[
C(\mu; u, v, w) = \sum_{r=1}^{R_G} \int_{\hat{\Omega}_{G}^r} u_i \left( G_{i\nu}^r(\mu) \det(G^r(\mu))^{-1} \right) \frac{\partial v_j}{\partial x_i} w d\Omega + \\
-\sum_{r=1}^{R_T} \int_{\hat{\Omega}_{T}^r} u_i \left( T_{i\nu}^r(\mu, x) \det(T^r(\mu, x))^{-1} \right) \frac{\partial v_j}{\partial x_i} w d\Omega \forall w \in Y,
\] (7.6.8)

\[
\langle F, w \rangle = \langle F, w \rangle + \langle F^0, w \rangle,
\] (7.6.9)
where

\[ \langle F_s, w \rangle = \sum_{r=1}^{R_G} \int_{\Omega_G} \left( \hat{f}^r \det(G^r(\mu))^{-1} \right) w d\Omega + \sum_{r=1}^{R_T} \int_{\Omega_T} \left( \hat{f}^r \det(T^r(\mu, x))^{-1} \right) w d\Omega; \tag{7.6.10} \]

\[ \langle F^0, w \rangle = -\langle A_{g_m}, w \rangle; \quad \langle G^0, q \rangle = \langle B_{q}, g_{in} \rangle. \]

The transformation tensors for diffusion bilinear forms with affine and non-affine mappings are defined, respectively, as follows:

\[ \nu_{G_{ij}}^r(\mu) = G_{ij}^{r(\mu)} \hat{\nu}_{ij} G_{jj}^{r(\mu)} \det(G^r(\mu))^{-1}, 1 \leq i, j \leq 2, r = 1, ..., R_G, \tag{7.6.11} \]

\[ \nu_{T_{ij}}^r(\mu, x) = T_{ij}^r(\mu, x) \hat{\nu}_{ij} T_{jj}^r(\mu, x) \det(T^r(\mu, x))^{-1}, 1 \leq i, j \leq 2, r = 1, ..., R_T. \tag{7.6.12} \]

The tensors for pressure, divergence and advection forms are defined, respectively, for affine and non-affine mappings as:

\[ \chi_{G_{ij}}^r(\mu) = \pi_{G_{ij}}^r(\mu) = G_{ij}^r \det(G^r(\mu))^{-1}, \tag{7.6.13} \]

\[ \chi_{T_{ij}}^r(\mu, x) = \pi_{T_{ij}}^r(\mu, x) = T_{ij}^r(\mu, x) \det(T^r(\mu, x))^{-1}. \tag{7.6.14} \]

For the non-affine parts we apply, as already proposed, the empirical interpolation procedure of Section 6.2 (and recalled in the previous section) to expand non-affine mapping terms and decouple the parameters dependent contribution from the one depending only on spacial coordinates. We write:

\[ \nu_{G_{ij}}^r(\mu, x) = \sum_{m=1}^{M_{G}} \pi_{G_{ij}}^r(\mu) \gamma_{ijm}^r(x), \tag{7.6.15} \]

\[ \chi_{T_{ij}}^r(\mu, x) = \sum_{m=1}^{M_{G}} \pi_{T_{ij}}^r(\mu) \omega_{ijm}^r(x), \tag{7.6.16} \]

where \( m \) refers to the number of interpolation functions we use for each form (related with max interpolation error), \( i \) and \( j \) are indexes related to linear/bilinear form, \( r \) is the subdomain index, \( \beta \) and \( \alpha \) are weighing quantities depending on the parameters \( \mu, \gamma \) and \( \omega \) are interpolation functions used as basis.

Furthermore, we define

\[ \Theta^{(i,j,r)}(\mu) = \nu_{G_{ij}}^r(\mu), \quad \mathcal{A}^{(i,j,r)}_G(u, w) = \int_{\Omega_G} \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_j} d\Omega, \tag{7.6.17} \]

\[ \Phi^{(i,j,r)}(\mu) = \chi_{G_{ij}}^r(\mu), \quad \mathcal{B}_G^{(i,j,r)}(p, w) = -\int_{\Omega_G} p \frac{\partial w}{\partial x_j} d\Omega, \tag{7.6.18} \]

\[ \mathcal{C}^{(i,j,r)}_G(u, v, w) = \int_{\Omega_G} u_i \frac{\partial v}{\partial x_j} w d\Omega, \tag{7.6.19} \]

for \( 1 \leq r \leq R_G, 1 \leq i, j \leq d = 2 \) (\( q \) and \( s \) are condensed indexes of \( i, j, r \) quantities), and

\[ \Psi^{(i,j,r,m)}(\mu) = \beta_{ijm}^r(\mu), \quad \mathcal{A}^{(i,j,r,m)}_T(\gamma(x), u, w) = \int_{\Omega_T} \gamma_{ijm}(x) \frac{\partial u}{\partial x_i} \frac{\partial w}{\partial x_j} d\Omega, \tag{7.6.20} \]
\[ \Upsilon^{p(i,j,r,m)}(\mu) = \alpha^r_{i;j,m}(\mu), \quad B_T^{p(i,j,r,m)}(\omega(x), p, w) = -\int_{\Omega_T} \omega_{ijm}(x) \frac{\partial w_i}{\partial x_j} d\Omega, \]  
\[ C_T^{p(i,j,r,m)}(\omega(x), u, v, w) = \int_{\Omega_T} \omega_{ijm}(x) u_i \frac{\partial v_i}{\partial x_j} w d\Omega, \]

for \( 1 \leq r \leq R_T, 1 \leq i, j \leq d = 2, 1 \leq m \leq \max(M^a_{ijr}, M^b_{ijr}) \) (\( t \) and \( p \) are condensed indexes of \( i, j, r, m \) quantities used to simplify notation: each value of \( t \) or \( p \) represents a different combination of the previous four indexes \( i, j, r, m \)). We apply an effectively affine decomposition:

\[ A(\mu, u, w) = \sum_{q=1}^{Q_G^a} \Theta^q(\mu) A^q_G(u, w) + \sum_{l=1}^{Q_T^a} \Psi^l(\mu) A^l_T(\gamma(x), u, w); \]

\[ B(\mu, p, w) = \sum_{s=1}^{Q_G^b} \Phi^s(\mu) B^s_G(p, w) + \sum_{p=1}^{Q_T^b} \Upsilon^p(\mu) B^p_T(\omega(x), p, w); \]

\[ C(\mu, u, v, w) = \sum_{s=1}^{Q_G^c} \Phi^s(\mu) C^s_G(u, v, w) + \sum_{p=1}^{Q_T^c} \Upsilon^p(\mu) C^p_T(\omega(x), u, v, w); \]

in general, \( \max(Q_G^p) = d \times d \times d \times R_G, \max(Q_T^p) = d \times d \times R_G \); \( Q_T^a = \Sigma_{j=1}^{d} \Sigma_{i=1}^{d} \Sigma_{r=1}^{R_T} M^a_{ijr}, Q_T^b = \Sigma_{j=1}^{d} \Sigma_{i=1}^{d} \Sigma_{r=1}^{R_T} M^b_{ijr}, Q_T^c = \Sigma_{j=1}^{d} \Sigma_{i=1}^{d} \Sigma_{r=1}^{R_T} M^c_{ijr} \).

The non-linear problem (7.6) has to be discretized and then linearized to be solved, by an iterative method as seen in Section 7.4.

In the reduced basis approximation we choose properly (i.e. by optimized algorithm as seen in Section 5.4) a set of sample parameters \( S_N^\mu = \{ \mu^1, \ldots, \mu^N \} \), where \( \mu^N \in \mathcal{D}^\mu, n = 1, \ldots, N \). Correspondingly, we take a set of couples \( (u(\mu^N), p(\mu^N)) \) which are approximate solutions of the Navier-Stokes problem (7.6.1). Then the reduced basis pressure space is always \( Q_N = \text{span} \{ \xi_n, n = 1, \ldots, N \} \), where \( \xi_n = p(\mu^N) \), while for the reduced basis velocity space we take into consideration the option in which the space is \( \mu \)-dependent:

\[ Y_N^\mu = \text{span} \{ \zeta_n, n = 1, \ldots, N; T^\mu \xi_n, n = 1, \ldots, N \} \], where \( \zeta_n = u(\mu^N) \).

The reduced basis approximation problem reads: find \( (u_N(\mu), p_N(\mu)) \in Y_N \times Q_N \) s.t.:

\[ \begin{cases} 
A(\mu; u_N(\mu), w) + B(\mu; p_N(\mu), w) + C(\mu; u_N(\mu), u_N(\mu), w) = \langle F, w \rangle \forall w \in Y_N, \\
B(\mu; q, u_N(\mu)) = \langle C^0, q \rangle \forall q \in Q_N. 
\end{cases} \]

We re-write for computational convenience \( Y_N^\mu \) using the effectively affine dependence of \( B(\mu; q, w) \) on the parameter and the linearity of \( T^\mu \):

\[ T^\mu \xi = \sum_{q=1}^{Q_G^q} \Phi^q(\mu) T^q_G \xi + \sum_{p=1}^{Q_T^p} \Upsilon^p(\mu) T^p_T \xi \]

for any \( \xi \) and \( \mu \), where:

\[ (T^q_G \xi, w)_Y = B^q_G(q, w) \forall w \in Y, \]
which allows us to write:

\[
Y^*_{N}\text{span}\left\{\sum_{k=1}^{Q^p_G} \Phi^k(\mu)\sigma_{kn} + \sum_{k'=1}^{Q^p_T} \Upsilon^{k'}(\mu)\bar{\sigma}_{k'n}, n = 1, \ldots, 2N\right\},
\]

where \(Q^p_G = Q^b_{G} + 1, \Phi^0_G = 1\).

For \(n = 1, \ldots, N:\)

\[
\sigma_{kn} = 0, \text{ for } k = 1, \ldots, Q^b_{G};
\]

\[
\bar{\sigma}_{k'n} = 0, \text{ for } k' = 1, \ldots, Q^b_{P};
\]

\[
\sigma_{Q_{G}^b n} = \zeta_n = u(\mu^n).
\]

For \(n = N + 1, \ldots, 2N:\)

\[
\sigma_{k'n} = 0;
\]

\[
(\sigma_{k'n}, w)_{Y} = B_{G}(\xi_{n-N}, w), \forall w \in Y, \text{ for } k = 1, \ldots, Q^b_{G};
\]

\[
(\bar{\sigma}_{k'n}, w)_{Y} = B_{T}(\omega, \xi_{n-N}, w), \forall w \in Y, \text{ for } k = 1, \ldots, Q^b_{P}.
\]

For a new "\(\mu\)" we want a solution given by a combination of previously computed stored solutions as basis functions, i.e.:

\[
\begin{aligned}
  u_N(\mu) &= \sum_{j=1}^{2N} u_{Nj}(\mu) \left(\sum_{k=1}^{Q^p_G} \Phi^k(\mu)\sigma_{kj} + \sum_{k'=1}^{Q^p_T} \Upsilon^{k'}(\mu)\bar{\sigma}_{k'j}\right), \\
  p_N(\mu) &= \sum_{i=1}^{N} p_{Ni}(\mu)\xi_i,
\end{aligned}
\]

whose unknowns \(u_{Nj}\) and \(p_{Ni}\) satisfy the following non-linear system:

\[
\begin{cases}
  \sum_{j=1}^{2N} A_{ij}^{\mu} u_{Nj}(\mu) + \sum_{i=1}^{N} B_{ij}^{\mu} p_{Ni}(\mu) + \sum_{h=1}^{2N} \sum_{j=1}^{2N} u_{Nh}(\mu) C_{ijh}^{\mu} u_{Nj}(\mu) = F_i, & 1 \leq i \leq 2N, \\
  \sum_{j=1}^{2N} B_{ij}^{\mu} u_{Nj}(\mu) = G_i, & 1 \leq i \leq N.
\end{cases}
\]

\[
(7.6.27)
\]

To solve it we apply the Newton method reads, yielding the following iteration: for \(k \geq 0\) given \(u_{Nj}^{(k)}\), find \(u_{Nj}^{(k+1)}\) and \(p_{Ni}^{(k+1)}\) such that

\[
\begin{cases}
  \sum_{j=1}^{2N} A_{ij}^{\mu} u_{Nj}^{(k+1)}(\mu) + \sum_{i=1}^{N} B_{ij}^{\mu} p_{Ni}^{(k+1)}(\mu) + \sum_{h=1}^{2N} \sum_{j=1}^{2N} u_{Nh}^{(k)}(\mu) C_{ijh}^{\mu} u_{Nj}^{(k+1)}(\mu) + \\
  \sum_{h=1}^{2N} \sum_{j=1}^{2N} u_{Nh}^{(k+1)}(\mu) C_{ijh}^{\mu} u_{Nj}^{(k)}(\mu) = F_i^{(k+1)} + \sum_{h=1}^{2N} \sum_{j=1}^{2N} u_{Nh}^{(k)}(\mu) C_{ijh}^{\mu} u_{Nj}^{(k)}(\mu), & 1 \leq i \leq 2N, \\
  \sum_{j=1}^{2N} B_{ij}^{\mu} u_{Nj}^{(k+1)}(\mu) = G_i^{(k+1)}, & 1 \leq i \leq N, 1 \leq i \leq 2N.
\end{cases}
\]

\[
(7.6.28)
\]

The sub-matrices \(A, B, C\) are given by:

\[
A_{ij}^{\mu} = \sum_{z=1}^{Q^p_G} \sum_{k'=1}^{Q^p_G} \sum_{k''=1}^{Q^p_G} \Theta^z(\mu) \Phi^{k'}(\mu) \Phi^{k''}(\mu) A_{G}^z(\sigma_{k'z}, \sigma_{k''z}) +
\]
\[
\begin{align*}
&+ \sum_{z=1}^{Q_T^p} \sum_{k'=1}^{Q_T^p} \sum_{k''=1}^{Q_T^p} \Psi^z(\mu) \mathcal{Y}^{k'}(\mu) \mathcal{Y}^{k''}(\mu) A_T^z(\gamma, \bar{\sigma}_{k'1}, \bar{\sigma}_{k''1}), \quad 1 \leq i, j \leq 2N; \\
&= \sum_{z=1}^{Q_T^p} \sum_{k'=1}^{Q_T^p} \Phi^z(\mu) \Phi^{k'}(\mu) B_G^z(\sigma_{k'1}, \xi_1) + \\
&+ \sum_{z=1}^{Q_T^p} \sum_{k'=1}^{Q_T^p} \sum_{k''=1}^{Q_T^p} \mathcal{Y}^z(\mu) \mathcal{Y}^{k'}(\mu) B_T^z(\omega, \bar{\sigma}_{k'1}, \xi_1), \quad 1 \leq i \leq 2N, \quad 1 \leq l \leq N; \\
&\sum_{z=1}^{Q_T^p} \sum_{k'=1}^{Q_T^p} \sum_{k''=1}^{Q_T^p} \sum_{k''''=1}^{Q_T^p} \Phi^z(\mu) \Phi^{k'}(\mu) \Phi^{k''}(\mu) \Phi^{k''''}(\mu) C_G^z(\sigma_{k'1}, \sigma_{k''1}, \sigma_{k''''1}) + \\
&+ \sum_{z=1}^{Q_T^p} \sum_{k'=1}^{Q_T^p} \sum_{k''=1}^{Q_T^p} \sum_{k''''=1}^{Q_T^p} \mathcal{Y}^{k'}(\mu) \mathcal{Y}^{k''}(\mu) \mathcal{Y}^{k'''}(\mu) C_T^z(\omega, \bar{\sigma}_{k'1}, \bar{\sigma}_{k''1}, \bar{\sigma}_{k''''1}), \quad 1 \leq i, j, h \leq 2N; \\
&F_i = \sum_{k'=1}^{Q_T^p} \Phi^{k'}(\mu) \langle F, \sigma_{k'1} \rangle + \sum_{k'=1}^{Q_T^p} \mathcal{Y}^{k'}(\mu) \langle F, \bar{\sigma}_{k'1} \rangle, \quad 1 \leq i \leq 2N; \\
&G_l = \langle G_0, \xi_1 \rangle, \quad 1 \leq l \leq N.
\end{align*}
\]

In compact form the linearized problem (7.6.28) can therefore be written as:

\[
\begin{pmatrix}
A + C^{(k+1)} & \text{\(B\)} \\
B^T & 0
\end{pmatrix}
\begin{pmatrix}
U_N^{(k+1)} \\
L_N^{(k+1)}
\end{pmatrix}
= \begin{pmatrix}
F^{(k)} \\
G
\end{pmatrix}.
\] (7.6.29)

**Remark 7.6.1** This reduced basis formulation seems to be more involved than the one introduced in Section 7.5 due to the coupling between affine and non-affine maps in different subdomains from one hand, the use of a different supremizer which is \(\mu\) dependent, yielding a different reduced basis velocity approximation space from the other hand. \(\square\)

**Remark 7.6.2** We have the following computational costs to build reduced basis matrices, accounting also for the computation of supremizer components in the velocity space: \(O(Q^a(Q_G^0)^2)\) for sub-matrix \(A\), \(O((Q_T^p)^2)\) for \(B\), \(O(Q_T^p(Q_G^0)^3)\) for \(C\), \(O(Q_T^p N)\) for \(F\) and \(O(9N^3)\) for the “inversion” of the full reduced basis matrix (7.6.29) at each Newton iteration, where \(Q^a = Q_T^a + Q_G^a\), \(Q_T^b = \mathcal{C}_G^T + Q_T^p, \mathcal{C}_T^b = \mathcal{C}_G^T + Q_T^p\). Note that the quantities \(Q_G^a, Q_T^b\), and \(Q_T^b\) are depending only on the number of subdomains with affine mappings \((R_G)\), while quantities \(Q_T^a, Q_T^b,\) and \(Q_T^b\) are depending also on the number of “shape functions” \((\gamma(x)\) and \(\omega(x)\) related with interpolation error \((\varepsilon_{max})\) and the number of subdomains with non-affine mappings \((R_T)\). \(\square\)

Some numerical results based on the reduced basis approximation introduced in this section will be reported in Section 7.9.
CHAPTER 7. REDUCED BASIS FOR NAVIER-STOKES EQUATIONS

7.7 Some preliminary results

For our first test on the use of reduced basis for Navier-Stokes equations we deal only with affine mapping on a rectangular domain $(R = 1)$, parametrized by two quantities $D$ and $t$, see Figure 7.1 (right). We have used the formulation introduced in Section 7.4 and have considered a forced flow. The parameters range is: $0.1 \leq D \leq 1.5$ and $0.1 \leq t \leq 1.5$.

- To solve the parametrized Navier-Stokes problem in the domain outlined in Figure 7.1 we have imposed zero Dirichlet conditions on the boundary $\Gamma_D$, Neumann homogeneous conditions on the inflow $\Gamma_{Ni}$ and outflow $\Gamma_{No}$ ($\tau_{\hat{n}} = 0$, $\tau_{\hat{t}} = 0$, where $\tau = (\nu \frac{\partial u}{\partial n} - p\frac{\hat{n}}{\nu})$, with $\hat{n}$ and $\hat{t}$ normal and tangential directions, respectively).

- We have considered a forcing term $f = 10 \cdot (x, y)^T$ in order to create a flow acceleration, to have a non-zero velocity in $y$-direction and not only a parabolic velocity profile. Using the compact notation of Section 7.4 (7.4.10 and 7.4.11) and transformation (7.4.8) we get the following tensor for diffusion, divergence and transport forms, respectively:

$$
\nu = \bar{\nu} \begin{bmatrix} \frac{\partial}{\partial y} & 0 \\ 0 & \frac{\partial}{\partial x} \end{bmatrix}, \chi = \pi = \begin{bmatrix} t & 0 \\ 0 & D \end{bmatrix},
$$

where $\bar{\nu} = 0.04 Ns m^{-2}$ is the viscosity.

- Taylor-Hood finite elements have been used to build offline approximation basis functions: $\mathbb{P}^2 - \mathbb{P}^1$ elements for velocity (with supremizer) and pressure, respectively [127]. The problem has been solved using the Pressure-Matrix Method. The mean Reynolds number considered was of order $10^2$ (a low Reynolds number). Note, in fact, that the change of the value of the parameter $D$ (channel diameter) implies also a Reynolds number variation (by definition).

- At this step we have applied the reduced basis method and assembled the approximation spaces as described in Section 7.4. The basis assembling is based on the optimized procedure introduced in Section 5.4, see Figure 7.1 for the basis construction when increasing $N$.

- Figure 7.2 shows numerical results (mean and max $H^1$ and $L^2$ relative errors on velocity and pressure, respectively, on a large number of configurations) at different $N$. The reduced basis solutions have been compared directly with the approximate finite element solutions: the associate $H^1$ relative error for velocity and $L^2$ relative error for pressure are computed.

7.8 Furrowed channel test

Our second test on the use of reduced basis for Navier-Stokes equations deals with non-affine mapping on a rectangular domain $(R = 1)$, considering a channel with an upper wall...
parametrized by a sinusoidal law, see Figure 7.3 (left). This may be regarded as a simplified stenotic arterial flow, that was already investigated e.g. by Sobey and Stephanoff [143] and [148]. We have used the formulation introduced in Section 7.5. The test case is the same considered in Section 6.5.1 however here we consider Navier-Stokes equations whereas there we used Stokes equations. We briefly recall the assumptions that we have made during this study:

- To solve the parametrized Navier-Stokes problem in the domain outlined in Figure 7.3 we have imposed zero Dirichlet conditions on the boundary $\Gamma_D$, Neumann non-homogeneous conditions on the inflow $\Gamma_{N_i}$ ($\tau_\mathbf{n} = 1, \tau_\mathbf{t} = 0$, where $\tau = (\nu \frac{\partial u}{\partial \mathbf{n}} - p\mathbf{n})$, with $\mathbf{n}$ and $\mathbf{t}$ normal and tangential directions, respectively) and Neumann homogeneous conditions on outflow $\Gamma_{N_o}$ ($\tau_\mathbf{n} = 0, \tau_\mathbf{t} = 0$).

- We consider one parameter $\mu$ ranging in $[-0.8,0.8]$ to describe the upper arterial wall in the physical domain, through $\hat{x}_2 = f(\hat{x}_1, \mu) = 1 + \mu \sin(2\pi \hat{x}_1)$ (we have a single
domain subject to a unique non-affine mapping). Figures 7.3-7.5 show some simulations to better describe the fluid dynamics phenomena involved in this case of study. The upper curved wall is responsible of the changing of the channel section and so of the velocity gradient. Interesting phenomena to be “captured” by reduced basis method are the vertical velocity behavior (see Figure 7.5 on the right) and the development of a secondary flow (to be added to the mainstream channel) when $|\mu|$ increases (see Figure 7.4 on the right). In Sobey [143] an accurate study of steady (and unsteady) flows has been carried out dealing with furrowed channels at different Reynolds number, in particular focusing the attention over the secondary flows in the hollow zone of the channel. In our case the Reynolds ranges between $10^2$ and $4 \cdot 10^3$.

![Figure 7.3: Geometrical scheme for curved wall test model problem (left) and velocity (absolute value) for $\mu = 0.4$, $Re = 100$ (right). The colorbar is the same as in Figure 6.1.](image)

![Figure 7.4: Pressure (left) and vertical velocity (right), for $\mu = 0.7$ and $Re = 100$.](image)

Referring to Section 7.5 we have $\Omega_G = \emptyset$ and $\Omega_T = \Omega$, so $R_T = 1$. The coordinate transformation is $T : \bar{\Omega} \rightarrow \Omega$, $x = T(\bar{x})$, with

$$(x_1, x_2) = T(\bar{x}_1, \bar{x}_2) = (\bar{x}_1, \frac{1}{f(\bar{x}_1, \mu)} \bar{x}_2)$$  \hspace{1cm} (7.8.1)$$

in $\Omega$. Then,

$$d\bar{x}_1 d\bar{x}_2 = f(x_1, \mu) dx_1 dx_2,$$
and the following relations hold:

\[
\begin{align*}
\frac{\partial \hat{u}}{\partial x_2} &= \frac{1}{f(x_1, \mu)} \frac{\partial u}{\partial x_2}, \\
\frac{\partial \hat{u}}{\partial x_1} &= \frac{\partial \hat{\phi}}{\partial x_1} - x_2 f_{x_1}(x_1, \mu) \frac{\partial u}{\partial x_2} \quad \text{(with } f_{x_1} := \frac{df}{dx_1}).
\end{align*}
\]  

(7.8.2)

\[
\nabla \cdot \hat{u} = \frac{\partial u_1}{\partial x_1} - x_2 \frac{f_{x_1}(x_1, \mu)}{f(x_1, \mu)} \frac{\partial u_1}{\partial x_2} + \frac{1}{f(x_1, \mu)} \frac{\partial u_2}{\partial x_2}.
\]

(7.8.3)

Using the compact notation of Section 7.5 (7.5.3 and 7.5.4) and transformation (7.8.1) we get the following tensor for diffusion and divergence (and advection) forms, respectively:

\[
\nu_T = \nu \begin{bmatrix} f(x_1, \mu) & -f'_{x_1}(x_1, \mu)x_2 \\ -f'_{x_1}(x_1, \mu)x_2 & \frac{1}{f(x_1, \mu)} + \frac{f'^2_{x_1}(x_1, \mu)x_2^2}{f(x_1, \mu)} \end{bmatrix};
\]

(7.8.4)

\[
\chi_T = \pi_T = \begin{bmatrix} f(x_1, \mu) & -f'_{x_1}(x_1, \mu)x_2 \\ 0 & 1 \end{bmatrix};
\]

(7.8.5)

where \(\nu\) is the viscosity \([N m^{-2}]\) whose value was varied in simulating flows at different Reynolds number. Referring to notation of Reminder 7.5.1 we get 5 different coefficient functions \(g^j_M(x, \mu)\) to expand.

**Reminder 7.8.1** We recall here that the coordinate transformation (7.8.2) used in this example is the same one used in Chapters 3 and 4. □

- We apply empirical interpolation (7.5.5 and 7.5.6) to the tensors (7.8.4 and 7.8.5) and we impose a maximum interpolation error \(\varepsilon_{\text{max}}\), thus considering different \(M_{\text{max}}\) “shape functions” for each \(g^j_M(x, \mu)\). Each \(g^j_M(x, \mu)\) represents a different coefficient for a different term of our problem \((j = 5\) in this test case). Owing to empirical interpolation we expand each tensor component to apply the effectively affine decomposition:

\[
\nu_T = \nu \begin{bmatrix} \Sigma_{m=1}^{M_1} \beta_{11m}(\mu)\gamma_{11m}(x) & \Sigma_{m=1}^{M_2} \beta_{12m}(\mu)\gamma_{12m}(x) \\ \Sigma_{m=1}^{M_1} \beta_{21m}(\mu)\gamma_{21m}(x) & \Sigma_{m=1}^{M_2} \beta_{22m}(\mu)\gamma_{22m}(x) \end{bmatrix}.
\]
Note that this tensor is symmetric. Moreover,

\[ \chi_T = \pi_T = \begin{bmatrix} \sum_{m=1}^{M^1_b} \alpha_{11m}(\mu) \omega_{11m}(x) & \sum_{m=1}^{M^1_b} \alpha_{12m}(\mu) \omega_{12m}(x) \\ 0 & 1 \end{bmatrix}. \]

The index \( r \) referring to every different subdomain is omitted (in this case \( R = 1 \)).

- At this step we may apply the reduced basis formulation to this case and assemble the approximation spaces \( Y_N \) (7.4.14) and \( Q_N \) (7.4.13).

### 7.8.1 Results for low Reynolds number

We report some numerical results dealing with the solution of the furrowed channel rebuilt by reduced basis method. First we considered low Reynolds number (\( \sim 100 \)). Figures 7.6 and 7.7 show convergence results (mean and max \( H^1 \) and \( L^2 \) relative errors on velocity and pressure, respectively, testing a large number of configurations) at different \( N \) and at different imposition of max interpolation error \( \varepsilon_{\text{max}} \). At the end of the test we have carried out also a comparison between empirical interpolation (using \( g^j_M(x, \mu) \)) and true functions (\( g^j(x, \mu) \)). We can see that for \( \varepsilon_{\text{max}} \leq 10^{-8} \) we have accurate results that are not dominated or affected by interpolation error. When the interpolation error is dominating, the reduced basis error is characterized by a constant “plateau” and is not diminished by increasing \( N \) (see for example the case in which \( \varepsilon_{\text{max}} \geq 10^{-6} \)).

![Figure 7.6: \( H^1 \) and \( L^2 \) relative mean error on velocity and pressure (\( \nu = 0.1 \)) at different \( \varepsilon_{\text{max}} \) imposed on all \( g^j_M(x, \mu) \) (great number of testing configurations).](image)

### 7.8.2 Results for higher Reynolds number

A further test stage has been devoted to the increase of Reynolds number (\( \sim 4 \cdot 10^3 \)). We have reported in Figure 7.8 convergence results using a max interpolation error \( \varepsilon_{\text{max}} \geq 10^{-6} \). The comparison is always made between the “true” approximated solution by finite element method and the “interpolated” one by reduced basis.
7.8. FURROWED CHANNEL TEST

Figure 7.7: $H^1$ and $L^2$ relative max error on velocity and pressure ($\nu = 0.1$) at different $\epsilon_{\text{max}}$ imposed on all $g_M^j(x, \mu)$ (great number of testing configuration).

Figure 7.8: $H^1$ and $L^2$ relative error (max and mean) on velocity and pressure ($\nu = 0.04$) (testing a great number of configurations).

7.8.3 Reduced basis computational costs

At the end of this section we make some remarks on the computational costs in using the reduced basis in the online stage at different Reynolds number. This completes Remark 7.6.2 that was concerned with the assembling and computational costs. For simplicity we consider the case of the problem of the furrowed channel with one varying parameter and at different Reynolds number (at different viscosity values). Figure 7.9 shows online reduced basis computational cost (cputime, Pentium M 1.80 GHz, 1.0 Gb of Ram, IBM®T42 ThinkPad) for increasing $N$, compared with the computational cost of a finite element numerical simulation; the Reynolds number is $Re \leq 400$. We can see that reduced basis computational costs are $\sim 20\%$ of the finite element offline solutions if we choose $N = 12$ corresponding to an $H^1$ error on velocity of $O(10^{-6})$, as shown in the picture on the right. The reduces basis approximation spaces have been optimized during the assembling procedures. Increasing the Reynolds number ($400 \leq Re \leq 4000$) we can see how the computational saving of reduced basis techniques
are at least of two orders of magnitude (i.e. 1%) as shown in Figure 7.10, where on the right we have also a zoom representing online computational cost when increasing $N$. Results in reality depend on the choice of the initial guess for Newton iterations: in our case we have used Stokes solution (already calculated and stored). Other elements influencing computational costs are the setting of the tolerance for the Newton algorithm (difference between the solutions of two successive iterations), in our case set to $10^{-8}$. The numerical approximation of steady Navier-Stokes equations has been carried out by using a parametrized version of MLife, a FEM library [139] in Matlab-PDE toolbox environment where we have considered $\mathbb{P}^2 - \mathbb{P}^1$ elements for velocity and pressure, respectively, over a triangulation of $O(10^4)$ elements (see also Gresho and Sani [46]). The sparse linear system has been solved using Pressure-Matrix Method [127] (which is a block LU type decomposition of the system matrix) to compute velocity in two steps and to decouple the calculation of pressure from velocity (see [127]). The iterative methods used to solve the linearized system at each iteration have been GMRES and Bi-CGSTAB [129]. To improve computational efficiency we have used Cahouet-Chabard Preconditioner, see [22].

![Graphs showing computational costs and velocity error](image)

**Figure 7.9:** Reduced basis online mean computational costs in comparison with finite element simulation (left) and $H^1$ reduced basis velocity error increasing $N$ (right) for channel flow at low Reynolds numbers.

We observe that the use of reduced basis is providing real time solutions (or related outputs) of a non-linear parametrized problems, whose solution would be quite expensive even if considering efficient FEM solvers. The importance of real time accurate solutions increases when considering optimization problems involving also non-affine geometrical parametric dependence, as seen in the example of the furrowed channel.

### 7.9 The bypass problem

We retake into consideration the parametrized bypass of Figure 7.11 configuration with the vector of parameters $\boldsymbol{\mu} = \{t, D, L, S, H, \theta, \nu\} \in \mathcal{D}^\mu \subset \mathbb{R}^P$ with $\mathcal{D}^\mu$ given by:

$$
[t_{\min}, t_{\max}] \times [D_{\min}, D_{\max}] \times [L_{\min}, L_{\max}] \times [S_{\min}, S_{\max}] \times [H_{\min}, H_{\max}] \times [\theta_{\min}, \theta_{\max}]
$$
[\[v_{\text{min}}, v_{\text{max}}\]], already considered in Section 6.5.2. This test problem deals both with affine and non-affine parameters dependence in different subdomains and it is the application of the formulation introduced in Section 7.6. The aim of the test is to combine the study of affine and non-affine terms in the same non-linear problem by varying different geometrical parameters and then to test reduced basis convergence, extract output information and a sensitivity analysis on parameters. Referring to notation in Section 7.6 we have \( R_G = 3 \) (number of subdomains with affine dependence, i.e. \( \Omega^2, \Omega^3, \Omega^4 \)), \( R_T = 1 \) (number of subdomains with non-affine dependence, i.e \( \Omega^1 \)). The coordinate transformation in \( \Omega^1 \) with non-affine parameter dependence is given by:

\[
\begin{align*}
x_1 &= \frac{1}{\pi} \hat{x}_1 \\
x_2 &= \frac{1}{\pi} (\hat{x}_2 - (v H^2 x_1 (x_1 - 1) + Hx_1 \tan(\theta))).
\end{align*}
\]

Figure 7.10: Reduced basis online mean computational costs for channel flow at higher Reynolds numbers.

We recall that the role of parameters \( t \) and \( H \) is to stretch subdomain \( \Omega^1 \) (as \( L, S, D \) stretch the remaining subdomains), the parameter \( v \) introduces a curvature in the walls of the incoming branch of the bypass and \( \theta \) is responsible for a rigid rotation by letting the graft angle vary. The tensors for viscous bilinear terms are given by:

\[
\begin{align*}
\nu_T^1 &= \nu \begin{bmatrix}
\frac{t}{\pi} & \frac{(\tan \theta + 2v H x_1 - v H)}{(1+(\tan \theta + 2v H x_1 - v H)^2)^{1/2}} \\
-\tan \theta & \frac{2v H x_1 - v H}{(1+(\tan \theta + 2v H x_1 - v H)^2)^{1/2}}
\end{bmatrix}; \\
\nu_G^2 &= \nu \begin{bmatrix}
S & 0 \\
D & 0
\end{bmatrix}; \\
\nu_G^3 &= \nu \begin{bmatrix}
0 & L \\
0 & 0
\end{bmatrix}; \\
\nu_G^4 &= \nu \begin{bmatrix}
0 & 0 \\
0 & D
\end{bmatrix}.
\end{align*}
\]

The tensors for pressure, divergence and transport terms are given by:

\[
\begin{align*}
\chi_T^1 &= \pi_T^1 = \begin{bmatrix}
t & -H(\tan \theta + 2v H x_1 - v H) \\
0 & H
\end{bmatrix}; \\
\chi_T^2 &= \pi_T^2 = \begin{bmatrix}
S & 0 \\
0 & 0
\end{bmatrix}; \\
\chi_T^3 &= \pi_T^3 = \begin{bmatrix}
t & 0 \\
0 & D
\end{bmatrix}; \\
\chi_T^4 &= \pi_T^4 = \begin{bmatrix}
L & 0 \\
0 & 0
\end{bmatrix}.
\end{align*}
\]
We apply empirical interpolation expansion to the components of tensors \( \nu_T^1 \), \( \chi_T^1 \) and \( \pi_T^1 \) and we build the reduced basis approximation spaces for velocity and pressure.

![Geometrical scheme for the bypass test problem (physical domain and reference one).](image)

Figure 7.11: Geometrical scheme for the bypass test problem (physical domain and reference one).

We have carried out some tests based on the same geometry considering five different varying parameters (we have frozen \( L \) and \( H \)). In particular we are interested in varying graft angle \( \theta \) and curvature \( \upsilon \) (defining the upstream geometry) and the ratio \( \frac{L}{D} \). In Figures 7.12 and 7.13 we report numerical results (max and mean \( H^1 \) errors on velocity and \( L^2 \) errors for pressure) considering several configurations at different \( N \) for two different maximum interpolation error \( \epsilon_{max} = 10^{-5} \) and then \( \epsilon_{max} = 10^{-8} \) to avoid to have interpolation error dominating our approximation with the constant “plateau” in error plots.

![Figure 7.12: \( H^1 \) relative errors on velocity with different \( \epsilon_{max} \) interpolation error imposed on all \( g^j_M(x, \mu) \) (testing hundreds of different configurations with 5 different parameters varying).](image)

Figure 7.12: \( H^1 \) relative errors on velocity with different \( \epsilon_{max} \) interpolation error imposed on all \( g^j_M(x, \mu) \) (testing hundreds of different configurations with 5 different parameters varying).

### 7.9.1 Outputs sensitivities

We conclude this section with two different studies on the bypass problem providing a comparison between Stokes and Navier-Stokes solutions (and outputs) and some considerations about the influence of curvature of the upper stream geometry. The ratio \( \frac{L}{D} \) is the most important parameter and it is responsible of recirculation in the host artery, but also curvature...
Figure 7.13: $L^2$ relative errors on pressure with different $\epsilon_{\text{max}}$ interpolation error imposed on all $g_M^j(x, \mu)$ (testing hundreds of configurations with 5 different parameters varying).

has a role. The ratio $\frac{S}{D}$ becomes important if we freeze $\frac{t_D}{D}$ and so the graft angle $\theta$.

Figure 7.14 (left) shows a comparison between the vorticity output functional calculated with Stokes and Navier-Stokes equations. We can see that the “low fidelity” method is a good approximation only if the ratio $\frac{t_D}{D}$ is less than unity and this is our case, otherwise if the bypass diameter $t$ is too small with respect to the arterial diameter $D$ a strong recirculation arises in the host vessel and also vorticity increases considerably. These phenomena are not well captured by the linearized model. An optimized value of quantity $\frac{t_D}{D}$ is in the range $[0.85 - 0.96]$. This test gives us important information about the fidelity of our model and it allows a generalization of results available in the literature, especially provided by experimental research activity and surgical experience.

Figure 7.14 (right) shows the behavior of vorticity varying the curvature of the upstream (inflow) geometry. Increasing the curvature $\nu$ the vorticity diminishes: this behavior can be explained by the fact that curvature is guiding the flow more smoothly. An interesting analysis can be obtained introducing the Dean number, representing the ratio of the square root of the product of the inertial and centrifugal forces to the viscous forces, defined as follows

$$De = 4 \left[ \frac{D}{R} \right]^{\frac{1}{2}} Re,$$  \hspace{1cm} (7.9.6)

where $R$ is the radius of curvature and $Re$ the Reynolds number. See also Doorly and Sherwin in [141]. If we do not have curvature the Dean number is equal to zero. In the case we have considered, the range of the Dean number was $[0, 2.31 \cdot Re]$. By increasing the Dean number (and so curvature, the inverse of the curvature radius) makes the peak velocity to increase, but at the same time the velocity peak is displaced away from the centre of curvature. Note that $De = 0$ corresponds to a case in which we have a centered velocity profile of Hagen-Poiseuille type. In our case the displacement of the peak velocity profile allows the blood to be driven into the host vessel more smoothly and to better adapt the upstream inflow condition at the junction geometry. In our case the critical zone of the bypass near the upper wall has lower mean velocity. The introduction of the upstream curvature has been discussed also in
Papaharilaou, Doorly and Sherwin [108]. Results in Figure 7.14 (right) refers to a graft angle of 45 degree and a ratio $\frac{t}{D} = 1$.

Figure 7.14: Distributed vorticity $[m^2 s^{-1}]$ varying $\frac{t}{D}$ and comparison between Stokes (*) and Navier-Stokes (Δ) flows (left); distributed vorticity $[m^2 s^{-1}]$ and curvature (right).

7.10 Conclusions

In this chapter we have extended the use of reduced basis methodology to non-linear problems in domains with non-affine parametric dependence. This extension has allowed us (i) to get useful and more realistic information about parametrized bypass configuration; (ii) to compare previous results from low fidelity method (Stokes based) with the new higher fidelity ones; (iii) to use the reduced basis as an optimization method itself in our haemodynamic application and not only as a pre-process optimization tool to investigate a coarse bridge configuration. When we are considering problems with an increasing complexity, such as the ones with non-linearities and non-affine parametrization, the use of reduced basis method becomes even more competitive and computational savings are more relevant.

In the next chapter we provide some perspectives in developing reduced basis for optimal control problems but with results not specifically oriented to our bypass application, but more oriented on haemodynamics related problems.
Chapter 8

Perspectives in Flow Control by Reduced Basis Methods

In this conclusive chapter we provide some perspectives on the use of reduced basis in parametrized flow control problems. We consider a different kind of problem: an advection-diffusion scalar problem in order to investigate possible extensions of the use of reduced basis for optimal control problems in parametrized domains by combining the solutions of state and adjoint problems. With this chapter we conclude the guidelines which led us from numerical simulation to optimal control by reduced basis methods.

8.1 Introduction

Control problems solved with the reduced basis methods were already presented by Ito and Ravindran since 1998 in [62], [63], and [64] without considering multi-parametric problems and geometrical parameters, and so adopting different solution procedures. Parameters can be divided into three classes: control (i.e. depending, in some way, on control function), physical (like velocity field or diffusivity) and geometrical (i.e. related to different domain configuration). We are going consider all these parameter classes to be able to solve many different “kind” of control problems.

As a case study, we consider an advection-diffusion problem which can be seen in the framework of life sciences, considering, for example, an application related with the release of drugs or chemical elements (solutes) in blood or, more generally, physiological flows. See the work by Quarteroni, Veneziani and Zunino in [130] and [131]. The modelling of advection-diffusion phenomena can be useful in biomedical field for the setting, for example, of appropriate peritoneal dialysis procedures or the design of drug eluting stents. See, for example, Zunino [170]. The goal can be the regulation of the substance emission by some sources in order to keep concentration at a desired and/or acceptable level over an observation area. A similar problem arises in environmental fluid dynamics. See, for example, [122].

In Section 8.2 we formulate a generic control problem for a linear time-independent advection-diffusion equation using Lagrangian formulation. In Section 8.3 we describe the reduced basis
approximation for the solution of the parametrized equations governing the control problem. In Section 8.4 we describe very briefly an example of optimal flow control problem by applying the formulation presented in Section 8.2 to derive our model. In Section 8.5, 8.6 and 8.7 we present the parametrized state and adjoint equations, some results and an example in the case of control, physical and geometrical input, respectively. In Section 8.8 we report some concluding remarks and further perspectives. Section 8.9 is a quick note on time dependent problems. Other very recent applications of reduced basis methods are provided in the field of optimal control by Grepl [44] (unsteady advection-diffusion problems) and by Nguyen [103] in the field of inverse problems with non-affine parametric dependence.

8.2 Optimal control problem for advection-diffusion equations

8.2.1 A formulation for optimal control problem based on Lagrangian

We recall that the goal of an optimal control problem is to find the function \( u \in U \) in order to minimize the cost functional \( J = J(u) \), where \( w \) is the solution of the state equation, i.e.:

\[
\text{find } u \in U \text{ so that } J(u) \text{ is minimized and } Aw = f + Bu, \forall v \in V,
\]

with \( U \) and \( V \) being two Hilbert spaces. The differential operator \( A \) is an elliptic operator defined on \( V \) with values in \( V' \), \( B \) is an operator defined on \( U \) and valued in \( V' \) which introduces control variable \( u \) in the state equation, \( f \) is a source field. To analyze optimal control problems, we adopt the Lagrangian approach [15], equivalent to optimal control classical analysis, developed for example by Lions in [82]. By introducing the Lagrangian multiplier \( p \) and the Lagrangian functional:

\[
\mathcal{L}(w, p, u) = J(u) + \langle p, Aw - f - Bu \rangle,
\]

the unique solution of problem (8.2.1) is determined by solving the PDE system:

\[
\nabla \mathcal{L}(w, p, u) = 0.
\]

The weak form of the state equation is:

\[
a(w, \varphi) = (f, \varphi) + b(u, \varphi), \forall \varphi \in V,
\]

where \( a(\cdot, \cdot) \) is the bilinear form associated to the elliptic operator and \( b(\cdot, \cdot) \) is the bilinear form associated to the control term \( Bu \). The cost functional can be expressed as:

\[
J(u) = \frac{1}{2}\|Cw - z_d\|^2 + \frac{1}{2}n(u, u),
\]

where operator \( C \) brings state variable \( w \) into the observation space (an Hilbert space), \( z_d \) is the desired observation function, that is the optimal control problem goal, and \( n(\cdot, \cdot) \) is an Hermitian form. Let’s denote the product \( \langle \cdot, \cdot \rangle \) as the \( L^2 \) scalar product. The Lagrangian functional can be expressed as:

\[
\mathcal{L}(w, p, u) = J(u) + b(u, p) + (f, p) - a(w, p),
\]
so that the optimal control solution can be calculated solving the following problem:

\[ \text{find } (w, p, u) \in \mathcal{V} \times \mathcal{V} \times \mathcal{U} : \nabla \mathcal{L}(w, p, u)[(\varphi, \phi, \psi)] = 0, \quad \forall (\varphi, \phi, \psi) \in \mathcal{V} \times \mathcal{V} \times \mathcal{U}. \]

We have then:

\[
\begin{align*}
\mathcal{L}_w[\varphi] &= (Cw - z_d, C\varphi) - a(p, \varphi) = 0, \quad \forall \varphi \in \mathcal{V}, \\
\mathcal{L}_p[\phi] &= (f, \phi) + b(u, \phi) - a(w, \phi) = 0, \quad \forall \phi \in \mathcal{V}, \\
\mathcal{L}_u[\psi] &= b(p, \psi) + n(u, \psi) = 0, \quad \forall \psi \in \mathcal{U}.
\end{align*}
\]

The differential \( \mathcal{L}_p \) corresponds to the weak form of the state equation, \( \mathcal{L}_w \) to the adjoint equation and \( \mathcal{L}_u \) is an optimal control constraint. The variable \( p \) is the adjoint variable and it expresses the cost functional sensitivity to the variations of the control variable \( u \).

The optimal control constraint \( \mathcal{L}_u[\psi] \) can be related to the strong derivative \( J'(u) \) by means of the cost functional definition (8.2.2) and the Riesz theorem. In particular we can write \( \mathcal{L}_u[\psi] = \langle J'(u), \psi \rangle = \langle J'(u), \psi \rangle \), since the functional \( \mathcal{L}_u[\psi] \) is linear and bounded and \( \psi \) belongs to an Hilbert space \([127]\).

To solve the control problem, we adopt an iterative method. Starting from an initial value for control variable \( u_0 \), we solve the state equation; then we compute cost functional value and we resolve the adjoint equation. Once \( p \) is known, we determine cost functional derivative \( J_0(u) \) and apply a suitable stopping criterium. If this criterium is not fulfilled, we start an iterative process on control function \( u \); we adopt, for instance, the steepest descent method:

\[ u^{k+1} = u^k - \tau^k J'(u^k), \quad (8.2.3) \]

where \( \tau^k \) is a relaxation parameter, whose value can be determined by control problem properties, see for example Agoshkov [6]. The iterative process ends with the optimal control criterium fulfillment.

### 8.2.2 Optimal control problem governed by advection-diffusion equations

Let us consider the following state equation of our problem:

\[
\begin{align*}
- \nabla \cdot (\nu \nabla w) + \mathbf{V} \cdot \nabla w &= u \quad \text{in } \Omega \subset \mathbb{R}^2, \\
w &= 0 \quad \text{on } \Gamma_D, \\
\frac{\partial w}{\partial n} &= 0 \quad \text{on } \Gamma_N,
\end{align*}
\]

(8.2.4)

where \( w \) is the state variable, \( u \) the control function defined on the domain, \( \mathbf{V} \) is the velocity field and \( \nu \) is the diffusivity, depending on the domain coordinates \( x = (x,y)^T \). A homogeneous Dirichlet condition is imposed on inflow boundary \( \Gamma_D := \{ x \in \partial \Omega : \mathbf{V}(x) \cdot \hat{n}(x) < 0 \} \), where \( \hat{n}(x) \) is the unit vector directed outward, and a homogeneous Neumann condition on the rest of the boundary \( \Gamma_N := \partial \Omega \setminus \Gamma_D \). Defining \( H^1_{\Gamma_D} := \{ v \in H^1(\Omega) : v|_{\Gamma_D} = 0 \} \), the weak form of the state equation is:

\[ \text{find } w \in H^1_{\Gamma_D} : a(w, \varphi) = F(\varphi; u), \quad \forall \varphi \in H^1_{\Gamma_D}, \]
where
\[
a(w, \varphi) := \int_{\Omega} \nu \nabla w \cdot \nabla \varphi \, d\Omega + \int_{\Omega} \mathbf{V} \cdot \nabla w \, \varphi \, d\Omega, \tag{8.2.5}
\]
\[
F(\varphi; u) := \int_{\Omega} u \varphi \, d\Omega. \tag{8.2.6}
\]

We then define the observation of the system on a part \(D\) of the domain \((D \subset \Omega)\) with the following cost functional:
\[
J(u) = \frac{1}{2} \int_{D} \left( w(u) - z_d \right)^2 dD,
\]
where \(z_d\) is the desired observation.

It is now possible to write the Lagrangian functional:
\[
L(w, p, u) = J(u) + F(p; u) - a(w, p),
\]
where \(w, p \in H^1_{1,D}(\Omega)\) and \(u \in L^2(\Omega)\). By differentiating it with respect to the state variable, we obtain the weak form for the adjoint equation:
\[
\text{find } p \in H^1_{1,D} : a^{ad}(p, \phi) = F^{ad}(\phi; w), \quad \forall \phi \in H^1_{1,D},
\]
where
\[
a^{ad}(p, \phi) := \int_{\Omega} \nu \nabla p \cdot \nabla \phi \, d\Omega + \int_{\Omega} \mathbf{V} \cdot \nabla \phi \, p \, d\Omega, \tag{8.2.7}
\]
\[
F^{ad}(\phi, w) := \int_{\Omega} (w - z_d) \, g(x, y) \, \phi \, dD, \tag{8.2.8}
\]
whose strong form is:
\[
\begin{cases}
-\nabla \cdot (\nu \nabla p + \mathbf{V} \cdot p) = m_D(w - z_d) \text{ in } \Omega, \\
p = 0 \text{ on } \Gamma_D, \\
\nu \frac{\partial p}{\partial n} + \mathbf{V} \cdot \mathbf{n} \, p = 0 \text{ on } \Gamma_N,
\end{cases}
\]
where \(m_D\) is the characteristic function of the subdomain \(D\).

By differentiating the Lagrangian functional with respect to the control function \(u\), we obtain the weak form of the optimal condition constraint:
\[
\mathcal{L}_u[\psi] = \langle J'(u), \psi \rangle = \langle J'(u), \psi \rangle = \int_{\Omega} p \psi \, d\Omega = 0, \quad \forall \psi \in L^2(\Omega).
\]

We solve our problem using an iterative method where the variation of control function is led by a gradient method. From the optimal control constraint, we can derive a stopping criterium for the iterative method. At \(k - th\) step of the iterative method:

- we solve state equation
\[
\text{find } w^k \in H^1_{1,D} : a(w^k, \varphi) = F(\varphi; u^k), \quad \forall \varphi \in H^1_{1,D};
\]

- we solve adjoint equation
\[
\text{find } p^k \in H^1_{1,D} : a^{ad}(p^k, \phi) = F^{ad}(\phi; w^k), \quad \forall \phi \in H^1_{1,D};
\]
Chapter 8.2. Optimal Control Problem for Advection-Diffusion Equations

- if the stopping criterium is not satisfied, we update control function

\[ u^{k+1} = u^k + \delta u^k \quad \delta u^k = -\tau^k p^k. \]

The stopping criterium adopted [32] is:

\[ \|p^k\|_{L^2} < \text{tol}, \quad (8.2.9) \]

that is we check if adjoint variable \( p \) \( L^2 \)-norm is too small (according to our fixed tolerance \( \text{tol} \)) to produce a significative variation \( \delta u^k \) on the new control function \( u^{k+1} \).

### 8.2.3 Numerical discretization and stabilization

Since the state and adjoint equations are both advection-diffusion equations and transport term can dominate the diffusion one, a suitable stabilization is needed. We adopt the Stabilized Lagrangian [32], instead of stabilizing separately state and adjoint equations in a conventional manner [127]. In this way, stabilization is not only based on a strongly consistent method, but also there is coherence between state and adjoint stabilized equations.

In this thesis we have adopted the approach “optimize-then-discretize” to solve optimal control problems, as seen in Chapter 2, 3 and 4; we have firstly formulated an optimality condition, from this condition we have built an adjoint problem and then we have discretized both state and adjoint equations. In this chapter we use the same approach to be coherent; an alternative approach would be the so-called “discretize then optimize” which has been considered for the same kind of problem in [124].

Indicating with the subscript \( h \) the discretized quantities, the stabilized state equation reads:

\[ \text{find } w_h \in X_h : \quad a_h(w_h, \varphi_h) = F_h(\varphi_h; u_h), \quad \forall \varphi_h \in X_h, \quad (8.2.10) \]

where:

\[ a_h(w_h, \varphi_h) := a(w_h, \varphi_h) - \sum_{K \in T_h} \delta_K \int_K \left( - \nabla \cdot (\nu \nabla w_h) + \mathbf{V} \cdot \nabla w_h \right) \left( - \nabla \cdot (\nu \nabla \varphi_h + \mathbf{V} \varphi_h) \right) dK, \]

\[ F_h(\varphi_h; u_h) := F(\varphi_h; u_h) - \sum_{K \in T_h} \delta_K \int_K u_h \left( - \nabla \cdot (\nu \nabla \varphi_h + \mathbf{V} \varphi_h) \right) dK. \]

The terms \( a(w_h, \varphi_h) \), \( F(\varphi_h; u_h) \) are defined in (8.2.5), (8.2.6) and \( w_h \), \( u_h \) are discrete approximation of the functions \( w \), \( u \). The space \( X_h \subset H^1_{\Gamma_D} \) is the finite element one. The finite element solution is computed over the grid \( T_h \), so that the computational domain is \( \overline{\Omega} = \bigcup_{K \in T_h} K \).

The adjoint equation is:

\[ \text{find } p_h \in X_h : \quad a_h^{ad}(p_h, \phi_h) = F_h^{ad}(\phi_h; w_h; u_h), \quad \forall \phi_h \in X_h, \quad (8.2.11) \]

where:

\[ a_h^{ad}(p_h, \phi_h) := a^{ad}(p_h, \phi_h) - \]
\[
\sum_{K \in T_h} \delta_K \int_K \left( -\nabla \cdot (\nu \nabla p_h + \mathbf{V} p_h) \right) \left( -\nabla \cdot (\nu \nabla \phi_h) + \mathbf{V} \cdot \nabla \phi_h \right) \, dK,
\]

\[
F_h^{ad}(\phi_h; w_h; u_h) := F_h^{ad}(\phi_h; w_h) - \sum_{K \in T_h} \delta_K \int_K \left( m_D (w_h - z_d) \right) \cdot 
\left( -\nabla \cdot (\nu \nabla \phi_h) + \mathbf{V} \cdot \nabla \phi_h \right) + \left( -\nabla \cdot (\nu \nabla w_h) + \mathbf{V} \cdot \nabla w_h - u_h \right) m_D \phi_h \, dK.
\]

Note that the terms \( F_h^{ad}(\phi_h; w_h) \) are defined in (8.2.7) and (8.2.8), while \( p_h \) is the discrete approximation of functions \( p \).

8.3 Reduced basis method for optimal control

As anticipated in Section 8.1, we can consider an input parameter of three different types:

- control input \( \mu_u \), which define control function \( u = u(\mu_u) \):
- physical input \( \mu_p \), for example, velocity field \( \mathbf{V} \) and/or viscosity \( \nu \);
- geometrical input \( \mu_g \), that is geometrical parameters which can vary the configuration of the domain.

We make this distinction into parameter classes in order to be able to consider several aspects (advection fields, diffusivity terms, localization of sources for releases, etc.) of the same problem: even if, in reality, we are solving different “types” of problem. Of course, input can be combined together to form, for example, a control-physical input. In this section, for convenience, we refer to a generic input \( \mathbf{\mu} = \{\mu_u, \mu_p, \mu_g\} \), without specifying its nature.

We introduce a set of parameter samples \( S_N = \{\mu^1, \ldots, \mu^N\} \), where \( \mu^n \in \mathcal{D}, n = 1, \ldots, N \). For each input vector in \( S_N \), we calculate using the finite element method a solution of the state equation \( w_h(\mu^n) \) in the space \( X_h \); we choose a discretization enough refined to ensure that the solution in the high-dimensional space \( X_h \) is “in good agreement” with the exact solution in \( H^1_D \). We do the same for the adjoint problem (note the similarity of this approach with the one used to set the dual residual method to correct the cost functional in Section 5.10): find a set of \( N \) samples \( S_N^{ad} = \{\mu_{ad}^1, \ldots, \mu_{ad}^N\} \), where \( \mu_{ad}^n \in \mathcal{D}, n = 1, \ldots, N \), and compute the finite element approximation of the adjoint variable \( p_h(\mu_{ad}^n) \in X_h \). The two sets \( S_N \) and \( S_N^{ad} \) are chosen independently. Also the reduced basis formulation and the basis construction procedure have been influenced by our choice of using the approach “optimize-then-discretize” for optimal control problem. We introduce the following reduced basis spaces:

\[
W_N = \text{span}\{\xi^n \equiv w_h(\mu^n), \ n = 1, \ldots, N\} \tag{8.3.1}
\]

for the state problem and

\[
Z_N = \text{span}\{\xi^n \equiv p_h(\mu_{ad}^n), \ n = 1, \ldots, N\} \tag{8.3.2}
\]
for the adjoint problem. Starting from the state variable, we look for an approximation \( w_N(\mu) \) to \( w_h(\mu) \) in \( W_N \); in particular we express \( w_N(\mu) \) as:

\[
w_N(\mu) = \sum_{j=1}^{N} w_{Nj}(\mu) \zeta^j = (\mathbf{w}_N(\mu))^T \zeta,
\]

(8.3.3)

where \( \mathbf{w}_N(\mu) \in \mathbb{R}^N \) is the column vector of the linear combination coefficients \( w_{Nj}, j = 1, ..., N \). Let \( p_N \) be the reduced basis approximation of the adjoint variable:

\[
p_N(\mu) = \sum_{j=1}^{N} p_{Nj}(\mu) \xi^j = (\mathbf{p}_N(\mu))^T \xi.
\]

(8.3.4)

The goal is to represent accurately the solution of state and adjoint problem at some new point in parameter space, \( \mu^{new} \), as an appropriate combination of solutions previously computed at a small number of sample points in parameter space (\( \mu^n \) and \( \mu^{adn}, n = 1, ..., N \)). In this case we are interested in solving the optimal control problem and finding the control function evaluating the cost functional in a rapid, reliable and repeated way.

At each iterative step of the method adopted to solve the control problem, for the given \( \mu \in \mathcal{D} \) and the given control function \( u \) (8.2.3):

- we compute the reduced basis approximation of the state variable \( w_N(\mu) \in W_N \), where

\[
a(w_N(\mu), v; \mu) = F(v; u), \ \forall v \in W_N;
\]

- once \( w_N \) is available, we determine the solution \( p_N \in Z_N \) of the adjoint equation:

\[
a^{ad}(p_N(\mu), \phi; \mu) = F^{ad}(\phi, w_N(\mu)), \ \forall \phi \in Z_N;
\]

- we evaluate the reduced basis approximation of our output, i.e. the cost function \( J(u, w_N) \) and the adjoint variable \( p_N \), which allows us to check whether the stopping criterion is satisfied.

To apply the reduced basis method, we suppose, as usual, that for some finite (preferably small) integers \( Q \) and \( Q^{ad} \), the bilinear forms \( a(\cdot, \cdot; \mu) \) and \( a^{ad}(\cdot, \cdot; \mu) \) may be expressed as:

\[
a(w, v; \mu) = \sum_{q=1}^{Q} \sigma^q(\mu) a^q(w, v), \ \forall w, v \in H^1_D, \ \forall \mu \in \mathcal{D},
\]

\[
a^{ad}(p, \phi; \mu) = \sum_{q=1}^{Q^{ad}} \sigma^{adq}(\mu) a^{adq}(p, \phi), \ \forall p, \phi \in H^1_D, \ \forall \mu \in \mathcal{D},
\]

for some \( \sigma^q(\mu) : \mathcal{D} \to \mathbb{R}, a^q : H^1_D \times H^1_D \to \mathbb{R}, q = 1, ..., Q \) and for some \( \sigma^{adq}(\mu) : \mathcal{D} \to \mathbb{R}, a^{adq} : H^1_D \times H^1_D \to \mathbb{R}, q = 1, ..., Q^{ad} \). We assume also affine parameter dependence for the functionals \( F \) and \( F^{ad} \).

Coming to the matrix form, we define the matrices \( A_N(\mu) = a(\zeta^i, \zeta^j; \mu) \) and \( A_N^{ad}(\mu) = a^{ad}(\zeta^i, \zeta^j; \mu) \).
\( a^{ad}(\xi^i, \xi^j; \mu), 1 \leq i, j \leq N \) and the vectors \( F_N = F(\xi^j, u) \) and \( F_N^{ad} = F^{ad}(\xi^j, u_N), 1 \leq j \leq N \). We observe that:

\[
A_N(\mu) = \sum_{q=1}^{Q} \sigma^q(\mu) A^q, \tag{8.3.5}
\]

\[
A_N^{ad}(\mu) = \sum_{q=1}^{Q^{ad}} \sigma_{ad}^q(\mu) A_{ad}^q,
\]

where \( A^q_{i,j} = a^q(\xi^i, \xi^j), 1 \leq i, j \leq N, 1 \leq q \leq Q \) and \( A_{ad}^q_{i,j} = a_{ad}^q(\xi^i, \xi^j), 1 \leq i, j \leq N, 1 \leq q \leq Q^{ad} \). Note that \( A^q, 1 \leq q \leq Q \) and \( A_{ad}^q, 1 \leq q \leq Q^{ad} \) are independent of input parameter \( \mu \).

We can then reformulate the state equation as: given \( \mu \in \mathcal{D} \), find the unique solution \( w_N(\mu) \) to

\[
A_N(\mu) w_N(\mu) = F_N; \tag{8.3.6}
\]

and the adjoint equation as: given \( \mu \in \mathcal{D} \) and \( w_N(\mu) \), find the unique solution \( p_N(\mu) \) to

\[
A_N^{ad}(\mu) p_N(\mu) = F_N^{ad}. \tag{8.3.7}
\]

### 8.3.1 Computational procedure: off-line/on-line decomposition

The matrices and vectors introduced in the previous section are assembled in two different steps: the parameter-dependent (online, many queries) and the parameter-independent ones (off-line, computed once).

In the off-line stage, we find the \( \xi^i, \xi^i, i = 1, \ldots, N \) and form the \( A^q \), for \( 1 \leq q \leq Q \), the \( A_{ad}^q \), for \( 1 \leq q \leq Q^{ad} \) and \( F_N, F_N^{ad} \). The online stage, for any given new \( \mu \), we only need to form \( A_N(\mu) \) from the \( A^q, A_N^{ad}(\mu) \) from the \( A_{ad}^q \), then solve (8.3.6) for \( w_N(\mu) \) and (8.3.7) for \( p_N(\mu) \) and finally evaluate \( J(u, w_N; \mu) \).

Note that the two processes are completely decoupled. Also in this case the expensive off-line computation be processed at an early stage and needs to be done only once. The efficient online computation can then be used for very fast evaluations of outputs at different point in the parameter space. The incremental cost to evaluate, for example, \( J_N(\mu) \) for any given new \( \mu \) is very small: (i) \( N \) is very small, typically \( \mathcal{O}(10) \) thanks to the good convergence properties of \( W_N \) and \( Z_N \) [134]; (ii) (8.3.6) and (8.3.7) can be assembled and inverted very rapidly.

### 8.3.2 Error on control and error on cost functional

As seen, while solving a simple equation, for example the state equation, the reduced basis approach looks for an approximation \( w_N(\mu) \) to \( w(\mu) \) in \( W_N \), presuming that \( w(\mu) \) is sufficiently close to \( w(\mu) \). This means that we make no distinction between the actual solution and the finite element solution computed on a grid refined enough for our purposes. When dealing with an optimal control problem, it is important to check that the problem solved with the finite element method and the one solved with the reduced basis method converge to
§8.4. AN APPLICATION TO FLOW CONTROL

8.4 An application to flow control

We consider now a type of control problem with the aim of regulating a generic substance emission by some sources into a fluid (for example a physiological flow) in order to achieve a desired concentration level or to maintain the concentration level below a fixed threshold over an observation area in the domain. This kind of problem is a parametrized extension of the one in [32]. Preliminary results are reported in [124] dealing with a different application.

We refer to the domain in Figure 8.1 where we have positioned three different sources $Q_1$, $Q_2$, and $Q_3$. The control problem is to find the optimal control function $u_h$ that minimizes the cost functional $J(u_h, w_h)$ over the domain $\Omega$. The control function $u_h$ is defined as the source strength at each source location $Q_i$. The cost functional $J(u_h, w_h)$ is given by

$$J(u_h, w_h) = \frac{1}{2} \int_{\Omega} (u_h - w_h)^2 \, d\Omega,$$

where $u_h$ is the control function corresponding to $Q_1$, $Q_2$, and $Q_3$, and $w_h$ is a known data function representing the desired concentration level. The problem is solved using an iterative method called the adjoint method, which involves solving the control problem at each iteration and updating the control function to minimize the cost functional.

The control function $u_h$ is updated at each iteration using the adjoint variable $p_h$ and the relaxation parameter $\tau$. The updated control function is given by

$$u_{h+1} = u_h + \tau \nabla J(u_h, w_h),$$

where $\nabla J(u_h, w_h)$ is the gradient of the cost functional with respect to the control function. The adjoint variable $p_h$ is updated at each iteration using the adjoint equation

$$\nabla^T p_h = \frac{\partial J(u_h, w_h)}{\partial u_h}.$$

The iteration process continues until the convergence criterion is met.

The error in the control function $\varepsilon_u$ and the error in the cost functional $\varepsilon_J$ are defined as

$$\varepsilon_u = \left\| u_{h+1} - u_h \right\|_2,$$

and

$$\varepsilon_J = \left| J(u_{h+1}, w_h) - J(u_h, w_h) \right|.$$

The errors are typically small for the solution to be considered acceptable. The errors are influenced by the relaxation parameter $\tau$ and the number of iterations. A larger relaxation parameter and a larger number of iterations tend to decrease the errors. The errors can be computed using the updated control function $u_{h+1}$ and the updated cost functional $J(u_{h+1}, w_h)$.

The errors $\varepsilon_u$ and $\varepsilon_J$ are small in any case, and the solution is considered acceptable. The errors are influenced by the relaxation parameter $\tau$ and the number of iterations. A larger relaxation parameter and a larger number of iterations tend to decrease the errors. The errors can be computed using the updated control function $u_{h+1}$ and the updated cost functional $J(u_{h+1}, w_h)$. The errors are typically small for the solution to be considered acceptable.
$Q_2, Q_3$ emitting certain substances, while $D$ is an observation area. Our goal is to provide an example of a systematical method to control emission by some release sources so that the concentration of the eluted substance over a certain area $D$ is acceptable (i.e., at a desired level), taking into account advection field and diffusivity conditions in a stationary frame (an unsteady frame is provided in Section 8.9). In Figure 8.2 we report three different functions describing the diffusivity $\nu$ considered in our test cases. We consider the state equation

$$\nabla \cdot (\nu \nabla \mathbf{V}) = \mathbf{Q}$$

(8.2.4), where $\nu$ can also be dependent on $\mathbf{V}$, describing the mean motion of the fluid, instead
of exactly representing pointwise velocity. Diffusivity $\nu$ may depend also on the problem type and domain geometry ($x$ coordinate). We assume $u = \sum_{i=1}^{3} u_i m_{Q_i}$, being $m_{Q_i}$ the characteristic function of the $i$-th source and $u_i$ the rate of emission from the $i$-th source. Then quantity (8.2.6) can be written as:

$$F(\varphi; u) = \sum_{i=1}^{3} \int_{Q_i} u_i \varphi \, dQ_i.$$ 

Solving the control problem with the iterative method, at each step we update control function in the following way:

$$\delta u^k = \sum_{i=1}^{3} \delta u_i^k = - \tau \sum_{i=1}^{3} p m_{Q_i}.$$

### 8.5 Control input: variable emission rates

As first test we consider a two component control input parameter $\mu_u = \{u_1, u_2\}$. We fix a total emission value $u_{tot} = 27 \text{ mgs}^{-1}$ corresponding to a certain release of substances in the physiological fluid, for example. Therefore we write $u_3 = 27 - u_1 - u_2$ and $\mu_u \in \mathcal{D}$, where $\mathcal{D} = [0, 27] \times [0, 27]$ with $u_1 - u_2 \leq 27$.

#### 8.5.1 Parametrized state equation

To compute the reduced basis approximation of the state variable (8.3.3), we need to find the $N$ unknown components $w_{N}(\mu_u)$ by solving the problem:

$$a_h(w_{N}(\mu_u), \zeta_i) = F_h(\zeta_i) \quad i = 1, ..., N,$$

where

$$a_h(w_{N}(\mu_u), \zeta_i) = \sum_{j=1}^{N} \sum_{i=1}^{N} w_{N}(\mu_u) \left[ \sum_{K \in T_h} \int_{K} \nu \nabla \zeta_j \cdot \nabla \zeta_i + \sum_{K \in T_h} \int_{K} (\mathbf{v} \cdot \nabla \zeta_j) \zeta_i \right] + \sum_{K \in T_h} \delta K \frac{h_K}{\|\mathbf{V}\|} \int_{K} \left( - \nabla \nu \cdot \nabla \zeta_j + \mathbf{V} \cdot \nabla \zeta_j \right) \left( \mathbf{V} \cdot \nabla \zeta_i \right),$$

and

$$F_h(\zeta_i) = \sum_{i=1}^{N} \left[ \sum_{K \in T_h} \int_{K} u(\mu_u) \zeta_i + \sum_{K \in T_h} \delta K \frac{h_K}{\|\mathbf{V}\|} \int_{K} u(\mu_u) \left( \mathbf{V} \cdot \nabla \zeta_i \right) \right].$$

We define now the following matrices, related to the diffusive term, convective term and stabilization term:

$$C_{i,j} = \sum_{K \in T_h} \int_{K} \nu \nabla \zeta_j \cdot \nabla \zeta_i,$$

$$B_{i,j} = \sum_{K \in T_h} \int_{K} (\mathbf{V} \cdot \nabla \zeta_j) \zeta_i,$$
The term $\nu_x$ has only the component $\nu_x$, being $\nu = \nu(x)$. The matrix $A_N$ (8.3.5) can thus be written as:

$$A_{N,i,j} = C_{i,j} + B_{i,j} + S_{i,j}.$$ (8.5.6)

In this simple and particular case, $A_N$ is independent of input parameter $\mu_u$.

Let us define also the following column vectors:

$$G_i(\mu_u) = \sum_{K \in \mathcal{T}_h} \int_K u(\mu_u) \zeta_i,$$ (8.5.7)

$$H_i(\mu_u) = \sum_{K \in \mathcal{T}_h} \delta_K \frac{h_K}{\| \nabla \|} \int_K u(\mu_u) \left( \mathbf{V} \cdot \nabla \zeta_i \right),$$ (8.5.8)

so that $F_N$ is:

$$F_N(\mu_u) = G_i(\mu_u) + H_i(\mu_u).$$

Both $G$ and $H$ depend on $\mu_u$ because $u_i = \mu_{u,i}$ on the $i$-th source and set to zero over the remaining part of the domain. This assembling procedure of the right-hand-side is very fast.

The unknown vector $\mathbf{w}_N(\mu_u)$ is the solution of the system:

$$A_N \mathbf{w}_N(\mu_u) = F_N(\mu_u).$$ (8.5.9)

### 8.5.2 Parametrized adjoint equation

The $N$ unknown components $p_{N,j}(\mu_u)$ for the reduced basis approximation of adjoint variable (8.3.4) are the solution of the problem:

$$a^{ad}_{h,i}(p_N(\mu_u), \xi_i) = F^{ad}_{h,i}(\xi_i; w_N, u_h) \quad i = 1, \ldots, N,$$ (8.5.10)

where

$$a^{ad}_{h,i}(p_N(\mu_u), \xi_i) = \sum_{j=1}^N \sum_{i=1}^N p_{N,j}(\mu_u) \left[ \sum_{K \in \mathcal{T}_i} \int_K \nu \nabla \xi_j \cdot \nabla \xi_i + \sum_{K \in \mathcal{T}_h} \int_K \left( \nabla \nu \cdot \nabla \xi_j + \nabla \cdot \mathbf{V} \xi_j + \mathbf{V} \cdot \nabla \xi_i \right) \left( -\nabla \nu \cdot \nabla \xi_i + \mathbf{V} \cdot \nabla \xi_i \right) \right],$$ (8.5.11)

and

$$F^{ad}_{h,i}(\xi_i; w_N(\mu_u); u_h) = \sum_{i=1}^N \left[ \sum_{K \in \mathcal{T}_h} \int_K m_D(w_N(\mu_u) - z_d) \xi_i - \sum_{K \in \mathcal{T}_h} \delta_K \int_K m_D(w_N(\mu_u) - z_d) \left( -\nabla \nu \cdot \nabla \xi_i + \mathbf{V} \cdot \nabla \xi_i \right) - \sum_{K \in \mathcal{T}_h} \delta_K \int_K m_D \left( -\nabla \nu \cdot \nabla w_N(\mu_u) + \mathbf{V} \cdot \nabla w_N(\mu_u) + u_h \right) \xi_i \right].$$
Let $C$ be the matrix (8.5.4), we now introduce $B^\text{ad}$ and $S^\text{ad}$:

$$B^\text{ad}_{i,j} = \sum_{K \in T_h} \int_K \left( \mathbf{V} \cdot \nabla \xi_i \right) \xi_j = B_{j,i}, \quad (8.5.13)$$

$$S^\text{ad}_{i,j} = \sum_{K \in T_h} \delta_K \int_K \left( \nabla \nu \cdot \nabla \xi_j + \mathbf{V} \cdot \nabla \xi_j + \mathbf{V} \cdot \nabla \xi_j \right) \left( - \nabla \nu \cdot \nabla \xi_i + \mathbf{V} \cdot \nabla \xi_i \right) = S_{j,i}. \quad (8.5.14)$$

where $B_{j,i}, S_{j,i}$ are defined by (8.5.5), (8.5.6).

The matrix $A^\text{ad}$ can be rewritten as:

$$A^\text{ad}_{N,i,j} = C_{i,j} + B^\text{ad}_{i,j} + S^\text{ad}_{i,j}. \quad (8.5.18)$$

Let $G^\text{ad}, H^\text{ad}, I^\text{ad} \in \mathbb{R}^N$ be the column vectors defined as:

$$G^\text{ad}_i = \sum_{K \in T_h} \int_K m_D \xi_i, \quad (8.5.15)$$

$$H^\text{ad}_i = - \sum_{K \in T_h} \delta_K \int_K m_D \left( - \nabla \nu \cdot \nabla \xi_i + \mathbf{V} \cdot \nabla \xi_i \right), \quad (8.5.16)$$

$$I^\text{ad}_i = - \sum_{K \in T_h} \delta_K \int_K m_D \left( - \nabla \nu \cdot \nabla \xi_i + \mathbf{V} \cdot \nabla \xi_i \right) \xi_i. \quad (8.5.17)$$

All the matrices $C, B^\text{ad}, S^\text{ad}$ and the vectors $G^\text{ad}, H^\text{ad}, I^\text{ad}$ are computed off-line (i.e only once and stored), while for every new $\mu_0$ we assemble online the right-hand-side $F^\text{ad}_N$.

The unknown vector $p^\text{ad}_{N}(\mu_0)$ is the solution of system:

$$A^\text{ad}_{N,i} p^\text{ad}_{N,i}(\mu_0) = F^\text{ad}_N(\mu_0). \quad (8.5.18)$$

### 8.5.3 Some results

We report in Table 8.1 the number of basis functions (for both state and adjoint equation), the mean error on cost functional and on control function at convergence (computed on a high number of random inputs) and the computational saving in the case of $\nu = \nu_2$ and $\nu = \nu_3$. We have imposed $tol = 10^{-7}$. The computational saving compares the time needed to perform the on-line steps with the one necessary to complete a finite element simulation, using a mesh with $O(10^4)$ elements.

The number of basis functions and the saving percentage are the same for the two cases. The orders of magnitude of the errors are nearly the same. In Table 8.2 we report some details regarding only state equation: the mean $H^1$ error with respect to the finite element solution (computed on a high number of random inputs) and the computational saving for the three different diffusivity conditions. The number of basis functions for the reduced basis approximation of state problem is $N = 7$ for the three cases.

Dealing with a test case we calculated several errors on control function, on cost functional and its gradient.
Table 8.1: Control input: number of basis functions, mean errors on cost functional and control function at convergence and time saving.

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>$N$</th>
<th>Mean error on $J$</th>
<th>Mean $H^1$ error on $u$</th>
<th>Saving</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_1$</td>
<td>7</td>
<td>$1.4E-11$</td>
<td>$2.5E-5$</td>
<td>90%</td>
</tr>
<tr>
<td>$\nu_2$</td>
<td>7</td>
<td>$1.9E-12$</td>
<td>$6.1E-6$</td>
<td>90%</td>
</tr>
</tbody>
</table>

Table 8.2: Control input for state equation: mean $H^1$ error and time saving for control input.

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>mean $H^1$ error</th>
<th>Saving</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_1$</td>
<td>$1E-8$</td>
<td>96%</td>
</tr>
<tr>
<td>$\nu_2$</td>
<td>$2.1E-8$</td>
<td>94%</td>
</tr>
<tr>
<td>$\nu_3$</td>
<td>$3.7E-8$</td>
<td>90%</td>
</tr>
</tbody>
</table>

8.5.4 Example

We start from the following test case, where $\nu = \nu_2$: the upper source is emitting at 45% of $u_{tot}$ and the central one is not active, that is $\mu_u = \{1215, 0\}$ [mgs$^{-1}$]. The control problem solved by the two methods (finite element and reduced basis) leads to the optimal solution: the upper source emission rate is reduced to 3.49%, central source remains inoperating and the lower one is at 55.02% of $u_{tot}$. Figure 8.3 shows the reduced basis approximation of initial and optimal solution.

For this particular choice of $\mu_u$, control problems solved with the two methods converge both after 21 iteration. Since the number of iteration is the same (in this case), we can compare cost functional values ($J(w_N, u_N)$ and $J(w_h, u_h)$) at each iteration and check the $L^2$ error on control function trend (Figure 8.4).

Figure 8.3: Control input: initial reduced basis solution (right) and final reduced basis solution (left) of state equation. Substance concentration is in [\mu g].
8.6 Physical input: variable emission rates and velocity field

This case is more complex: we consider an input made of four component \( \mu_p = \{u_1, u_2, V_x, V_y\} \). Once again \( u_3 = 27 - u_1 - u_2 \text{ mgs}^{-1} \), having imposed \( u_1 + u_2 \leq u_{tot} \), and we fix the velocity absolute value \( |\mathbf{V}| = 1 \). We assume that the velocity field of the advection term can vary in direction of an angle ranging in \([-40^\circ, 90^\circ]\). As physical input one could also choose the “diffusivity coefficient” \( \nu \).

8.6.1 Parametrized state equation

The \( N \) unknowns \( w_{Ni}(\mu_p) \) (8.3.3) are the solution of the problem (8.5.1), where \( a_h \) and \( F_h \) are respectively defined in (8.5.2) and (8.5.3).

The matrix \( A_N \) can be written as follows:

\[
A_{Ni,j}(\mu_p) = C_{i,j} + V_x \cdot B_{x,i,j} + V_y \cdot B_{y,i,j} + S_{i,j}(\mu_p),
\]

where \( C \) is the matrix (8.5.4), \( S \) is defined by (8.5.6) and \( B_x \) and \( B_y \) are given by:

\[
B_{x,i,j} = \sum_{K \in T_h} \int_K (\mathbf{U}_x \cdot \nabla \zeta_j) \zeta_i,
\]

\[
B_{y,i,j} = \sum_{K \in T_h} \int_K (\mathbf{U}_y \cdot \nabla \zeta_j) \zeta_i,
\]

with \( \mathbf{U}_x = (1, 0) \) and \( \mathbf{U}_y = (0, 1) \); \( \mathbf{B} = V_x \cdot \mathbf{B}_x + V_y \cdot \mathbf{B}_y \), where \( \mathbf{B} \) (8.5.5) is the matrix related to the convective term.

The source term \( F_N(\mu_p) \) is given by:

\[
F_{Ni}(\mu_p) = G_i(\mu_p) + H_i(\mu_p).
\]
All the parameter-independent matrices \((C, B_x, B_y)\) are assembled in the off-line stage, while matrices \(S, G, H\), depending on the parameter in a non-affine way, must be built on-line for each new \(\mu_p\). Always in the on-line stage, matrices \(A_N(\mu_p)\) and \(F_N(\mu_p)\) are assembled.

The unknown vector \(w_N(\mu_p)\) is the solution of the system:

\[
A_N(\mu_p)w_N(\mu_p) = F_N(\mu_p).
\]  
(8.6.4)

The difference between (8.5.9) and (8.6.4) is that in the latter also \(A_N\) depends on input parameter \(\mu_p\). To improve efficiency of the assembling procedure described above we may apply the decomposition of non-affine terms by the empirical interpolation method described in Chapter 6 while assembling \(S, G\) and \(H\), specified in (8.5.6)-(8.5.8).

### 8.6.2 Parametrized adjoint equation

We want to find the weights of the combination (8.3.4), in order to have the reduced basis approximation of the adjoint variable, solution of (8.5.10). Bilinear form \(a^a_h\) and functional \(F^a_h\) are the same as in (8.5.11) and (8.5.12).

Let \(A^a_N \in \mathbb{R}^{N \times N}\) be the matrix:

\[
A^a_{N,ij}(\mu_p) = C_{i,j} + V_x \cdot B^a_{x,i,j} + V_y \cdot B^a_{y,i,j} + S^a_{i,j}(\mu_p),
\]

where \(C\) is defined by (8.5.4), \(S^a\) is the stabilization matrix (8.5.14), now parameter-dependent. The elements of matrices \(B^a_x\) and \(B^a_y\) are given by:

\[
B^a_{x,i,j} = \sum_{K \in T_h} \int_K (U_x \cdot \nabla \xi_i) \xi_j = B_{x,i,j},
\]

\[
B^a_{y,i,j} = \sum_{K \in T_h} \int_K (U_y \cdot \nabla \xi_i) \xi_j = B_{y,i,j},
\]

with \(U_x = (1, 0)\) and \(U_y = (0, 1)\), while \(B_{x,i,j}\) and \(B_{y,i,j}\) are (8.6.2) and (8.6.3). Moreover \(B^a = V_x \cdot B^a_x + V_y \cdot B^a_y\), where \(B^a\) is the matrix related to the convective term (8.5.13).

Source term \(F^a_N\) is given by:

\[
F^a_N(\mu_p) = G^a_i + H^a_i(\mu_p) + I^a_i(\mu_p).
\]

In this case, \(H^a\) (8.5.16) and \(I^a\) (8.5.17) are parameter-dependent, since they depend on velocity vector, while \(G^a\) (8.5.15) is parameter-independent. Also in this case an efficient computational procedure should be used.

Unknown vector \(p_N(\mu_p)\) is the solution of system:

\[
A^a_N(\mu_p)p_N(\mu_p) = F^a_N(\mu_p).
\]  
(8.6.5)

Even in this case, the difference between (8.5.18) and (8.6.5) is that in the latter also \(A^a_N\) depends on input parameter \(\mu_p\).
8.6.3 Some results

The use of the reduced basis method to solve both state and adjoint equations, at each step of our iterative method to solve control problem, implies several advantages from a computational point of view. In the case of control and physical input we have time savings up to $65 - 70\%$, which means that in the same time the finite element method solves just one iteration (state + adjoint equation using a mesh with $O(10^4)$ elements), the reduced basis method solves 3 iterations. This is a good result dealing with optimal control problems, which are not real-time problems, but time savings could be even improved if we can adopt a stabilization method based on terms which can be built off-line. At the present, stabilization is needed also for the reduced basis formulation; indeed, when using the non-stabilized Galerkin method, we have a “plateau” as $N \to \infty$, corresponding to the Galerkin residual evaluated for the stabilized “truth” solutions.

In our case, we can fix velocity field (the desired online value) before applying optimal control (at first iteration) and so we can build $S$ and $S^{ad}$ offline or at a step we can call “pre-online” for our optimal control problem. Note that the basis has been assembled offline considering different values for velocity field.

We report in Table 8.3 the number of basis function, the mean error on cost functional and on control function at convergence (computed on a high number of random inputs) and the computational saving (having fixed the “online” velocity field at first iteration) for $\nu = \nu_2$ and $\nu = \nu_3$, having imposed $tol = 10^{-7}$. To understand the slight difference in the number of basis functions, we need to compare the weight of the diffusive term and the one of the convective term in $A_N$ (8.6.1). As diffusivity increases, diffusive term becomes dominant and so just “few” basis functions are needed to have a good approximation of the solution depending on convective velocity. Since for $\nu = \nu_3$ diffusivity absolute value is greater, the convective term is less influent and fewer basis functions are needed.

In Table 8.4 we report some details about only state equation: the mean $H^1$ error with respect to the finite element solution (computed on a high number of random inputs) and the computational saving for the three different diffusivity conditions. The number of basis functions for the reduced basis approximation of state variable is $N = 81$ for $\nu = \nu_3$ and $N = 132$ for $\nu = \nu_2$.

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>$N$</th>
<th>Mean error on $J$</th>
<th>Mean $H^1$ error on $u$</th>
<th>Saving</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_2$</td>
<td>132</td>
<td>$0.9E - 9$</td>
<td>$2.5E - 7$</td>
<td>80%</td>
</tr>
<tr>
<td>$\nu_3$</td>
<td>81</td>
<td>$0.5E - 9$</td>
<td>$1.1E - 7$</td>
<td>80%</td>
</tr>
</tbody>
</table>

Table 8.3: Physical input: number of basis functions, mean errors on cost functional and control function, at convergence, and time saving (including also error calculation).
Table 8.4: Physical input for state equation during optimal control (velocity field is fixed at first iteration): mean $H^1$ error and time saving. Without fixing the online velocity field at first iteration computational saving are $\sim 65 - 70\%$.

8.6.4 Example

We choose the following input: $u_1 = 30\%$, and $u_2 = 40\%$, respectively, of $u_{tot}$ and advection field direction at $45^\circ$ with respect to $x$-axis. Figure 8.5 shows the initial solution and the solution at convergence, with the upper source emitting at $30.02\%$, the central one at $38.81\%$ and the lower one at $7.27\%$.

Also in this second case we choose a particular input parameter so that the control problems solved with the reduced basis method and with the finite element method converge in the same number of iteration. Once again we can compare cost functional values ($J(w_N, u_N)$ and $J(w_h, u_h)$) at each iteration and check the $L^2$ error on control function trend (Figure 8.6). We notice that till the $5^{th}$ iteration $J(w_N, u_N) < J(w_h, u_h)$, while from the $5^{th}$ till convergence $J(w_N, u_N) > J(w_h, u_h)$. The absolute value $|J(w_N, u_N) - J(w_h, u_h)|$ increases at each iteration till it sets on a constant value. At the beginning, $J(w_N, u_N) < J(w_h, u_h)$ because, as seen, reduced basis solution is always an under-estimation of the finite element solution (see Appendix $A$). After some iteration, instead, $J(w_N, u_N) > J(w_h, u_h)$ because, since reduced basis solution is an under-estimation, the source term of the adjoint problem is inferior and so $p_N(\mu) < p_h(\mu)$ on the whole domain; since $J' = p$, if $J'(u_N) < J'(u_h)$, it means that $J(w_N, u_N)$ decreases less rapidly and so $J(w_N, u_N)$ is greater.

Figure 8.5: Physical input: initial reduced basis solution (right) and final reduced basis solution (left) of state equation. Concentration is in $[\mu g]$. 
8.7 Geometrical input: parametrized domains

The goal of this section is to show how to combine optimal control problems with the calculation of geometrical sensitivity analysis and to solve a problem with a double aim: to achieve the desired concentration into a zone and to optimize the distribution of sources by positioning them in the domain. The ultimate goal is, for example, the optimization of substances release in the domain or in a subset of it.

We refer to Figure 8.7, the physical domain $\Omega \subset \mathbb{R}^2$ is divided in seven subdomains $\hat{\Omega}_r$, $r = 1, \ldots, 7$. We have chosen the parameters $\mu_\theta$ as $C_1, C_2, C_3$ and $C_4$ [cm], with $C_1 + C_2 = 3$ and $C_3 + C_4 = 3$. The position of the central source is fixed, while the position $C_3$ of the upper one and the position $C_2$ of the lower one can vary.

Moreover, we considered this kind of parametrized domain in order to be able to handle some practical situations: for example where to place a new emission source. The method is based on the affine mapping procedures from reference subdomains (the ones with $C_1 = 2$, $C_2 = 1$, $C_3 = 1$ e $C_4 = 2$) to the true ones ($\Omega^r \rightarrow \hat{\Omega}^r$). Also in this case we can extend this methodology to non-affine mappings dealing with non-Cartesian geometry.

8.7.1 Parametrized state equation

Let $R$ be the number of subdomains in which the real domain is divided: $\hat{\Omega} = \bigcup_{r=1}^R \hat{\Omega}^r$. From the weak form of the state equation (8.2.10), we define the following bilinear and linear forms:

$$
\hat{A}(\hat{w}, \hat{v}) = \sum_{r=1}^R \int_{\hat{\Omega}^r} \frac{\partial \hat{w}}{\partial \hat{x}_i} \frac{\partial \hat{v}}{\partial \hat{x}_j} d\hat{\Omega},
$$

$$
\hat{B}(\hat{w}, \hat{v}) = \sum_{r=1}^R \int_{\hat{\Omega}^r} V_i \frac{\partial \hat{w}}{\partial \hat{x}_i} d\hat{\Omega}, \quad (8.7.1)
$$
Figure 8.7: Scheme for the computational domain: subdomains and parameters.

\[
\hat{F}(\hat{v}) = \sum_{r=1}^{R} \int_{\Omega_r} \hat{u}\hat{v}d\Omega_r,
\]

\[
\hat{S}(\hat{w}, \hat{v}) = \sum_{r=1}^{R} \sum_{K \in T_h}^{R} \delta_K \frac{h_K}{|V|} \int_{K} \left( V_i \frac{\partial \hat{w}}{\partial x_i} \right) \left( V_j \frac{\partial \hat{v}}{\partial x_j} \right) d\hat{K},
\]

\[
\hat{G}(\hat{v}) = \sum_{r=1}^{R} \sum_{K \in T_h}^{R} \delta_K \frac{h_K}{|V|} \int_{K} \hat{u} \left( V_i \frac{\partial \hat{v}}{\partial x_i} \right) d\hat{K},
\]

where \(1 \leq i, j \leq d = 2\) and \(\delta_{i,j} = \hat{\nu} \delta_{i,j}\). For the sake of simplicity, we assume that diffusivity \(\nu\) and \(|V|\) are constant over the whole domain. The weak form of the stabilized state equation in the real domain can be written as:

\[
\text{find } \hat{w}_N(\mu) \in \hat{W}_N : \hat{A}(\hat{w}_N, \hat{v}) + \hat{B}(\hat{w}_N, \hat{v}) + \hat{S}(\hat{w}_N, \hat{v}) = \hat{F}(\hat{v}) + \hat{G}(\hat{v}) \quad \forall \hat{v} \in \hat{W}_N.
\]

We are interested in writing a partial differential equation depending on the set of geometrical parameters given as input. We introduce an affine mapping from the “true” subdomains \(\Omega_r\) into the corresponding \(\Omega^r\). For any \(\hat{x} \in \hat{\Omega}_r\), \(r = 1, \ldots, R\), its image \(x \in \Omega^r\) is given by:

\[
x = G^r(\mu_g; \hat{x}) = G^r(\mu)\hat{x} + g^r, \quad 1 \leq r \leq R.
\]

For convenience, we rename \(\mu_g\) with \(\mu\). We thus write:

\[
\frac{\partial}{\partial \hat{x}_i} = \frac{\partial x_j}{\partial \hat{x}_i} \frac{\partial}{\partial x_j} = G^r_{ji}(\mu) \frac{\partial}{\partial x_j}, 1 \leq i, j \leq d = 2
\]

and in the reference domain \(\Omega\) we have:

\[
A(w, v; \mu) = \sum_{r=1}^{R} \int_{\Omega_r} \frac{\partial w}{\partial x_i} \left( G^r_{ir}(\mu) \frac{\partial v}{\partial x_j} G^r_{jr}(\mu) \left( (G^r(\mu))^{-1} \right) \right) \frac{\partial v}{\partial x_j} d\Omega_r,
\]
We can now apply ane decomposition:

\[ \mathcal{E}(w, v; \mu) = \sum_{r=1}^{R} \int_{\Omega} V_i \frac{\partial w}{\partial x_i} \left( G_{ij}^{r}(\mu) |(G^{r}(\mu))^{-1} | \right) v d\Omega, \]

\[ \mathcal{F}(v; \mu) = \sum_{r=1}^{R} \int_{\Omega} \left( \hat{u} |(G^{r}(\mu))^{-1} | \right) v d\Omega, \]

\[ \mathcal{S}(w, v; \mu) = \sum_{r=1}^{R} \sum_{K \in T_h} \delta K \frac{h_K}{\|V\|} \int_{K} V_i \frac{\partial w}{\partial x_i} \left( G_{ij}^{r}(\mu) G_{jj}^{r}(\mu) |(G^{r}(\mu))^{-1} | \right) \frac{\partial v}{\partial x_j} V_j dK, \]

\[ \mathcal{G}(v; \mu) = \sum_{r=1}^{R} \sum_{K \in T_h} \delta K \frac{h_K}{\|V\|} \int_{K} \hat{u} V_i \frac{\partial v}{\partial x_i} \left( G_{ij}^{r}(\mu) |(G^{r}(\mu))^{-1} | \right) dK, \]

for \( 1 \leq i, j \leq 2, \ r = 1, ..., R. \)

The transformation tensors for bilinear forms are defined as follows:

\[ \nu_{ij}^{r}(\mu) = G_{ij}^{r}(\mu) \tilde{v}_{ij}^{r}(\mu) |(G^{r}(\mu))^{-1} |, \ 1 \leq i, j \leq 2, \ r = 1, ..., R, \]

\[ \chi_{ij}^{r}(\mu) = G_{ij}^{r}(\mu) G_{jj}^{r}(\mu) |(G^{r}(\mu))^{-1} | = \frac{\nu_{ij}^{r}}{\nu_{jj}^{r}}, \ 1 \leq i, j \leq 2, \ r = 1, ..., R, \]

where \( \nu^{r} \) is the diffusivity in the reference subdomains. For the linear forms we define:

\[ \chi_{ij}^{r}(\mu) = G_{ij}^{r}(\mu) |(G^{r}(\mu))^{-1} |, \ 1 \leq i, j \leq 2, \ r = 1, ..., R. \]

Furthermore, we may define:

\[ \sigma^{q(i,j,r)}(\mu) = \nu_{ij}^{r}(\mu), \ \mathcal{A}^{q(i,j,r)}(w, v) = \int_{\Omega} \frac{\partial w}{\partial x_i} \frac{\partial v}{\partial x_j} d\Omega, \]

\[ \Phi^{s(i,r)}(\mu) = \chi_{ij}^{r}(\mu), \ \mathcal{B}^{s(i,r)}(w, v) = \int_{\Omega} V_i \frac{\partial w}{\partial x_i} v d\Omega, \]

\[ \Upsilon^{q(i,j,r)}(\mu) = \chi_{ij}^{r}(\mu), \ \mathcal{S}^{q(i,j,r)}(w, v) = \sum_{K \in T_h} \delta K \frac{h_K}{\|V\|} \int_{K} V_i \frac{\partial w}{\partial x_i} V_j dK, \]

\[ \mathcal{G}^{s(i,r)}(v) = \sum_{K \in T_h} \delta K \frac{h_K}{\|V\|} \int_{K} \hat{u} V_i \frac{\partial v}{\partial x_i} dK, \]

for \( 1 \leq i, j \leq 2, \ r = 1, ..., R, \) with \( q \) and \( s \) “condensed” indexes for combinations of \( i, j, r \) and \( i, r. \)

We can now apply affine decomposition:

\[ \mathcal{A}(\sigma(\mu), w, v) = \sum_{q=1}^{Q^a} \sigma^{q(\mu)} \mathcal{A}^{q}(w, v), \ \mathcal{B}(\Phi(\mu), w, v) = \sum_{s=1}^{Q^b} \Phi^{s(\mu)} \mathcal{B}^{s}(w, v), \]

\[ \mathcal{S}(\Upsilon(\mu), w_N, v) = \sum_{q=1}^{Q^c} \Upsilon^{q(\mu)} \mathcal{S}^{q}(w_N, v), \ \mathcal{G}(\Phi(\mu), v) = \sum_{s=1}^{Q^b} \Phi^{s(\mu)} \mathcal{G}^{s}(v), \]

where \( \max(Q^a) = d \times d \times R \) and \( \max(Q^b) = d \times R. \)

The reduced basis approximation of the stabilized state equation in the reference domain \( \Omega \) reads: find \( w_N(\mu) \in W_N \) such that

\[ \mathcal{A}(w_N, v; \mu) + \mathcal{B}(w_N, v; \mu) + \mathcal{S}(w_N, v; \mu) = \mathcal{F}(v; \mu) + \mathcal{G}(v; \mu), \ \forall v \in W_N. \]
8.7.2 Parametrized adjoint equation

Let $\hat{A}$ be the form (8.7.1); from the weak formulation of the adjoint problem (8.2.11), we define:

$$\mathcal{B}^{ad}(\hat{p}, \hat{v}) = \sum_{r=1}^{R} \int_{\Omega_r} \hat{p} V_i \frac{\partial \hat{v}}{\partial x_i} d\Omega = \hat{B}(\hat{v}, \hat{p}),$$

$$\tilde{\mathcal{F}}^{ad}(\hat{v}) = \sum_{r=1}^{R} \int_{\Omega_r} \hat{m}_D (\hat{w} - \hat{z}_d) \hat{v} d\Omega,$$

$$\tilde{\mathcal{S}}^{ad}(\hat{p}, \hat{v}) = \sum_{r=1}^{R} \sum_{K \in T_h^r} \delta_K \int_K (V_i \frac{\partial \hat{p}}{\partial x_i}) (V_j \frac{\partial \hat{v}}{\partial x_j}) dK,$$

$$\tilde{\mathcal{G}}^{ad}(\hat{v}) = -\sum_{r=1}^{R} \sum_{K \in T_h^r} \delta_K \int_K \hat{m}_D (\hat{w}_N - \hat{z}_d) (V_i \frac{\partial \hat{v}}{\partial x_i}) dK,$$

$$\tilde{\mathcal{H}}^{ad}(\hat{v}) = -\sum_{r=1}^{R} \sum_{K \in T_h^r} \delta_K \int_K \hat{m}_D (V_i \frac{\partial \hat{w}_N}{\partial x_i}) \hat{v} dK,$$

for $1 \leq r \leq R, 1 \leq i, j \leq 2$, where $\hat{B}$ is defined by (8.7.1).

The reduced basis approximation of the stabilized adjoint problem in the real domain is: find $\hat{p}_N(\mu) \in \tilde{Z}_N$ such that:

$$\hat{A}(\hat{p}_N, \hat{v}) + \mathcal{B}^{ad}(\hat{p}_N, \hat{v}) + \tilde{\mathcal{S}}^{ad}(\hat{p}_N, \hat{v}) = \tilde{\mathcal{F}}^{ad}(\hat{v}) + \tilde{\mathcal{G}}^{ad}(\hat{v}) + \tilde{\mathcal{H}}^{ad}(\hat{v}) \forall \hat{v} \in \tilde{Z}_N.$$

In the reference domain, we have:

$$\mathcal{B}^{ad}(p, v; \mu) = \sum_{r=1}^{R} \int_{\Omega_r} V_i \frac{\partial v}{\partial x_i} \left( G_{iv}(\mu) \left[ (G^v(\mu))^{-1} \right] \right) p d\Omega,$$

$$\mathcal{F}^{ad}(v; \mu) = \sum_{r=1}^{R} \int_{\Omega_r} \left( \hat{m}_D (\hat{w} - \hat{z}_d) \left[ (G^v(\mu))^{-1} \right] \right) v d\Omega,$$

$$\mathcal{S}^{ad}(p, v; \mu) = \sum_{r=1}^{R} \sum_{K \in T_h^r} \delta_K \int_K V_i \frac{\partial p}{\partial x_i} \left( G_{iv}(\mu) G_{jj}(\mu) \left[ (G^v(\mu))^{-1} \right] \right) \frac{\partial v}{\partial x_j} V_j dK,$$

$$\mathcal{G}^{ad}(v; \mu) = -\sum_{r=1}^{R} \sum_{K \in T_h^r} \delta_K \int_K \hat{m}_D (\hat{w}_N - \hat{z}_d) \left( V_i \frac{\partial v}{\partial x_i} \left[ (G^v(\mu))^{-1} \right] \right) dK,$$

$$\mathcal{H}^{ad}(v; \mu) = -\sum_{r=1}^{R} \sum_{K \in T_h^r} \delta_K \int_K \hat{m}_D \left( V_i \frac{\partial \hat{w}_N}{\partial x_i} \left[ (G^v(\mu))^{-1} \right] \right) v dK.$$

We introduce:

$$\Phi_s^{(i,r)}(\mu) = \chi_s^{(i)}(\mu), \quad B_{ad}^{(i,r)}(p, v) = \int_{\Omega_r} V_i \frac{\partial v}{\partial x_i} p d\Omega,$$
\[ S_{ad}^{(i,j,r)}(p_N, v) = \sum_{K \in T_h} \delta_K \int_K V_i \frac{\partial p_N}{\partial x_i} \frac{\partial v}{\partial x_j} V_j \, dK, \quad \forall v \in Z_N, \]

\[ G_{ad}^{(i,r)}(v) = -\sum_{K \in T_h} \delta_K \int_K \hat{m}_D \left( w_N - \hat{z}_d \right) \left( V_i \frac{\partial v}{\partial x_i} \right) dK, \]

\[ H_{ad}^{(i,r)}(v) = -\sum_{K \in T_h} \delta_K \int_K \hat{m}_D \left( V_i \frac{\partial w_N}{\partial x_i} \right) v \, dK, \]

for \( 1 \leq i, j \leq 2, 1 \leq r \leq R \), where \( \chi_\alpha^r(\mu) \) is defined by (8.7.1).

By the definitions above, we write:

\[ B^{ad}(\Phi(\mu), p, v) = \sum_{s=1}^{Q_b} \Phi^s(\mu) B_{ad}^s(p, v), \quad S^{ad}(\Psi(\mu), p, v) = \sum_{q=1}^{Q_c} \Psi^q(\mu) S_{ad}^q(p, v), \]

\[ G^{ad}(\Phi(\mu), v) = \sum_{s=1}^{Q_b} \Phi^s(\mu) G_{ad}^s(v), \quad H^{ad}(\Phi(\mu), v) = \sum_{s=1}^{Q_b} \Phi^s(\mu) H_{ad}^s(v), \]

where \( \Psi^q(\mu) \in \Phi^s(\mu) \) are respectively defined in (8.7.3) and (8.7.2).

The reduced basis approximation of the stabilized adjoint equation in the reference domain \( \Omega \) is: \textit{find} \( p_N(\mu) \in Z_N \text{ such that } \)

\[ A(p_N, v) + B^{ad}(p_N, v) + S^{ad}(p_N, v) = F^{ad}(v) + G^{ad}(v) + H^{ad}(v) \quad \forall v \in Z_N. \]

At this point we solve a \textit{parametrized optimal control problem}.

### 8.7.3 Geometrical sensitivity analysis

We fix the source release rates \((u_1 = 20\% \text{ of } u_{tot} = 27 \text{ mgs}^{-1}, u_2 = 5\% \text{ and so } u_3 = 75\%)\), so that the only variable parameters will be the geometrical quantities in order to get important information on geometrical parameters sensitivity. With the adaptive assembling procedure of Section 5.4, we find 35 basis functions.

We report in Figure 8.8 the result of the analysis for the parameterized domain of Figure 8.7. We note that, fixing \( C_3 \) (i.e. for a determined upper source position), substance concentration over the observation zone decreases when \( C_2 \) goes from 0.1 to 1.3, while for \( 1.3 \leq C_2 \leq 2.9 \) there are no important variations. This is due to the fact that for \( C_2 \geq 1.3 \) the observation area is out of the lower emission zone. If we keep constant the position \( C_2 \) of the lower source, we note that the concentration over the observation zone increases rapidly till it reaches its maximum around \( C_3 = 1 \) and then it decreases till \( C_3 = 2.9 \). The explanation of this behavior is that for \( C_3 = 0.1 \) the observation zone is partially inside the emission wake, while for \( C_3 = 1 \) it’s totally inside of it. For \( C_3 \geq 1 \) the zone comes gradually out of the emission wake.
8.7.4 Sensitivity analysis applied to control problem

The choice of geometrical inputs allows us to study the state variable sensitivity to domain variations, to know how the concentration over the observation area varies according to the geometrical changes. We want now to exploit this sensitivity analysis in solving the control problem, in order to optimize the source emission level while keeping the concentration level at an optimal level.

At this purpose we modify the iterative method adopted to solve the control problem. Starting always from an initial value for control variable \( u_0 \), we solve state and adjoint equations. Once \( p \) is known, we check if the stopping criterium (8.2.9) is satisfied. At every iterative step in which the adjoint variable does not satisfy the stopping criterium, instead of starting directly an iterative process on control variable \( u \), we try first to fulfill the criterium only by varying the positions of upper and lower sources, for example. If we cannot satisfy it just by modifying the geometry, then we update control variable value adopting the steepest descent method (8.2.3). In this way we minimize the number of iteration on \( u \). Figure 8.9 clarifies this new combined technique for solving the control problem.

The number of input parameters for this test problem is six: the emission rates of the two sources and the geometrical parameters of our model, i.e. \( \mu = \{u_1, u_2, C_1, C_2, C_3, C_4\} \). The set of parameters \( \mu \) belongs to \( D = [0, 27] \times [0, 27] \times [0.1, 2.9] \times [0.1, 2.9] \times [0.1, 2.9] \times [0.1, 2.9] \). We notice that for a number of basis functions \( N = 80 \), the reduced basis solution of the control problem is a “good” approximation of the finite element solution, i.e. the mean \( H^1 \) error for random input is about \( 10^{-5} \). To verify that the control problem solved with the finite element method and the one solved with the reduced basis method converged to the same solution, we compute the two errors \( \varepsilon_u \) and \( \varepsilon_J \) for a certain number of random input.

Figure 8.8: Variations of substance concentration (in \( \mu g \)) over the observation zone when the upper (\( C_3 \)) and lower (\( C_2 \)) source positions change.
Imposing $tol = 10^{-7}$, we find that the order of magnitude for $\varepsilon_u$ mean value is $10^{-5}$, while for $\varepsilon_J$ mean value is $10^{-10}$.

For each of the random input we solve also the simple control problem without geometrical sensitivity, that is the one which iterates only on control function: on an average the sensitivity analysis increases the diffusivity of the substances emitted of around 15% on the total.

Figure 8.9: Scheme of the iterative process to solve control problem combined with sensitivity analysis.

8.8 Some concluding remarks

We found out that adopting the reduced basis method for the solution of both state and adjoint equations, at every steps of the iterative method used to solve the optimal control problem, implies important advantages from a computational point of view. In the analyzed
cases we found time savings up to 70%.

The reduced basis method shows also in this case great versatility. In fact we analyzed different input parameter classes:

- **control input** - in our application this kind of input is related to the emissions from sources, for example;
- **physical input** - this kind of input allows us to take into account advection and diffusivity conditions;
- **geometrical input** - in order to consider the position of the emitting sources with respect to the observation zone.

With all these different parametrization we are able to solve the control problem from a global and decisional point of view, being able to consider, at the same time, several aspects of the same problem. All the extensions to reduced basis methods introduced in the previous chapters (Stokes and Navier-Stokes problems in domains with also non-affine parametric dependence) are possible.

An other interesting aspect to consider in the efficient resolution of optimal control problems is the mesh adaptivity and the influence of the mesh on the accuracy of cost functional estimation. This aspect is developed in the work by Dedè and Quarteroni [32].

A further extension is the development of a posteriori error estimation in the reduced basis context dealing with problem with non-affine geometrical parametric dependence.

### 8.9 Appendix: reduced basis for time-dependent problem

In this appendix we show very briefly how to solve an unsteady problem with reduced basis considering also time as a parameter. An unsteady problem has been recently studied using reduced basis method by Grepl in [44] and before by Ito and Ravindran in [64].

Indicating with $w$ the concentration of a substance, with $V$ a velocity field (for example given by a Navier-Stokes simulation in the same domain) and with $\nu$ the diffusivity coefficient, we consider the following state equation in $\Omega$ completed with properly initial and boundary condition:

$$\begin{cases}
\frac{\partial w}{\partial t} - \nu \Delta w + V \cdot \nabla w = \beta(x)u(t), \\
w = 0 \text{ on } \Gamma_D, \\
\nabla w \cdot \mathbf{n} = 0 \text{ on } \Gamma_N, \\
y(t = 0) = 0.
\end{cases} \tag{8.9.1}$$

The right-hand-side is built as a product of two functions: $\beta$, depending on spacial coordinates, indicating for example the zone of emission of a source and $u$, time dependent, indicating an emission rate in time. Introducing $H^1_{\Gamma_D} = \{ v \in H^1 : v|_{\Gamma_D} = 0 \}$, we can write the weak formulation for the problem (8.9.1):

$$\forall \ 0 < t \leq 1 \ \text{find } w(t) \in H^1_{\Gamma_D} : \ a(w(t), \varphi) = F(\varphi, u(t)), \ \forall \varphi \in H^1_{\Gamma_D},$$
§8.9. APPENDIX: REDUCED BASIS FOR TIME-DEPENDENT PROBLEM

where
\[ a(w(t), \varphi) := \int_{\Omega} \frac{\partial w}{\partial t} \varphi \, d\Omega + \int_{\Omega} \nu \nabla w \cdot \nabla \varphi \, d\Omega + \int_{\Omega} \mathbf{V} \cdot \nabla w \varphi \, d\Omega, \]
\[ F(\varphi, u) := \int_{\Omega} \beta(x)u(t)\varphi \, d\Omega. \]

We approximate the temporal derivative with:
\[ \frac{\partial w}{\partial t} = \frac{w^k - w^{k-1}}{\Delta t}, \]
where \( w^k = w(t^k = k \cdot \Delta t) \), for \( k = 1, ..., N_k \) with \( N_k = 1/\Delta t \). Using a strategy based on backward Euler ([127]) we find the following discretized problem:

\[
\text{for } 1 \leq k \leq N_k \text{ find } w^k \in H^1_{\Gamma_D} : \quad a^k(w^k, \varphi) = F^k(\varphi, u^k), \quad \forall \varphi \in H^1_{\Gamma_D},
\]

where
\[
a^k(w^k, \varphi) := \Delta t^{-1} \int_{\Omega} w^k \varphi \, d\Omega + \int_{\Omega} \nu \nabla w^k \cdot \nabla \varphi \, d\Omega + \int_{\Omega} \mathbf{V}^k \cdot \nabla w^k \varphi \, d\Omega,
\]
\[
F^k(\varphi, u^k) := \int_{\Omega} \beta(x)u^k\varphi \, d\Omega + \Delta t^{-1} \int_{\Omega} w^{k-1} \varphi \, d\Omega,
\]
and \( \mathbf{V}^k \) is the velocity field at the instance \( t^k \), while \( u^k = u(t^k) \).

For values of \( \nu \in [0.1, 10] \) we note that the local Péclet number is always greater than unity and we need stabilization procedure (see [127]). Indicating with \( w^k \in X_h \) the Galerkin-Finite Element approximation at instance \( t^k \), the problem becomes the following:

\[
\text{for } 1 \leq k \leq N_k \text{ find } w^k \in X_h : \quad a^k(w^k, \varphi_h) + L_h^{(\rho)}(w^k, f, \varphi_h) = F^k(\varphi, u^k), \quad \forall \varphi_h \in X_h.
\]

so that \( L_h^{(\rho)}(w^k, f, \varphi_h) = 0, \forall \varphi_h \in X_h \). The formulation for \( L_h^{(\rho)} \) is:

\[
L_h^{(\rho)}(w_h, f, \varphi_h) = \sum_{K \in T_h} \delta_K(Aw_h - f, S_h^{(\rho)}(\varphi_h))_{L^2(K)}, \tag{8.9.2}
\]

where \( Aw = \frac{\partial w}{\partial t} - \nu \Delta w + \mathbf{V} \cdot \nabla w \) and \( f = \beta(x)u(t) \), while
\[
S_h^{\rho} = \frac{h_K}{\| \mathbf{V} \|} [A_{SS} \varphi_h + \rho \cdot A_s \varphi_h],
\]
with \( A_{SS} \varphi_h \) and \( A_s \varphi_h \) we indicate, respectively, the antisymmetric part and the symmetric part of the operator \( A \varphi_h \). We decide to use SUPG method so that \( \rho = 0 \) and we consider only the following antisymmetric part of the operator \( A \) which is:

\[
A_{SS} \varphi_h = \frac{1}{2} (\nabla \cdot (\mathbf{V} \varphi_h) + \mathbf{V} \cdot \nabla \varphi_h) = \frac{1}{2} (\nabla \cdot \mathbf{V} + \mathbf{V} \cdot \nabla \varphi_h + \mathbf{V} \cdot \nabla \varphi_h) = \mathbf{V} \cdot \nabla \varphi_h,
\]
and the stabilized problem becomes:

\[
\text{for } 1 \leq k \leq N_k \text{ find } w^k \in X_h : \quad a^k(w^h, \varphi_h) + s_h(w^k, \varphi_h; u^k) = F^k(\varphi_h; u^k), \quad \forall \varphi_h \in X_h,
\]

where \( s_h \) is the stabilization function.
or its equivalent formulation:

\[ s_h(w^k_h, \varphi_h; u^k) = \sum_{K \in T_h} \delta_K \frac{h_K}{\| \mathbf{V} \|} \int_K (Aw^k_h - f)(A_{SS}\varphi_h) \, dK \]

for \( 1 \leq k \leq N_k \) find \( w^k_h \in X_h \) : \( a^k_h(w^k_h, \varphi_h) = F^k_h(\varphi_h; u^k), \forall \varphi_h \in X_h \). \hspace{1cm} (8.9.3)

where

\[
a^k_h(w^k_h, \varphi_h) := a^k(w^k_h, \varphi_h) + \sum_{K \in T_h} \delta_K \frac{h_K}{\| \mathbf{V} \|} \int_K (\Delta t^{-1}w^k_h + \mathbf{V}^k \cdot \nabla w^k_h)(\mathbf{V}^k \cdot \nabla \varphi_h) \, dK,
\]

\[
F^k_h(\varphi_h; u^k) := F(\varphi_h; u^k) + \sum_{K \in T_h} \delta_K \frac{h_K}{\| \mathbf{V} \|} \left[ \int_K (\Delta t^{-1}w^k_{h,-1})\mathbf{V}^k \cdot \nabla \varphi_h) \, dK + \int_K \beta_h u^k(\mathbf{V}^k \cdot \nabla \varphi_h) \, dK \right],
\]

where \( h_K \) is the diameter of \( K \), \( \delta_K \) is a parameter depending on the local Péclet number, \( \beta_h \) is the discrete approximation of \( \beta(x) \) function and \( X_h \subset H^1_{TD} \) is the finite element space (for example \( \mathbb{P}^1 \)). The contribution to stabilization is zero on the elements where \( \mathbf{V}^k \) is zero. The output of interest (concentration) into a subdomain \( \Omega_s \subset \Omega \) is:

\[
s(\nu, t) = |\Omega_s|^{-1} \int_{\Omega_s} w(\nu, t) \, d\Omega.
\]

(8.9.4)

### 8.9.1 Reduced basis formulation

We consider the input made up of the following vector of parameters: \( \mu = (\nu, t^f) \), where \( t^f \in [0, 1] \) is the time instance of which we want to know the solution, \( \mu \in D \subset \mathbb{R}^2 \), \( D = [0,1,1] \times [0,1] \).

The reduced basis solution of problem (8.9.3) is indicated with \( y_N(\mu) \) and given by:

\[
w_N(\nu, t^k) = \sum_{j=1}^N w^k_{N_j}(\mu) \zeta_j = \sum_{j=1}^N w^k_{N_j}(\mu) \sum_{i=1}^N \zeta^h_i \varphi_i.
\]

(8.9.5)

The \( N \) unknowns \( w^k_{N_j}(\mu) \) are calculated by solving the following problem:

\[
a^k_h(w_N(\mu), \zeta; \mu) = F^k_h(\zeta_h) \quad 1 \leq h \leq N_e \quad 1 \leq k \leq N_k,
\]

where

\[
a^k_h(w_N, \zeta; \mu) = \sum_{j=1}^N \sum_{h=1}^N w^k_{N_j}(\mu) \sum_{i=1}^N \zeta^h_i \varphi_i \sum_{l=1}^N \int_K (\mathbf{V}^k \cdot \nabla \varphi_i) \varphi_l + \sum_{K \in T_h} \delta_K \frac{h_K}{\| \mathbf{V} \|} \int_K (\Delta t^{-1}\varphi_i + \mathbf{V}^k \cdot \nabla \varphi_i)(\mathbf{V}^k \cdot \nabla \varphi_l)
\]

\[
\sum_{K \in T_h} \int_K (\mathbf{V}^k \cdot \nabla \varphi_i) \varphi_l + \sum_{K \in T_h} \delta_K \frac{h_K}{\| \mathbf{V} \|} \int_K (\Delta t^{-1}\varphi_i + \mathbf{V}^k \cdot \nabla \varphi_i)(\mathbf{V}^k \cdot \nabla \varphi_l)
\]
and
\[ F_k^h(\zeta_h) = \sum_{j=1}^{N} \sum_{h=1}^{N} w_{N_j}^{k-1} \sum_{i=1}^{N} \sum_{l=1}^{N} \zeta_i^j \zeta_l^h \left[ \Delta t^{-1} \sum_{K \in T_h} \varphi_i \varphi_l + \Delta t^{-1} \sum_{K \in T_h} \delta_K \frac{h_K}{\|V^k\|} \int_K \varphi_l (V^k \cdot \nabla \varphi_l) \right] + \sum_{h=1}^{N} \sum_{l=1}^{N} \zeta_i^j \left[ \int_K \beta_h u^k \varphi_i + \sum_{K \in T_h} \delta_K \frac{h_K}{\|V^k\|} \int_K \beta_h u^k (V^k \cdot \nabla \varphi_l) \right]. \]

We define \(M, C, B, S\) as mass matrix, stiffness matrix, transport related matrix and stabilization one, respectively:
\[
M_{i,i} = \sum_{K \in T_h} \int_K \varphi_i \varphi_i, \\
C_{i,i} = \sum_{K \in T_h} \int_K \nabla \varphi_i \cdot \nabla \varphi_i, \\
B_{i,i} = \sum_{K \in T_h} \int_K (V^k \cdot \nabla \varphi_i) \varphi_i, \\
S_{i,i}(\mu) = \sum_{K \in T_h} \delta_K \frac{h_K}{\|V^k\|} \int_K (\Delta t^{-1} \varphi_i + V^k \cdot \nabla \varphi_i)(V^k \cdot \nabla \varphi_i).
\]

We introduce matrix \(A_h\) given by:
\[
A_{h_{i,i}}(\mu) = \Delta t^{-1} \cdot M_{i,i} + \nu \cdot C_{i,i} + B_{i,i} + S_{i,i}(\mu).
\]

To write also the right-hand-side we define \(D e G, H\):
\[
D_{i,i}(\mu) = \sum_{K \in T_h} \delta_K \frac{h_K}{\|V^k\|} \int_K \varphi_i (V^k \cdot \nabla \varphi_i), \\
G_i = \sum_{K \in T_h} \int_K \beta_h u^k \varphi_l, \\
H_i(\mu) = \sum_{K \in T_h} \delta_K \frac{h_K}{\|V^k\|} \int_K \beta_h u^k (V^k \cdot \nabla \varphi_l),
\]
\(S, D\) and \(H\) are depending on \(\mu\), because \(\delta_K\) is depending on the Péclet number (an on \(\nu\)) and they should be calculated online. All the other quantities are calculated offline.

The reduced basis solution \(u_N^k(\mu)\) is obtained by solving the following problem:
\[
(W^T A_h(\mu) W) w_N^k(\mu) = \Delta t^{-1} \cdot [W^T (M + D(\mu)) W] w_N^{k-1} + W^T (G + H(\mu)), \tag{8.9.6}
\]
where \(w_N^{k-1}\) is known by the solution at the previous time step \(t = t^{k-1}\).

From solution (8.9.5) we can calculate the output approximation:
\[
s_N^k(\nu, t^k) = |\Omega_s|^{-1} \int_{\Omega_s} w_N^k(\mu) d\Omega = |\Omega_s|^{-1} \sum_{j=1}^{N} w_{N_j}^k(\mu) \sum_{i=1}^{N} \zeta_i^j \int_{\Omega_s} \varphi_i d\Omega. \tag{8.9.7}
\]


Basis assembling

As basis function we take solutions calculated at different time steps ($\delta t = 0.1$ s) and at different values of $\nu = 0.1, 0.3, 0.5, 0.7, 1$ as shown in Figure 8.10. With this kind of basis the relative $H^1$ max error for a great number of random configurations during the optimized assembling process is shown in Figure 8.11.

Figure 8.10: Basis construction.

Figure 8.11: Max $H^1$ relative error during basis assembling and optimization.

Error on solutions

We can calculate the relative $H^1$ error considering a great number of different random configuration and using the basis assembled previously. Figure 8.12 shows max and mean error varying $N$. The same calculation is carried out considering the error on outputs and it is shown in Figure 8.13.
Figure 8.12: Max and mean $H^1$ relative error on solutions.

Figure 8.13: Max and mean $H^1$ relative error on output.
Appendix A

An Introduction to Reduced Basis Methods: Elliptic Equations in Parametrized Domains

In this appendix we review several aspects of reduced basis methods and we deal for simplicity with multi-parametrized subdomains in the elliptic case (introduced also in [135]). The essential components of the method are (i) (provably) rapidly convergent global reduced-basis approximations — Galerkin projection onto a space $W_N$ spanned by solutions of the governing partial differential equation at $N$ selected points in parameter space; (ii) a posteriori error estimation — relaxations of the error-residual equation that provide inexpensive bounds for the error in the outputs of interest; and (iii) off-line/on-line computational procedures — methods which decouple the generation and projection stages of the approximation process. The operation count for the on-line stage — in which, given a new parameter value, we calculate the output of interest and associated error bound — depends only on $N$ (typically very small) and the parametric complexity of the problem; the method is thus ideally suited for the repeated and rapid evaluations required in the context of parameter estimation, design, optimization, and real-time control. In [118] a rigorous a posteriori error bound framework for reduced-basis approximations of elliptic coercive equations is developed. The resulting error estimates are, in some cases, quite sharp: the ratio of the estimated error in the output to the true error in the output, or effectivity, is close to (but always greater than) unity. We use a posteriori bound error estimator applied also to an adaptive procedure in choosing the approximation space and its dimension, minimizing the estimated error or the effectivity [161].

A.1 Overview on reduced basis

The optimization, control, design and characterization of an engineering component or system requires, as seen also in the first part of the thesis, the prediction of certain “quantities of interest,” or performance metrics, which we shall denote outputs — for example velocity field, maximum stresses, maximum temperatures, heat transfer rates, flow rates, vorticity, or lifts
and drags. These outputs are typically expressed as functionals of field variables associated with a parametrized partial differential equation which describes the physical behavior of the component or system. The parameters, which we shall denote inputs, serve to identify a particular “configuration” of the components: these inputs may represent design or decision variables, such as geometry, or characterization variables, such as physical properties — for example in inverse design problems. We thus get an implicit input-output relationship, evaluation of which demands solution of the underlying partial differential equations. See [158] for a detailed presentation of the state of the art of the design problem and some examples. The development of computational methods is permitting rapid and reliable evaluation of this partial-differential-equation-induced input-output relationship in the design, optimization and control contexts. See recent developments in [134]. The approach used is based on the reduced-basis method, first introduced in the late 1970s for non-linear structural analysis, and later developed more extensively in the 1980s and 1990s [11, 13, 37, 113, 117, 133]. The reduced basis method recognizes that the field variable is not, in fact, some arbitrary member of the infinite-dimensional solution space associated with the partial differential equation; rather, the field variable resides, or evolves, on a much lower-dimensional manifold induced by the parametric dependence.

In the application we use global approximation spaces; second, we make rigorous a posteriori error estimations; and third, we exploit off-line/on-line computational decompositions (see Balmes [11] for application of this strategy within the reduced-basis context). These three steps allow us — for the restricted but important class of “parameter-affine” problems, that it has been extended in Chapters 6 and 7 — to reliably decouple the generation and projection stages of reduced-basis approximation, thereby effecting computational economies of several orders of magnitude.

In Section A.2 we present the problem statement. In Section A.3 we describe the a posteriori error estimation framework. In Section A.4 we present the a priori convergence theory applied also to our output bounds and not only to approximate solution. In Section A.5 we study our procedure to control $N$ more tightly and to apply our error bound adaptively in the choice of $\mu$ parameters family. Finally, in Section A.6, we present the application and in the Section A.7 the numerical results for our “model-problem” example, in Section A.8, we provide our guidelines for the development of reduced basis method in fluid mechanics and precisely in haemodynamics. These aspects are the subject of Chapters 5, 6, 7 and 8. Finally, in Section A.9 we provide an historical review on model-order reduction and some remarks on its development.

### A.2 Problem formulation

We first introduce a Hilbert space $Y$, and an associated inner product and a norm, $(\cdot, \cdot)$ and $\| \cdot \| = (\cdot, \cdot)^{1/2}$, respectively. We next introduce the dual space of $Y$, $Y'$, and the associated duality pairing between $Y$ and $Y'$, $\langle \cdot, \cdot \rangle_Y = (\cdot, \cdot)$.

We then define, for any $\mu \in D^\mu \subset \mathbb{R}^P$, the parametrized (distributional) operator $A(\mu): Y \rightarrow$
Y'. We assume that \( A(\mu) = A(\Theta(\mu)) \), where, for any \( \Theta \in \mathbb{R}_+^Q \), \( A(\Theta): Y \rightarrow Y' \) is given by

\[
A(\Theta) = A_0 + \sum_{q=1}^{Q} \Theta_q A_q ,
\]

and the \( \Theta_q: \mathcal{D}^u \rightarrow \mathbb{R}_+ \), \( q = 1, \ldots, Q \), are non-negative functions. Here \( \mathbb{R}_+ \) refers to the non-negative real numbers. The range of \( \Theta \) is denoted \( \mathcal{D}^\theta \); and we define \( \theta_{\text{min}}^{\text{max}} (\geq 0) \), \( \theta_{\text{max}} \) (assumed finite), and \( \mathcal{D}_\text{box}^\theta \subseteq \mathbb{R}_+^Q \) as

\[
\theta_q^{\text{min}} = \sup_{\{t \in \mathbb{R}_+ | \Theta_q(\mu) \geq t, \forall \mu \in \mathcal{D}^u\}}, \quad q = 1, \ldots, Q ,
\]

\[
\theta_q^{\text{max}} = \inf_{\{t \in \mathbb{R}_+ | \Theta_q(\mu) \leq t, \forall \mu \in \mathcal{D}^u\}}, \quad q = 1, \ldots, Q ,
\]

and \( \mathcal{D}_\text{box}^\theta = \prod_{q=1}^{Q} [\theta_q^{\text{min}}, \theta_q^{\text{max}}] \), respectively.

Finally, we require that \( A_0 \) is continuous, symmetric, and coercive, and that the \( A_q, q = 1, \ldots, Q \), are continuous, symmetric, and positive-semidefinite \( (A_q v, v) \geq 0, \forall v \in Y) \); it follows that \( A(\Theta) \) (respectively, \( A(\mu) \)) is continuous, symmetric, and coercive for all \( \theta \) in \( \mathcal{D}_\text{box}^\theta \) (respectively, for all \( \mu \) in \( \mathcal{D}^u \)).

The general formulation for the problem can then be stated as: given \( \mu \in \mathcal{D}^u \), and linear functional \( F \in Y' \), evaluate the output

\[
s(\mu) = \langle F, u(\mu) \rangle ,
\]

where \( u(\mu) \in Y \) is the unique solution of \( A(\Theta(\mu)) u(\mu) = F \); we shall interpret the latter as

\[
\langle A(\Theta(\mu)) u(\mu), v \rangle = \langle F, v \rangle , \quad \forall v \in Y. \tag{A.2.1}
\]

Output \( s(\mu) \) may also be interpreted as the energy of the solution — \( s(\mu) = \langle F, u(\mu) \rangle = \langle A(\Theta(\mu)) u(\mu), u(\mu) \rangle \) — and is hence strictly positive. The output \( s(\mu) \) is “compliant,” and the operator \( A(\Theta) \) is symmetric; however, the formulation will be readily extended [118] to treat both non-compliant outputs, \( s(\mu) = \langle L, u(\mu) \rangle \) for given \( L \in Y' \), and non-symmetric, but still coercive, operators. We may also express our output as

\[
s(\mu) = \langle F, A^{-1}(\Theta(\mu)) F \rangle . \tag{A.2.2}
\]

Here, for any \( \Theta \in \mathcal{D}_\text{box}^\theta \), \( A^{-1}(\Theta): Y' \rightarrow Y \) is the (continuous, symmetric, coercive) inverse of \( A(\Theta) \); further, \( \forall G \in Y', \langle A(\Theta) A^{-1}(\Theta) G, v \rangle = \langle G, v \rangle , \forall v \in Y \).

### A.2.1 Galerkin approximation

The \( u(\mu) \) of (A.2.1) are, in general, not known exactly. In order to construct our reduced-basis space we will therefore require a finite-dimensional “truth” approximation to \( Y \), which we shall denote \( \hat{Y}_h \); \( \hat{Y}_h \) is an \( N \)-dimensional subspace of \( Y \). For example, for \( \Omega \subseteq \mathbb{R}^{d=1, 2, \text{or} 3} \), and \( \hat{Y}_h \subset H^1(\Omega) \), \( \hat{Y}_h \) will typically be a finite element approximation space associated with a very fine triangulation \( \mathcal{T}_h \) of the computational domain. We assume that the triangles, denoted with \( T_h^j \), also referred to as elements, cover the computational domain \( \Omega \), \( \Omega = \bigcup_{T_h \in \mathcal{T}_h} T_h \) (\( \mathcal{T}_h \)}}
is the closure $T_h$) and that each of the elements do not overlap, $T_h^i \cap T_h^j = 0$, $\forall T_h^i, T_h^j \in T_h$.

The subscript $h$ denotes the diameter of the triangulation defined as:

$$h = \sup_{T_h \in T_h} \sup_{x,y \in T_h} |x - y|;$$

(A.2.3)

here $| \cdot |$ is the Euclidean norm.

**Discrete Problem**

Using then the triangulation $T_h$, we define the space $\tilde{Y}_h$ as the space of continuous functions which are piecewise linear over each of the elements $T_h^i \subset T_h^j$:

$$\tilde{Y}_h = \{ v \in C^0(\Omega) | v|_{T_h} \in \mathbb{P}^1(T_h) \text{, } \forall T_h \in T_h \}. $$

(A.2.4)

If $\mathcal{N}$ is the number of nodes in the triangulation, we introduce the Lagrangian basis functions $\phi_i \in \tilde{Y}_h$, such that $\phi_i(x_j) = \delta_{ij}$, $i,j = 1,\ldots,\mathcal{N}$, where $x_j$ are the coordinates of node $j$, and $\delta_{ii} = 1$ if $i = j$, or $\delta_{ij} = 0$ if $i \neq j$. Each function $\phi_i$ has compact support over the region defined by the elements surrounding node $i$. In general, we expect that $\mathcal{N}$ will be very large.

See also [127] for more detailed studies on FEM approximation.

Our Galerkin-FEM approximation can be stated as: given a $\mu \in \mathcal{D}^\mu$, evaluate the output

$$\tilde{s}(\mu) = \langle F, \tilde{u}(\mu) \rangle, $$

(A.2.5)

where $\tilde{u}(\mu) \in \tilde{Y}_h$ is the unique solution of

$$\langle A(\Theta(\mu)) \tilde{u}(\mu), v \rangle = \langle F, v \rangle, \text{ } \forall v \in \tilde{Y}_h.$$ (A.2.6)

As before, the output can be expressed as a (strictly positive) energy: $\tilde{s}(\mu) = \langle F, \tilde{u}(\mu) \rangle = \langle A(\Theta(\mu)) \tilde{u}(\mu), \tilde{u}(\mu) \rangle$.

It is convenient to express (A.2.5)–(A.2.6) in terms of the basis for $\tilde{Y}_h = \{ \phi_i, i = 1,\ldots,\mathcal{N} \}$.

We first introduce the matrices $\tilde{A}_q \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}, q = 0,\ldots,Q$, as $\tilde{A}_{q,i,j} = \langle A\phi_j, \phi_i \rangle , 1 \leq i,j \leq \mathcal{N}$; it is readily shown that $\tilde{A}_0$ (respectively, $\tilde{A}_q, q = 1,\ldots,Q$) is symmetric positive-definite (respectively, symmetric positive-semidefinite). For any $\Theta \in \mathcal{D}_\text{box}^\theta$, we then define $\tilde{A}(\Theta) \in \mathbb{R}^{\mathcal{N} \times \mathcal{N}}$ as

$$\tilde{A}(\Theta) = \tilde{A}_0 + \sum_{q=1}^Q \Theta_q \tilde{A}_q;$$

where $\tilde{A}(\Theta)$ is symmetric positive-definite for all $\Theta \in \mathcal{D}_\text{box}^\theta$. In the same way we introduce $\tilde{F} \in \mathbb{R}^\mathcal{N}$ as $\tilde{F}_i = \langle F, \phi_i \rangle , 1 \leq i \leq \mathcal{N}$.

Our approximation can then be restated as: given $\mu \in \mathcal{D}^\mu$, evaluate the output

$$\tilde{s}(\mu) = \tilde{F}^T \tilde{u}(\mu),$$

where $^T$ refers to the algebraic transpose and $\tilde{u}(\mu) \in \mathbb{R}^\mathcal{N}$ is the unique solution of

$$\tilde{A}(\Theta(\mu)) \tilde{u}(\mu) = \tilde{F}.$$ (A.2.7)
Note that $\tilde{u}(\mu)$ and $\tilde{\mathbf{u}}(\mu) = (\tilde{u}_1, \ldots, \tilde{u}_N)^T$ are related via

$$\tilde{u}(\mu) = \sum_{j=1}^{N} \tilde{u}_j(\mu) \phi_j.$$ 

As always, our compliance output can be expressed as an energy:

$$\tilde{s}(\mu) = \tilde{\mathbf{u}}^T(\mu) \tilde{A}(\Theta(\mu)) \tilde{\mathbf{u}}(\mu),$$

or, equivalently,

$$\tilde{s}(\mu) = \tilde{F}^T \tilde{A}^{-1}(\Theta(\mu)) \tilde{F}, \tag{A.2.8}$$

where $\tilde{A}^{-1}(\Theta)$ is the (symmetric, positive-definite) inverse of $\tilde{A}(\Theta)$. Note that since $N$ can be large, solution of (A.2.7), and hence evaluation of $\tilde{s}(\mu)$, will be computationally expensive.

### Computational Complexity

We see that the original problem has been replaced by a finite-dimensional one. Assuming sufficient regularity of the solution $u(\mu)$ the a priori convergence theory for this type of finite-elements suggests that the error in the output $|s(\mu) - \tilde{s}(\mu)|$ will converge as $h^2$ (see Theorem 4.1 in Patera and Renquist [110]), where $h$ is defined in (A.2.3). Moreover as $h \to 0$, we get $\tilde{u}(\mu) \to u(\mu)$ and $\tilde{s}(\mu) \to s(\mu)$. The above a priori result suggests also that to decrease the error in the output by a factor $C > 0$, we need to increase the number of elements and therefore $N$ roughly by the same factor. We see that as the requirements for accuracy increase or the geometric complexity increases, we need higher $N$ to obtain accurate and reliable results (to ascertain the accuracy we need a posteriori error estimators). Moreover, in the presence of singularities or boundary layers, local refinement is essential, further increasing the required degrees of freedom.

The discussion above suggests that even for relatively simple problems in simple domains $N$ is large, at least $O(10^3)$, but it is not uncommon for $N$ to be $O(10^6)$ or higher. We also see the difficulty, as $N$ increases, so does the size of the linear system (A.2.7), which has to be inverted. By virtue of the compact support of $\phi_i$, the matrix $\tilde{A}$ is sparse and therefore iterative solvers can be used to obtain a solution. The computational complexity scales as $O(N^a)$, where $a$ depends on the condition number of the problem (which increases quadratically with $1/h$). Especially in contexts where repeated solution of (A.2.7) is required, the computational requirements soon become unacceptably large.

### A.2.2 Reduced basis method

Identifying the problem in the high dimensionality of the finite-element spaces, we look for ways to further reduce the computational complexity. The large number of degrees of freedom required in the case of finite-element methods, is attributed to the particular choice of basis functions, which have general approximation properties for functions in $Y$. To further reduce the computational complexity we look for spaces with approximation properties specific to the problem of interest.
The critical observation is that the solution and the output evolve in a low-dimensional manifold induced by the parametric dependence of the problem. Central to reduced-basis methods, first introduced in Nagy [102], is constructing an approximation to this manifold. In our approach, slightly different from earlier approaches, we construct linear reduced-basis spaces comprising of solutions to \((A.2.7)\) at different parameter points. We then use these spaces to find an approximation \(u_N(\mu)\) to the exact solution.

Earlier approaches viewed the reduced-basis method as a combined projection and continuation method. A different view, suitable for our purposes, is that of multi-dimensional parameter-space interpolation. The required interpolation weights are obtained by solving suitably defined low-dimensional problems chosen to minimize the approximation error measured in problem-specific energy norms.

**Reduced-basis space**

We introduce some “\(\mu\)” samples \(S^\mu_N = \{\mu^1, \ldots, \mu^N\}\), where \(\mu^n \in \mathcal{D}^\mu\), \(n = 1, \ldots, N\). We then define our reduced basis space \(W_N = \text{span}\{\tilde{\zeta}_n, n = 1, \ldots, N\}\), where \(\tilde{\zeta}_n = \tilde{u}(\mu^n), n = 1, \ldots, N\). Recall that \(\tilde{u}(\mu^n)\) is the solution of \((A.2.6)\) for \(\mu = \mu^n\). We denote \(\tilde{\zeta}_n = \tilde{u}(\mu^n), n = 1, \ldots, N\).

In the first step, given a \(\mu \in \mathcal{D}^\mu\), we find \(s_N(\mu) = \langle F, u_N(\mu) \rangle\), where \(u_N(\mu) \in W_N\) satisfies

\[
\langle A(\Theta(\mu)) u_N(\mu), v \rangle = \langle F, v \rangle, \quad \forall \ v \in W_N.
\]

Also in this case we can express the output as an energy, \(s_N(\mu) = \langle A(\Theta(\mu)) u_N(\mu), u_N(\mu) \rangle\).

In terms of our basis functions, we define the symmetric positive-definite matrix \(\mathbf{A}_N(\mu) \in \mathbb{R}^{N \times N}\) as \(A_{N,1}(\mu) = \langle A(\Theta(\mu)) \tilde{\zeta}_j, \tilde{\zeta}_i \rangle\), \(1 \leq i, j \leq N\), and the vector \(\mathbf{F}_N \in \mathbb{R}^N\) as \(F_N = \langle F, \tilde{\zeta}_i \rangle\), \(1 \leq i \leq N\). It is a simple matter to observe that

\[
\mathbf{A}_N(\Theta) = \mathbf{A}_N \Theta + \sum_{q=1}^{Q} \Theta_q \mathbf{A}_N q, \quad (A.2.9)
\]

where \((\mathbf{A}_N q)_{i,j} = \langle A_q \tilde{\zeta}_j, \tilde{\zeta}_i \rangle, 1 \leq i, j \leq N, 0 \leq q \leq Q\); note that the \(\mathbf{A}_N q \in \mathbb{R}^{N \times N}\), \(0 \leq q \leq Q\), are independent of \(\theta\).

We can then restate the formulation as: given \(\mu \in \mathcal{D}^\mu\), find \(s_N(\mu) = \mathbf{F}_N \mathbf{u}_N(\mu)\), where \(\mathbf{u}_N(\mu) \in \mathbb{R}^N\) is the unique solution to

\[
\mathbf{A}_N(\Theta(\mu)) \mathbf{u}_N(\mu) = \mathbf{F}_N.
\]

Note that \(u_N(\mu) = \sum_{j=1}^{N} u_N j(\mu) \tilde{\zeta}_j\). The output may also be expressed as

\[
s_N(\mu) = \mathbf{u}_N^T(\mu) \mathbf{A}_N(\Theta(\mu)) \mathbf{u}_N(\mu) = \mathbf{F}_N^T \mathbf{A}_N^{-1}(\Theta(\mu)) \mathbf{F}_N. \quad (A.2.10)
\]

**Remark A.2.1** The a priori convergence theory, and extensive numerical tests, suggest that the convergence of the reduced-basis approximation to the exact will be very fast. In fact, exponential convergence is observed in all the numerical tests. This suggests that even with a very modest \(N\), we can expect to achieve good accuracy. The linear system above can be formed and solved very efficiently in the case where the operator depends affinely on the parameters. In this case we can separate the computational steps into two stages:
A.3. REDUCED BASIS OUTPUT BOUNDS

- The off-line stage, in which the reduced-basis space is constructed and some pre-processing is performed. This is an expensive step, that needs to be performed only once, requiring solutions of finite-element problems.

- The on-line stage, in which for each new parameter value, the reduced-basis approximation for the output of interest is calculated.

The on-line stage is “blackbox” in the sense that there is no longer any reference to the original problem formulation: the computational complexity of this stage scales only with the dimension of the reduced-basis space and the parametric complexity of the partial differential operator. The “blackbox” nature of the on-line component of the procedure has other advantages. In particular, the on-line code is simple, non-proprietary, and completely decoupled from the (often complicated) off-line “truth” code. This is particularly important in multidisciplinary design optimization, in which various models and approximations must be integrated.

A.3 Reduced basis output bounds

The computational relaxation introduced in the previous section, allows us to compute very efficiently accurate approximations to the solution and the output of interest. Thanks to the expected rapid convergence $N$ could, in theory, be chosen quite small. However, in practice we do not know how small $N$ should be: this will depend on the desired accuracy, the choice of $\mu$ in the construction of the reduced-basis spaces, the output of interest and the particular problem in question. Either too many or too few basis functions will be retained: the former results in computational inefficiency; the later in unacceptable uncertainty. For the successful application of reduced-basis methods it is therefore critical that we can ascertain the accuracy of our predictions; for this reason rigorous error-estimation approaches are needed, directly for outputs of interest, to validate a posteriori the accuracy of our predictions.

We prove that these estimators $s_N^+(\mu)$ and $s_N^-\mu$ are upper and lower bounds, respectively, to the “true” output $\tilde{s}(\mu)$ that would be obtained by solution of the expensive finite-element problem:

$$s_N^-\mu \leq \tilde{s}(\mu) \leq s_N^+\mu.$$  \hspace{1cm} (A.3.1)

Unlike the exact value, these error estimators can be computed inexpensively — with a complexity that scales only with the dimension of the reduced-basis space.

In reality the error in the output has two components:

$$|s(\mu) - s_N(\mu)| \leq |s(\mu) - \tilde{s}(\mu)| + |\tilde{s}(\mu) - s_N(\mu)|;$$

the first related to the discretization error (see in Section A.9.2); and the second to the reduced-basis error. In practice, both of these errors have to be estimated for reliability in our predictions. Estimation of the discretization error has been treated extensively in the literature; see Patera and Peraire [109] for a review. For our purposes, we assume that $h$ is chosen very conservatively such that $\tilde{s}(\mu) \approx s(\mu)$ and the dominant error is due to the reduced-basis approximation.
A.3.1 Bound conditioner

In the previous section we have seen how to compute our predictor, \( s_N(\mu) \) from equation (A.2.10); in the next steps we are going to compute our bounds, \( s_N(\mu) \leq \tilde{s}(\mu) \leq \tilde{s}_N(\mu) \). The latter may be interpreted also as a posteriori estimators. We first define the error \( \tilde{e}(\mu) = u(\mu) - e_N(\mu) \) and residual \( R \in Y' \) as \( \langle R(\mu), v \rangle \equiv \langle F - A(\Theta(\mu)) u_N(\mu), v \rangle, \forall v \in Y \); and then \( \tilde{R}(\mu) \in \mathbb{R}^N \) as \( \tilde{R}_i(\mu) = \langle R(\mu), \phi_i \rangle, i = 1, \ldots, N \). We note that

\[
\tilde{R}(\mu) = F - A(\Theta(\mu)) u_N(\mu),
\]

and that \( u_N(\mu) \) is given by

\[
u_N(\mu) = \sum_{n=1}^{N} u_{Nn}(\mu) \tilde{\zeta}_n;
\]

by construction: \( \tilde{\zeta}_n = \sum_{i=1}^{N} \tilde{u}_i(\mu^n) \phi_i \). We then introduce a symmetric, continuous, and coercive bound conditioner as in Veroy [158] \( C(\mu) : Y \to Y' \) such that the minimum and maximum eigenvalues

\[
\rho_{\min}(\mu) \equiv \min_{v \in Y} \frac{\langle \tilde{A}(\mu)v, v \rangle}{\langle C(\mu)v, v \rangle},
\]

\[
\rho_{\max}(\mu) \equiv \max_{v \in Y} \frac{\langle \tilde{A}(\mu)v, v \rangle}{\langle C(\mu)v, v \rangle},
\]

satisfy the following spectral condition:

\[
1 \leq \rho_{\min}(\mu), \quad \rho_{\max} \leq \rho,
\]

for some (preferably small) constant \( \rho \in \mathbb{R} \). We note that:

\[
\rho_{\min}(\mu) \leq \frac{\langle \tilde{A}(\mu)v, v \rangle}{\langle C(\mu)v, v \rangle} \leq \rho_{\max}(\mu) \quad \forall v \in Y,
\]

and for any \( \tilde{e} \in Y \) and \( \tilde{\epsilon} \in Y \) such that:

\[
\langle \tilde{A}(\mu)\tilde{e}, v \rangle = \langle \tilde{R}(\mu), v \rangle, \forall v \in Y,
\]

and

\[
\langle C(\mu)\tilde{e}, v \rangle = \langle \tilde{R}(\mu), v \rangle, \forall v \in Y,
\]

we can show that:

\[
\rho_{\min}(\mu) \leq \frac{\langle \tilde{A}(\mu)\tilde{e}, \tilde{v} \rangle}{\langle C(\mu)\tilde{e}, \tilde{v} \rangle} \leq \rho_{\max}(\mu).
\]

In addition to the spectral condition (A.3.6), we also require a “computational invertibility” hypothesis, in particular that \( C^{-1}(\mu) \) be of the form:

\[
C^{-1}(\mu) = \sum_{i \in I(\mu)} \alpha_i(\mu) C_i^{-1}
\]

where (i) \( I(\mu) \subset \{1, \ldots, I\} \) is a parameter-dependent set of indices, \( I \) is a finite (preferably small) integer, and each \( \mu \in \mathcal{D}^n \). (ii) The \( C_i : Y \to Y' \) are parameter-indipendent symmetric, coercive operators.
A.3.2 Error and output bounds

We now find $\hat{e}(\mu) \in \mathbb{R}^N$ such that

$$C(\mu) \hat{e}(\mu) = \hat{R}(\mu); \quad (A.3.11)$$

this equation will of course have a unique solution since $C(\mu)$ is symmetric positive-definite.

We can now define our lower and upper bounds as

$$s_N(\mu) = s_N(\mu),$$

and

$$s_N(\mu) = s_N(\mu) + \Delta_N(\mu),$$

where $\Delta_N(\mu)$, the bound gap, is given by

$$\Delta_N(\mu) = \hat{e}^T(\mu) C(\mu) \hat{e}(\mu)$$

$$= \hat{R}^T(\mu) C^{-1}(\mu) \hat{R}(\mu)$$

$$= \hat{R}^T(\mu) \hat{e}(\mu).$$

The previous expressions for the bound gap will be useful in the theoretical and computational context.

A.3.3 Minimum coefficient bound conditioner

We now consider a method for choosing the $C_i$ and the associated $\alpha_i$ called Minimum Coefficient Bound Conditioner.

To begin we recall our separability assumption on $\tilde{A}(\cdot)$:

$$\tilde{A}(\Theta(\mu)) = \tilde{A}_0 + \sum_{q=1}^Q \Theta_q(\mu) \tilde{A}_q \quad \forall \mu \in D^\mu_{\text{box}},$$

where $\Theta(\mu) : D^\mu \to \mathbb{R}$ and the $\tilde{A}_q : Y \to Y'$.

We now define

$$\hat{A}(\theta) = \hat{A}_0 + \sum_{q=1}^Q \theta_q \tilde{A}_q,$$

where $\theta \in \mathbb{R}^Q$ and $\hat{A}(\theta) : Y \to Y'$ and $\hat{A}_q \equiv \tilde{A}_q$. If $\Theta_q(\mu) = \theta_q$, we may then write

$$\hat{A}(\Theta(\mu)) = \hat{A}(\theta),$$

where $\Theta : D^\mu \to D^\theta$ and $D^\theta \equiv \text{Range}(\Theta) \in \mathbb{R}^Q$. Then we introduce $I$ points $\theta_i$ as samples $S^Q_I = \{\theta_1, \ldots, \theta_I\}$, where $\theta_i \in D^\theta_{\text{box}}, i = 1, \ldots, I$. We choose

$$\alpha_i(\mu) = \left( \min_{1 \leq q \leq Q} \left( \frac{\Theta_q(\mu)}{\theta^q_i} \right) \right)^{-1}, \quad (A.3.12)$$
and
\[ \mathcal{C}_i = \sum_{q=1}^{Q} \theta_i^q \tilde{A}_q. \]  
(A.3.13)

Our bound conditioner \( C(\mu) \in \mathbb{R}^{N \times N} \) is
\[ C(\mu) = \left( \sum_{i \in I(\mu)} \alpha_i(\mu) C_i^{-1}(\theta^i) \right)^{-1}. \]  
(A.3.14)

Clearly, \( C^{-1} \) and hence \( C \) are symmetric positive-definite. In words, \( C^{-1} \) is an approximation to \( \tilde{A}^{-1}(\Theta(\mu)) \). In our application we used a single-point conditioner. We set \( I = 1, S_1^\theta = \{ \theta \}, |I(\mu)| = 1, I(\mu) = \{ 1 \} \). For further information and a detailed understanding of bound conditioner families (and special cases) we suggest to see [90, 118, 120].

A.3.4 Bounding properties

It remains to demonstrate our claim that \( s^+_{N}(\mu) \leq \tilde{s}(\mu) \leq s^-_{N}(\mu) \) for all \( N \geq 1 \). We prove at first that for all \( \mu \in \mathcal{D}^\mu \), and all \( N \geq 1 \), \( s^-_{N}(\mu) \leq \tilde{s}(\mu) \). We have that
\[ \tilde{s}(\mu) - s_N(\mu) = \langle F, \tilde{u}(\mu) - u_N(\mu) \rangle \]
\[ = \langle A(\Theta(\mu)) \tilde{u}(\mu), \tilde{u}(\mu) - u_N(\mu) \rangle \]
\[ = \langle A(\Theta(\mu)) (\tilde{u}(\mu) - u_N(\mu)), \tilde{u}(\mu) \rangle \]
\[ = \langle A(\Theta(\mu)) (\tilde{u}(\mu) - u_N(\mu)), \tilde{u}(\mu) - u_N(\mu) \rangle \]
\[ \geq 0 \]
(A.3.15)

having used the definition of \( s(\mu) \), (A.2.6), the symmetry of \( A \), Galerkin orthogonality, and coercivity, respectively.

This lower bound proof is a standard result in variational approximation theory. We now turn to the upper bound to demonstrate that for all \( \mu \in \mathcal{D}^\mu \), and all \( N \geq 1 \), \( s^+_{N}(\mu) \geq \tilde{s}(\mu) \).

We first define \( \tilde{e} \in \mathbb{R}^N \) as \( \tilde{e} = \tilde{u} - u_N \); it then follows from (A.2.7) and (A.3.2) that
\[ \tilde{A}(\Theta(\mu)) \tilde{e}(\mu) = \tilde{R}(\mu), \]  
(A.3.16)

which is the usual error-residual relationship. It then follows from (A.3.15) that
\[ \tilde{s}(\mu) - s_N(\mu) = \tilde{e}^T(\mu) \tilde{A}(\Theta(\mu)) \tilde{e}(\mu) \]
\[ = \tilde{R}^T(\mu) \tilde{A}^{-1}(\Theta(\mu)) \tilde{R}(\mu). \]
(A.3.17)

It thus only remains to write the effectivity:
\[ \eta_{N}(\mu) \equiv \frac{s^+_{N}(\mu) - s_N(\mu)}{\tilde{s}(\mu) - s_N(\mu)} = \frac{\Delta_N(\mu)}{\tilde{s}(\mu) - s_N(\mu)} = \frac{\tilde{R}^T(\mu) \tilde{A}^{-1}(\Theta(\mu)) \tilde{R}(\mu)}{\tilde{R}^T(\mu) \tilde{A}^{-1}(\Theta(\mu)) \tilde{R}(\mu)}, \]  
(A.3.18)
using the boundaries properties (A.3.7) and (A.3.10) and equations (A.3.11)-(A.3.16) we have:
\[
\eta_N(\mu) = \frac{\tilde{e}^T(\mu) C(\Theta(\mu)) \tilde{e}(\mu)}{\tilde{e}^T(\mu) A^{-1}(\Theta(\mu)) \tilde{e}(\mu)} = \frac{\tilde{C}(\mu) \tilde{e}}{\langle A(\Theta(\mu)) \tilde{e}, \tilde{e} \rangle},
\]
therefore:
\[
\rho_{\text{min}}(\mu) \leq \eta_N(\mu) \leq \rho_{\text{max}}(\mu);
\] (A.3.20)
by construction, \(\rho_{\text{min}}(\mu) \geq 1\) for all \(\mu \in D^\mu\) and therefore:
\[
\eta_N(\mu) \geq 1,
\] (A.3.21)
and \(s_N^+(\mu) \geq s(\mu)\) as required. Note that the result (A.3.20) also indicates the sharpness of our bounds: it follows from (A.3.6) that:
\[
\eta_N(\mu) \leq \rho.
\] (A.3.22)
This result provides insight as to the properties of a good bound conditioner. Clearly, we wish \(\rho_{\text{max}}(\mu)\) to be as close to unity, and hence as close to \(\rho_{\text{min}}\), as possible.

### A.3.5 Computational procedure: off-line/on-line decomposition

We review here arguments given in detail in [118]; early applications of this approach may be found in [11]. The theoretical and empirical results let us to apply separate off-line and on-line computational procedures that exploit the dimension reduction.

In an off-line stage, we find the \(\tilde{z}_n\), \(n = 1, \ldots, N\) \((N \times \tilde{A}\)-solves), and form the \(A_{N,q}\), \(0 \leq q \leq Q\) \(((Q + 1)N^2 \times \tilde{A}\)-inner products), and \(\tilde{F}_N\) \((N \times \tilde{N}\) operations). In the on-line stage — given any new \(\mu\) — we need only form \(A_N(\mu)\) from the \(A_{N,q}\) \(((Q + 1)N^2\) operations), find \(u_N(\mu)\) \((O(N^3)\) operations), and evaluate \(s_N(\mu)\) \((N\) operations). The most important point is that the on-line complexity (and storage — \(O(QN^2)\)) is independent of the very large dimension of the truth space \(\tilde{Y}_h, \tilde{N}\); in particular, since \(N\) is typically very small (as suggested in the previous sections), “real-time” response is obtained.

**Computational Procedures for the upper bound \(s_N^+(\mu)\)**

A computational procedure for the upper bound is very important and very useful in view of adaptivity procedures applied to basis \(\tilde{z}_n\) storage and to test results with a cheap and fast procedure. See [118] for detailed bound conditioners presentation. We first note from (A.3.14)–(A.3.11) that
\[
\tilde{e}(\mu) = \sum_{i \in I(\mu)} \alpha_i(\mu) \tilde{A}_i^{-1}(\theta) \left[ \tilde{F} - \sum_{q=0}^Q \sum_{n=1}^N \Theta_q(\mu) u_{N,n}(\mu) \tilde{A}_n \right] \tilde{z}_n;
\]
recall that \(\Theta_0 = 1\). It follows that we may express \(\tilde{e}(\mu)\) as
\[
\tilde{e}(\mu) = \sum_{i \in I(\mu)} \alpha_i(\mu) \left[ \tilde{z}_i^0 + \sum_{q=0}^Q \sum_{n=1}^N \Theta_q(\mu) u_{N,n}(\mu) \tilde{z}_n^i \right],
\]
where for all \( i \in \{1, \ldots, M\} \), \( \tilde{A}(\theta^i) \tilde{z}^i_{mn} = \tilde{F}_n \) and \( \tilde{A}(\theta^i) \tilde{z}^i_{qm} = -\tilde{A}_q \tilde{\zeta}_n \), \( 0 \leq q \leq Q, 1 \leq n \leq N \). We may thus express our bound gap \( \Delta_N(\mu) \) as the following (A.3.23):

\[
\Delta_N(\mu) = \tilde{F}^T(\mu) \tilde{\xi}(\mu)
\]

\[
= \sum_{i \in I(\mu)} \alpha_i(\mu) \left[ \tilde{F} - \sum_{q=0}^{Q} \sum_{n=1}^{N} \Theta_q(\mu) u_{Nn}(\mu) \tilde{A}_q \tilde{\zeta}_n \right]^T \sum_{q=0}^{Q} \sum_{n=1}^{N} \Theta_q(\mu) u_{Nn}(\mu) \tilde{z}^i_{qn} + \sum_{q=0}^{Q} \sum_{n=1}^{N} \sum_{q'=0}^{Q} \sum_{n'=1}^{N} \Theta_q(\mu) \Theta_q'(\mu) u_{Nn}(\mu) u_{Nn'}(\mu) \Gamma^i_{qq'nn'} \right],
\]

where for all \( i \in \{1, \ldots, M\} \), \( c^i = \tilde{F}^T \tilde{z}^i_{mn} \), \( \Lambda^i_{qn} = \tilde{F}^T \tilde{z}^i_{qm} - \tilde{F}^T \tilde{A}_q \tilde{z}^i_{00} \) for \( 0 \leq q \leq Q, 1 \leq n \leq N \), and \( \Gamma^i_{qq'nn'} = -\tilde{\zeta}_n \tilde{A}_q \tilde{z}^i_{0n'} \) for \( 0 \leq q, q' \leq Q, 1 \leq n, n' \leq N \).

The off-line/on-line decomposition is now clear. In the off-line stage we compute the \( \tilde{z}^i_{00} \) and \( \tilde{z}^i_{mn}(M((Q+1)N+1) \tilde{A}\text{-solves})\) and the \( c^i, \Lambda^i_{qn}, \) and \( \Gamma^i_{qq'nn'} \) (predominated by \( M((Q+1)^2N^2 + (Q+1)N) \tilde{A}\text{-inner products} \)). In the on-line stage we need “only” perform the sum (A.3.23), which requires \( |I(\mu)|((Q+1)^2N^2 + (Q+1)N + 1) \) operations. The essential point is that the on-line complexity (and storage — \( O(M(Q+1)^2N^2) \)) is independent of \( N \).

We note that the off-line/on-line decomposition depends critically on the “separability” of \( \tilde{C}^{-1} \) as a sum of products of parameter-dependent functions (the \( \alpha_i(\mu) \)) and parameter-independent operators (the \( \tilde{A}^{-1}(\theta^i) \)). In turn, it is the direct approximation of \( \tilde{A}^{-1}(\theta^i) \) (i.e., by a convex combination of \( \tilde{A}^{-1}(\theta^i) \)) rather than of \( \tilde{A}(\theta(\mu)) \) (e.g., by a convex combination of \( \tilde{A}(\theta^i) \)) that permits us to achieve this separability while simultaneously pursuing a “high-order” bound conditioner achieving some fixed (known, certain) accuracy — as measured by \( \Delta_N(\mu) \) — at a lower computational effort. See [119].

### A.4 A priori convergence theory

We recall a priori framework for a general a priori convergence theory. Depending on the context and application, we will either invoke the lower bound (\( s_N^- (\mu) \)) or upper bound (\( s_N^+ (\mu) \)) as our estimator for \( \tilde{s}(\mu) \). For example, in an optimization problem in which \( \tilde{s}(\mu) \) enters as a constraint \( \tilde{s}(\mu) \leq s^\text{max} \) (respectively, \( \tilde{s}(\mu) \geq s^\text{min} \)), we will replace this condition with \( s_N^+ (\mu) \leq s^\text{max} \) (respectively, \( s_N^- (\mu) \geq s^\text{min} \)) so as to ensure satisfaction/feasibility even in the presence of approximation errors. The rigorous bounding properties proven in Section A.3.4 provide the requisite certainty.
§A.4. A PRIORI CONVERGENCE THEORY

But we of course also require accuracy: if, in the optimization context cited above, \( s_N^+(\mu) \) or \( s_N^-(\mu) \) is not close to \( \tilde{s}(\mu) \), then our design may be seriously suboptimal. Since \( |s_N^+(\mu) - \tilde{s}(\mu)| \leq |s_N^+(\mu) - s_N^-(\mu)| = \Delta_N(\mu) \) and \( |\tilde{s}(\mu) - s_N^-(\mu)| \leq |s_N^+(\mu) - s_N^-(\mu)| = \Delta_N(\mu) \), it is the convergence of \( \Delta_N(\mu) \) as a function of \( N \) that we must understand. In particular, from (A.3.19) and (A.3.15) we may write
\[
\Delta_N(\mu) = s_N^+(\mu) - s_N^-(\mu) = (\tilde{s}(\mu) - s_N^-(\mu)) \left( \frac{s_N^+(\mu) - s_N^-(\mu)}{\tilde{s}(\mu) - s_N^-(\mu)} \right)
\]
where \( \tilde{e}(\mu) = \tilde{u}(\mu) - u_N(\mu) \). In some sense, the first factor, \( \langle A(\Theta(\mu)) \tilde{e}(\mu), \tilde{e}(\mu) \rangle \), measures the error in the solution \( \tilde{u}(\mu) - u_N(\mu) \), while the second factor, the effectivity \( \eta_N(\mu) \), measures the ratio of the actual and estimated errors; the former should be small, while the latter should be close to unity. As we shall see, this two-step factorization is important not only as a theoretical construct: it is this factorization which permits us to achieve high accuracy while simultaneously honoring our bound requirements. We would thus like to understand the convergence of \( \Delta_N(\mu) \) to zero as a function of \( N \). In particular, we consider the case in which \( A(\mu) = A_0 + \mu A_1 \) (and hence \( \Theta_1(\mu) = \mu \)), and \( \mu \in D^\mu \equiv [0, \mu_{\max}] \). From our continuity and coercivity assumptions, there exists a positive real constant \( \gamma_1 \) such that
\[
\langle A_1 v, v \rangle \leq \gamma_1 \langle A_0 v, v \rangle ; \quad \text{(A.4.1)}
\]
it thus follows that \( \langle A(\mu)v, v \rangle \leq (1 + \mu_{\max} \gamma_1) \langle A_0 v, v \rangle \). Defining \( \| \cdot \| = \langle A_0 \cdot, \cdot \rangle \), we may thus write
\[
\Delta_N(\mu) \leq (1 + \mu_{\max} \gamma_1) \| \tilde{u}(\mu) - u_N(\mu) \|^2 \eta_N(\mu). \quad \text{(A.4.2)}
\]

A.4.1 Best approximation

It remains to bound \( \| \tilde{u}(\mu) - u_N(\mu) \| \) and \( \eta_N(\mu) \); and, in particular, to understand the convergence rate of \( \| \tilde{u}(\mu) - u_N(\mu) \| \rightarrow 0 \) and \( \eta_N(\mu) \rightarrow 1 \) (or at least a constant) as \( N \) increases. The proofs for both \( \| \tilde{u}(\mu) - u_N(\mu) \| \) [91] and \( \eta_N(\mu) \) implicate a particular “optimal” logarithmic point distribution which we thus impose \textit{a priori}. In particular, we introduce an upper bound for \( \gamma_1, \gamma \), and a “log increment” \( \delta_N = (\ln(\gamma \mu_{\max} + 1))/(N - 1) \); we then define
\[
\mu^n = \exp\{-\ln \gamma + (n - 1)\delta_N\} - \gamma^{-1}, \quad 1 \leq n \leq N ,
\]
and take \( S_N^\mu = \{\mu^1, \ldots, \mu^N\} \}. \) Clearly, \( \ln(\mu^n + \gamma^{-1}) \) is uniformly distributed. Note that for \( N \geq N_{\text{crit}} \equiv 1 + e \ln(\gamma \mu_{\max} + 1) \), \( \delta_N \leq e^{-1} < 1 \). We remind to the main result in [91, 92]: for \( N \geq N_{\text{crit}} \equiv 1 + e \ln(\gamma \mu_{\max} + 1) \) and all \( \mu \in D^\mu \) we find:
\[
\| \tilde{u}(\mu) - u_N(\mu) \| \leq (1 + \mu_{\max} \gamma_1)^{1/2} \| \tilde{u}(0) \| e^{-\delta_N N_{\text{crit}}}.
\]
See Theorem 3 of [91] (for \( e^* = 1 \)). We see that we obtain \textit{exponential} convergence, \textit{uniformly} for all \( \mu \) in \( D^\mu \). Furthermore, our convergence threshold parameter \( N_{\text{crit}} = 1 + e \ln(\gamma \mu_{\max} + 1) \), and exponential convergence rate \( 1/N_{\text{crit}} \), depend only weakly — \textit{logarithmically} — on \( \gamma_1 \) and
\( \mu^{\text{max}} \) (which together comprise the continuity-coercivity ratio). In short, we expect extremely rapid convergence even for large parameter ranges. To obtain a bound for \( \eta_N(\mu) \) (limited and close to unity) see [157].

### A.5 Adaptive procedure for basis construction

Given the higher powers of \( N \) that appear in our complexity estimates, it is crucial (both as regards online and offline effort) to control \( N \) more tightly. To this end, we may gainfully apply our \textit{a posteriori} error bounds adaptively. We first construct, offline, an approximation that, over most of the domain, exhibits an error (in the \( H^1 \)-norm) less than \( \epsilon_d^{\text{prior}} \); we begin with a first point \( \mu^1(S_{N'=1} = \{\mu^1\}) \); we next (inexpensively) evaluate \( \Delta_{N'=1}(\mu) \) over a large test sample of parameter points in \( \mathcal{D}^\mu, \Sigma^{\text{prior}} \); we then choose for \( \mu^2 \) (and hence \( S_{N'-2} = \{\mu^1, \mu^2\} \)) the maximizer of \( \Delta_{N'=1}(\mu) \) over \( \Sigma^{\text{prior}} \). We repeat this process until the maximum of \( \Delta_{N'=N^{\text{prior}}}(\mu) \) over \( \Sigma^{\text{prior}} \) is less than \( \epsilon_d^{\text{prior}} \). Then, online, given a new value of the parameter, \( \mu \), and an error tolerance \( \epsilon_d^{\text{post}}(\mu) \), we essentially repeat this adaptive process - but now our sample points are drawn from \( S_{N^{\text{prior}}} \), and the test sample is a singleton - \( \mu \). Typically we choose \( \epsilon_d^{\text{prior}} \ll \epsilon_d^{\text{post}}(\mu) \) since our test is not exhaustive; and therefore, typically, \( N^{\text{post}}(\mu) \ll N^{\text{prior}} \). With the adaptive process we get higher accuracy at lower \( N \): modest reductions in \( N \) can translate into measurable performance improvements. This procedure is very important not only to get a computationally cheaper and faster method but also to avoid ill-conditioning in matrix assembling procedures.

### A.6 A multi-parameter application

We consider a physical domain \( \Omega \subset \mathbb{R}^2 \) is divided in four subdomains \( \Omega^r, r = 1, \ldots, 4 \), a \textit{“T”}-shaped region with boundary \( \Gamma \), divided in \( \Gamma_D \) and \( \Gamma_N \) and associated, respectively, Dirichlet or Neumann condition. See Figures A.1 and A.2.

#### A.6.1 Geometrical model

The geometrical model is based on a simple and standard configuration made up of four square subdomains parametrized in their dimensions (lengths and angles). This is a first example towards a more relevant one concerning bypass anastomosis used in Chapters 6 and 7. We have chosen five geometrical parameters.

- The angle \( \theta \) for the incoming branch of the bypass.
- The diameters \( t \) and \( D \), respectively, of the bypass and the artery.
- The lengths \( L \) and \( S \), respectively, the outflow length and the distance between the incoming new branch and the occlusion caused by a stenosis. See Figure A.1.

We underline that this is a macro geometrical structure, useful to study with low computational cost and sharp error bounds a possible configuration to be optimized by optimal shape
A.6. A MULTI-PARAMETER APPLICATION

Preliminary tests were made studying a simpler case based only on a square domain.

A.6.2 Formulation

We present the strong form of the equations governing our preliminary problem (in $\hat{\Omega}$), from which we derive the weak statement; we then reformulate the problem in terms of a reference (parameter-independent) domain ($\Omega$), thus recovering the abstract formulation of Section A.2. We consider an elliptic equation, for example the equation of steady heat transfer with symmetric thermal diffusivity tensor $\hat{k}_{ij}$ (Poisson equation when $\hat{k}_{ij} = 1$) in $\hat{\Omega}^r \subset \mathbb{R}^d$, for $r = 1, \ldots, 4$ with boundary $\hat{\Gamma}$. The field $\hat{u}$ (i.e. for example temperature) satisfies the partial differential equation

$$ -\frac{\partial}{\partial \hat{x}_i} (\hat{k}_{ij} \frac{\partial \hat{u}}{\partial \hat{x}_j}) = \hat{f} \text{ in } \hat{\Omega}, \quad (A.6.1) $$

where $1 \leq i, j \leq 2$ and with boundary conditions (see Figure A.1)

$$ \hat{u} = 0 \text{ on } \hat{\Gamma}_D, \quad \hat{k}_{ij} \frac{\partial \hat{u}}{\partial \hat{x}_j} \hat{n}_i = \hat{b} \text{ on } \hat{\Gamma}_N, \quad (A.6.2) $$

where $\hat{f}$ can be seen as the rate of heat generated per unit volume, $\hat{b}$ is the prescribed heat flux input on the surface $\hat{\Gamma}_N$, and $\hat{n}_i$ is the $i$-th component of the unit outward normal. We now derive the weak form of the governing equations. We introduce the function space:

$$ \hat{Y} = \{ \hat{v} \in H^1(\hat{\Omega}) | \hat{v} = 0 \text{ on } \hat{\Gamma}_D \}, \quad (A.6.3) $$

Figure A.1: Scheme for the reduced-basis multi-parameter problem.
and the associated norm

$$||\hat{v}||_{\hat{Y}} = \left( \int_{\Omega} \sum_{i=1}^{d} \left( \frac{\partial \hat{u}}{\partial \hat{x}_i} \right)^2 d\hat{\Omega} \right)^{1/2}.$$  \hfill (A.6.4)

Multiplying (A.6.1) by a test function \( \hat{v} \in \hat{Y} \), integrating over \( \hat{\Omega} \) by parts and applying the divergence theorem we have for \( 1 \leq i, j \leq 2 \):

$$- \int_{\hat{\Omega}} \hat{v} \frac{\partial}{\partial \hat{x}_i} \left( \hat{k}_{ij} \frac{\partial \hat{u}}{\partial \hat{x}_j} \right) d\hat{\Omega} = - \int_{\hat{\Gamma}} \hat{v} \hat{k}_{ij} \frac{\partial \hat{u}}{\partial \hat{x}_j} n_i d\hat{\Gamma} + \int_{\hat{\Omega}} \frac{\partial \hat{v}}{\partial \hat{x}_i} \hat{k}_{ij} \frac{\partial \hat{u}}{\partial \hat{x}_j} d\hat{\Omega} = \int_{\hat{\Omega}} \hat{f} \hat{v} d\hat{\Omega} \quad \forall \hat{v} \in \hat{Y}. \hfill (A.6.5)$$

Substituting boundary condition and by the fact that \( \hat{v} = 0 \) on \( \hat{\Gamma}_D \), we obtain as our weak statement

$$\langle \hat{A} \hat{u}, \hat{v} \rangle = \langle \hat{F}, \hat{v} \rangle, \quad \forall \hat{v} \in \hat{Y},$$  \hfill (A.6.6)

where

$$\langle \hat{A} \hat{w}, \hat{v} \rangle = \int_{\hat{\Omega}} \frac{\partial \hat{w}}{\partial \hat{x}_i} \hat{k}_{ij} \frac{\partial \hat{w}}{\partial \hat{x}_j} d\hat{\Omega},$$  \hfill (A.6.7)

$$\langle \hat{F}, \hat{v} \rangle = \langle \hat{F}_f, \hat{v} \rangle + \langle \hat{F}_b, \hat{v} \rangle,$$  \hfill (A.6.8)

here,

$$\langle \hat{F}_f, \hat{v} \rangle = \int_{\hat{\Omega}} \hat{f} \hat{v} d\hat{\Omega}, \quad \langle \hat{F}_b, \hat{v} \rangle = \int_{\hat{\Gamma}_N} \hat{b} \hat{v} d\hat{\Gamma}.$$  \hfill (A.6.9)

In our case \( \hat{\Omega} = \bigcup_{r=1}^{R} \hat{\Omega}^r, \ R = 4 \), so that the weak statement takes the form A.6.6 where

$$\langle \hat{A} \hat{w}, \hat{v} \rangle = \sum_{r=1}^{R} \int_{\hat{\Omega}^r} \frac{\partial \hat{w}}{\partial \hat{x}_i} \hat{k}_{ij}^r \frac{\partial \hat{w}}{\partial \hat{x}_j} d\hat{\Omega},$$  \hfill (A.6.10)

$$\langle \hat{F}, \hat{v} \rangle = \langle \hat{F}_f, \hat{v} \rangle + \langle \hat{F}_b, \hat{v} \rangle,$$  \hfill (A.6.11)
and

$$\langle \hat{F}_f, \hat{v} \rangle = \sum_{r=1}^{R} \int_{\hat{\Omega}^r} \hat{f}^r \hat{v} d\hat{\Omega}, \quad \langle \hat{F}_b, \hat{v} \rangle = \sum_{r=1}^{R} \int_{\Gamma^r} \hat{b}^r \hat{v} d\hat{\Gamma}. \quad (A.6.12)$$

### A.6.3 Affine mapping

The partition in subdomains $\hat{\Omega}^r$ is done such that there exists a reference domain $\Omega$ where, for any $\hat{x} \in \hat{\Omega}^r$, $r = 1, \ldots, R$, its image $x \in \Omega^r$ is given by

$$x = G^r(\mu; \hat{x}) = G^r(\mu) \hat{x} + g^r, \quad 1 \leq r \leq R;$$

we thus write

$$\frac{\partial}{\partial \hat{x}_i} = \frac{\partial x_j}{\partial \hat{x}_i} \frac{\partial}{\partial x_j} = G_{ji}(\mu) \frac{\partial}{\partial x_j}. \quad (A.6.14)$$

where $x \in \Omega$, $\hat{x} \in \hat{\Omega}$, $G^r(\mu) \in \mathbb{R}^{d \times d}$ is a piecewise-constant matrix, $g^r(\mu) \in \mathbb{R}^d$ is a piecewise-constant vector, and $G(\mu) : \hat{\Omega} \rightarrow \Omega$ is a piecewise-affine geometric mapping. We then denote the boundary of $\Omega$ as $\Gamma$, where $\Gamma(\mu, \hat{\Gamma})$. We now define the function space $Y$ as $Y(\Omega) = \overline{Y}(G^{-1}(\mu; \Omega)) = \overline{Y}(\hat{\Omega})$ such that

$$Y = \{ v \in H^1(\Omega) | v = 0 \text{ on } \Gamma_D \}, \quad (A.6.15)$$

and for any function $\overline{w} \in \overline{Y}$, we define $w \in Y$ such that $w(x) = \overline{w}(G^{-1}(\mu; x))$. Furthermore, we have

$$d\hat{\Omega} = detG^{-1}(\mu)d\Omega, \quad d\Gamma = |G^{-1}(\mu)\hat{t}|d\Gamma; \quad (A.6.16)$$

where $\hat{t}$ is the unit vector tangent to the boundary $\Gamma$, and

$$|G^{-1}\hat{t}| = \left( \sum_{i=1}^{d} (G_{ij}\hat{t}_j)^2 \right)^{1/2}. \quad (A.6.17)$$

It then follows that $\langle A(\mu)w, v \rangle = \langle \hat{A}\hat{w}, \hat{v} \rangle$ and $A(\mu)$ given by

$$\langle A(\mu)w, v \rangle = \sum_{r=1}^{R} \int_{\Omega^r} \left( G_{ij}^r(\mu) \frac{\partial w}{\partial x_i} \right) \hat{k}_{ij}^r \left( G_{ji}^r(\mu) \frac{\partial v}{\partial x_j} \right) det(G^r(\mu))^{-1} d\Omega, \quad (A.6.18)$$

or

$$\langle \hat{A}(\mu)w, v \rangle = \sum_{r=1}^{R} \int_{\hat{\Omega}^r} \frac{\partial w}{\partial x_i} \left( G_{ij}^r(\mu) \hat{k}_{ij}^r G_{ji}^r(\mu)det(G^r(\mu))^{-1} \right) \frac{\partial v}{\partial x_j} d\Omega \quad \forall w, v \in Y, \quad (A.6.19)$$

and $\langle F(\mu)w, v \rangle = \langle \hat{F}\hat{w}, \hat{v} \rangle$ and $F(\mu)$ given by

$$\langle F(\mu), v \rangle = \langle F_f, v \rangle + \langle F_b, v \rangle, \quad (A.6.20)$$

where

$$\langle F_f, v \rangle = \sum_{r=1}^{R} \int_{\Omega^r} \left( \hat{f}^r det(G^r(\mu))^{-1} \right) v d\Omega, \quad \langle F_b, v \rangle = \sum_{r=1}^{R} \int_{\Gamma^r} \left( \hat{b}^r |(G^r(\mu))^{-1}\hat{t}| \right) v d\Gamma. \quad (A.6.21)$$
The abstract problem is then recovered for
\[
\langle Aw, v \rangle = \sum_{r=1}^{R} \int_{\Omega} \frac{\partial w}{\partial x_i} k_{ij}^r \frac{\partial v}{\partial x_j} d\Omega \quad \forall w, v \in Y, \quad (A.6.22)
\]
\[
\langle F(\mu), v \rangle = \langle F_f(\mu), v \rangle + \langle F_b(\mu), v \rangle, \quad (A.6.23)
\]
\[
\langle F_f, v \rangle = \sum_{r=1}^{R} \int_{\Omega} f^r v d\Omega, \quad \langle F_b, v \rangle = \sum_{r=1}^{R} \int_{\Gamma_{N^r}} b^r v d\Gamma. \quad (A.6.24)
\]

Here \( k_{ij}^r(\mu) \) is given by
\[
k_{ij}^r = G_{ii'}^r(\mu) \hat{k}_{ij'}^r G_{jj'}^r(\mu) \text{det}(G^r(\mu))^{-1}, \quad (A.6.25)
\]
and \( b^r(\mu), f^r(\mu) \) are given by
\[
f^r(\mu) = \hat{f}^r \text{det}(G^r(\mu))^{-1}, \quad b^r(\mu) = \hat{b}^r |(G^r(\mu))^{-1}|. \quad (A.6.26)
\]
Furthermore, we may define
\[
\Theta^{q(i,j,r)}(\mu) = k_{ij}^r(\mu), \quad \langle A^{q(i,j,r)} w, v \rangle = \int_{\Omega^r} \frac{\partial v}{\partial x_i} \frac{\partial w}{\partial x_j} d\Omega, \quad (A.6.27)
\]
for \( 1 \leq r \leq R, 1 \leq i, j \leq d \).

### A.6.4 Model problem

We now consider our model problem in detail. As already said, the problem can be seen as a problem involving the flow of heat in a T-shaped region containing an internal heat source as shown in Figure A.2. Our output of interest is

\[
s(\mu) = \sum_{r=1}^{R} \frac{1}{\xi_r} \int_{\Omega^r} \hat{u}^r d\Omega, \quad (A.6.28)
\]

for \( \mu = \{t, D, L, S, \theta\} \in D^\mu \subset \mathbb{R}^P \), where \( \xi_r \) is a normalizing factor (related to geometrical quantities in the subdomains \( \Omega^r \)). Our problem can then be formulated as: given a \( \mu \in D^\mu \subset \mathbb{R}^P \), find \( s(\mu) = \langle \hat{L}, \hat{u} \rangle \) where \( \hat{u} \in \hat{Y} \) is the solution to:

\[
\langle \hat{A} \hat{w}, \hat{v} \rangle = \langle \hat{F}, \hat{v} \rangle, \quad \forall \hat{v} \in \hat{Y}; \quad (A.6.29)
\]

here, \( \langle \hat{L}, \hat{v} \rangle = \langle \hat{F}, \hat{v} \rangle, \quad \forall \hat{v} \in \hat{Y} \). In our case \( D^\mu \) is given by \([t_{\min}, t_{\max}] \times [D_{\min}, D_{\max}] \times [L_{\min}, L_{\max}] \times [S_{\min}, S_{\max}] \times [\theta_{\min}, \theta_{\max}] \), i.e: \([0.1, 1.5] \times [0.1, 1.5] \times [0.1, 5.0] \times [0.1, 5.0] \times [0^\circ, 60^\circ] \). We have \( \hat{b} = 0, \hat{f} = \{1\} \) defined on \( \hat{\Omega}^r \), so that:

\[
\xi_1 = t, \quad \xi_2 = SD, \quad \xi_3 = tD, \quad \xi_4 = LD. \quad (A.6.30)
\]

The affine mapping \( G(\hat{x})(\mu) : \hat{\Omega} \to \Omega \) is given by (A.6.13) and we have
\[
G^1(\mu) = \begin{pmatrix} 1 & -\tan(\theta) \\ 0 & 1 \end{pmatrix}, \quad G^2(\mu) = \begin{pmatrix} \frac{1}{\theta} & 0 \\ 0 & \frac{1}{\theta} \end{pmatrix}. \quad (A.6.31)
\]
A.6. A MULTI-PARAMETER APPLICATION

\[
G^3(\mu) = \begin{pmatrix}
\frac{1}{\mu} & 0 \\
0 & \frac{1}{\mu}
\end{pmatrix}, \quad G^4(\mu) = \begin{pmatrix}
\frac{1}{\mu} & 0 \\
0 & \frac{1}{\mu}
\end{pmatrix}.
\]  

(A.6.32)

Each mapping function is defined in \( \hat{\Omega}' \), furthermore \( g^r = 0 \ \forall r \). We have

\[
d\hat{\Omega}^1 = det(G^1)^{-1}(\mu)d\Omega^1 = td\Omega, \quad d\hat{\Gamma}^1 = |det(G^1)^{-1}(\mu)|d\Gamma = td\Gamma, \]  

(A.6.33)

\[
d\hat{\Omega}^2 = det(G^2)^{-1}(\mu)d\Omega^2 = LDd\Omega, \quad d\hat{\Gamma}^2 = |det(G^2)^{-1}(\mu)|d\Gamma^2 = LDd\Gamma, \]  

(A.6.34)

\[
d\hat{\Omega}^3 = det(G^3)^{-1}(\mu)d\Omega^3 = tDd\Omega, \quad d\hat{\Gamma}^3 = |det(G^3)^{-1}(\mu)|d\Gamma^3 = tDd\Gamma, \]  

(A.6.35)

\[
d\hat{\Omega}^4 = det(G^4)^{-1}(\mu)d\Omega^4 = SDd\Omega, \quad d\hat{\Gamma}^4 = |det(G^4)^{-1}(\mu)|d\Gamma^4 = SDd\Gamma, \]  

(A.6.36)

Figures A.3 and A.4 show the affine mapping procedures on two subdomains (\( \Omega^1 \) and \( \Omega^3 \) for example).

---

Figure A.3: Scheme of parameter-dependent domain \( \Omega^1 \) undergoing vertical shear and rotation.

We may now re-formulate our problem in terms of our reference domain: find \( s(\mu) = \langle L, u \rangle \) where \( u \in Y \) is the solution to:

\[
\langle Aw, v \rangle = \langle F, v \rangle, \forall v \in Y, \]  

(A.6.37)

where \( \langle L, v \rangle = \langle F, v \rangle \ \forall v \in Y \) and

\[
\langle Aw, v \rangle = \sum_{r=1}^{R} \langle A_r w, v \rangle = \langle F, v \rangle, \forall v \in Y, \]  

(A.6.38)

\[
\langle F, v \rangle = \sum_{r=1}^{R} \int_{\Omega_r} vd\Omega, \ \forall v \in Y, \]  

(A.6.39)
Figure A.4: Scheme for parameter-dependent domain \( \Omega^3 \) and reference domain undergoing both stretch and shear.

\[
\langle A'w, v \rangle = \int_{\Omega_r} \frac{\partial v}{\partial x_i} k_{ij}^r \frac{\partial w}{\partial x_j} d\Omega, \ \forall w, v \in Y, \quad (A.6.40)
\]

and the effective diffusivity tensors \( k_{ij}^r(\mu) = G_{ij}'(\mu)k_{ij}^r(\mu)G_{jj}'(\mu)\det G^{-1}(\mu) \) are given by:

\[
k^1 = \begin{bmatrix} t & -\tan \theta \\ -\tan \theta & \frac{1+\tan^2 \theta}{t} \end{bmatrix}; \quad (A.6.41)
\]

\[
k^2 = \begin{bmatrix} \frac{S}{\gamma} & 0 \\ 0 & \frac{\rho}{\gamma} \end{bmatrix}; \quad (A.6.42)
\]

\[
k^3 = \begin{bmatrix} \frac{t}{\gamma} & 0 \\ 0 & \frac{\rho}{\gamma} \end{bmatrix}; \quad (A.6.43)
\]

\[
k^4 = \begin{bmatrix} \frac{L_{ij}}{\gamma} & 0 \\ 0 & \frac{\rho}{\gamma} \end{bmatrix}. \quad (A.6.44)
\]

The abstract problem formulation is then given for \( P = 5 \) (number of parameters), \( R = 4 \) (subdomains) and \( Q = 9 \) (different bilinear forms contributes to \( A \) over different portions of domain).

### A.7 Some numerical results

In this section we present some numerical results obtained with the configuration previously described. We used a posteriori error bounds and adaptive procedures in basis-building process. Table A.1 shows a preliminary test based on two-parameters configuration and adaptivity procedures. We can see that the dimension of \( N \) is very small in reduced-basis
model. With adaptive procedures we get a 60% saving in computational cost related to off-line procedures using a triangulation with $O(10^3)$ elements. The information about the error bound is provided by $\Delta_N$. Table A.2 gives us information about convergence ($\Delta_N$), varying $N$.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\Delta_N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.1</td>
</tr>
<tr>
<td>9</td>
<td>0.01</td>
</tr>
<tr>
<td>13</td>
<td>0.001</td>
</tr>
<tr>
<td>20</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Table A.1: $\Delta_N$ and $N$ using two-parameters configuration ($L$ and $D$) and adaptive procedure (Section A.5). Without adaptive procedure we would need $N = 50$ to get a $\Delta_N < 10^{-3}$, we reach a consistent computational load reduction.

and about the effectivity $\eta_N$, studying a complete five-parameters configuration. With $N = 45$, for example, we get $\Delta_N \leq 10^{-4} = \varepsilon_d^{\text{prior}}$. Effectivity is near unity (always $\geq 1$). The results are obtained testing at least 1000 different configurations $\Sigma^{\text{prior}}$ and different parameters combinations (see Section A.5), the ones in the table are the worst gotten testing all the random configurations considered. It’s important to underline that the adaptive procedure tested in the reduced-basis off-line building phase has permitted us to keep under control the condition number of the reduced-basis matrix, avoiding ill-conditioning problems caused by the random choice of parameters.

Figure A.5 shows the solution of the problem given by finite element method and by reduced basis for a certain combination of parameters (i.e $L = 1.0$, $D = 1.5$, $S = 1.0$, $t = 1.31$, $\theta = 16^\circ$). Figure A.6 shows the distribution of the error $\hat{e}$ in $\Omega$ (difference between $u_N$ and $\tilde{u}$) using $N = 60$ basis. This is not the error over the output $s(\mu)$ calculated by a posteriori error bounds, but it can be considered as a good indication how reduced-basis method is able to provide results close to the Finite-Elements solution. Figure A.7 shows the convergence of the method used in logarithmic scale increasing $N$, while Figure A.8 shows the upper and lower bound error estimator and their convergence to the true value of the output quantity of interest. Figures A.9 and A.10 represent effectivity $\eta_N$ and the upper ($\eta_{\text{max}}$) and lower limit ($\eta_{\text{min}}$, i.e. unity). Figures A.11 and A.12 show parameters distribution in the parameters spaces during the off-line reduced-basis matrix construction. The former shows a two-parameters distribution region ($D^\mu = \mathbb{R}^2$) around the first couple of parameters chosen in the centre of the region. Note that the parameters are distributed around the original starting point. The latter is a possible five-parameters distribution in the case studied applying the adaptive procedure.
Table A.2: $\Delta_{N_{max}}$ and $N$ using five-parameters configuration and adaptivity procedure to get a $\Delta_{N_{max}} < 10^{-4} = \varepsilon_d^{prior}$ in the worst case. Note values of the effectivity $\eta_N$. 

<table>
<thead>
<tr>
<th>$N$</th>
<th>$\eta_N$</th>
<th>$\Delta_{N_{max}}$</th>
<th>$N$</th>
<th>$\eta_N$</th>
<th>$\Delta_{N_{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.1438</td>
<td>0.064589558</td>
<td>31</td>
<td>1.2802</td>
<td>0.000234237</td>
</tr>
<tr>
<td>2</td>
<td>4.4864</td>
<td>0.051445346</td>
<td>32</td>
<td>1.4427</td>
<td>0.000220116</td>
</tr>
<tr>
<td>3</td>
<td>4.6687</td>
<td>0.052944041</td>
<td>33</td>
<td>1.4321</td>
<td>0.000192997</td>
</tr>
<tr>
<td>4</td>
<td>4.5023</td>
<td>0.050876316</td>
<td>34</td>
<td>1.3848</td>
<td>0.000171898</td>
</tr>
<tr>
<td>5</td>
<td>4.0258</td>
<td>0.043368854</td>
<td>35</td>
<td>1.6566</td>
<td>0.000189594</td>
</tr>
<tr>
<td>6</td>
<td>2.9801</td>
<td>0.025041318</td>
<td>36</td>
<td>1.651</td>
<td>0.000176161</td>
</tr>
<tr>
<td>7</td>
<td>2.521</td>
<td>0.019830625</td>
<td>37</td>
<td>1.814</td>
<td>0.000185761</td>
</tr>
<tr>
<td>8</td>
<td>2.3531</td>
<td>0.018128973</td>
<td>38</td>
<td>1.8877</td>
<td>0.000192532</td>
</tr>
<tr>
<td>9</td>
<td>1.8848</td>
<td>0.012419358</td>
<td>39</td>
<td>1.9003</td>
<td>0.000187474</td>
</tr>
<tr>
<td>10</td>
<td>1.7454</td>
<td>0.010821934</td>
<td>40</td>
<td>1.8567</td>
<td>0.000179674</td>
</tr>
<tr>
<td>11</td>
<td>1.5322</td>
<td>0.008960269</td>
<td>41</td>
<td>1.6172</td>
<td>0.000140013</td>
</tr>
<tr>
<td>12</td>
<td>1.5339</td>
<td>0.008963254</td>
<td>42</td>
<td>1.6207</td>
<td>0.000139235</td>
</tr>
<tr>
<td>13</td>
<td>1.5327</td>
<td>0.008874298</td>
<td>43</td>
<td>1.6361</td>
<td>0.000140171</td>
</tr>
<tr>
<td>14</td>
<td>1.5052</td>
<td>0.008636955</td>
<td>44</td>
<td>1.7266</td>
<td>0.000146752</td>
</tr>
<tr>
<td>15</td>
<td>2.2945</td>
<td>0.008465466</td>
<td>45</td>
<td>1.6297</td>
<td>0.000101771</td>
</tr>
<tr>
<td>16</td>
<td>2.1233</td>
<td>0.008245741</td>
<td>46</td>
<td>1.6366</td>
<td>9.74091E-05</td>
</tr>
<tr>
<td>17</td>
<td>2.0121</td>
<td>0.007101543</td>
<td>47</td>
<td>1.6187</td>
<td>9.63095E-05</td>
</tr>
<tr>
<td>18</td>
<td>1.9571</td>
<td>0.005058438</td>
<td>48</td>
<td>1.6283</td>
<td>9.64255E-05</td>
</tr>
<tr>
<td>19</td>
<td>2.2285</td>
<td>0.003577518</td>
<td>49</td>
<td>1.481</td>
<td>8.87554E-05</td>
</tr>
<tr>
<td>20</td>
<td>2.3294</td>
<td>0.003595009</td>
<td>50</td>
<td>1.4757</td>
<td>8.78055E-05</td>
</tr>
<tr>
<td>21</td>
<td>1.9196</td>
<td>0.001392253</td>
<td>51</td>
<td>1.4802</td>
<td>8.46522E-05</td>
</tr>
<tr>
<td>22</td>
<td>1.882</td>
<td>0.00132827</td>
<td>52</td>
<td>2.8702</td>
<td>8.46041E-05</td>
</tr>
<tr>
<td>23</td>
<td>1.673</td>
<td>0.000950676</td>
<td>53</td>
<td>2.9196</td>
<td>8.46041E-05</td>
</tr>
<tr>
<td>24</td>
<td>1.6219</td>
<td>0.000783639</td>
<td>54</td>
<td>2.8235</td>
<td>8.46041E-05</td>
</tr>
<tr>
<td>25</td>
<td>1.655</td>
<td>0.000795108</td>
<td>55</td>
<td>2.6366</td>
<td>8.47554E-05</td>
</tr>
<tr>
<td>26</td>
<td>1.2313</td>
<td>0.000360924</td>
<td>56</td>
<td>2.7151</td>
<td>7.8055E-05</td>
</tr>
<tr>
<td>27</td>
<td>1.1243</td>
<td>0.000301648</td>
<td>57</td>
<td>2.6354</td>
<td>7.65224E-05</td>
</tr>
<tr>
<td>28</td>
<td>1.1096</td>
<td>0.000269211</td>
<td>58</td>
<td>2.6772</td>
<td>7.46041E-05</td>
</tr>
<tr>
<td>29</td>
<td>1.108</td>
<td>0.000268716</td>
<td>59</td>
<td>2.3726</td>
<td>4.26101E-05</td>
</tr>
<tr>
<td>30</td>
<td>1.1698</td>
<td>0.000275998</td>
<td>60</td>
<td>2.4882</td>
<td>4.51111E-05</td>
</tr>
</tbody>
</table>
Figure A.5: Solution using FEM-Galerkin method for a 5-parameters configuration (left) and solution using reduced basis method with $N = 60$.

Figure A.6: Distributed error $\tilde{e}$ over the domain $\Omega$ between $\tilde{u}$ and $u_N$. 
Figure A.7: Convergence of the difference $s_N - \tilde{s}$ increasing N.

Figure A.8: Error Bounds $s_N^+, s_N^-$ and their convergence to $s(\ast)$.
Figure A.9: Effectivity coefficient $\eta_N$ and its bounds $\eta_{\text{min}} (= 1)$ and $\eta_{\text{max}}$.

Figure A.10: Effectivity coefficient $\eta_N$ with lower Bound (unity). Zoom.
Figure A.11: 2-Parameters \((t, D)\) distribution using adaptivity in basis building off-line process. The starting configuration is in the centre of the 2-parameters space.

Figure A.12: 5-parameters distribution using adaptivity for reduced-basis matrix assembling.
A.8 Developments on reduced basis

In Chapters 5, 6 and 7 we extend the reduced basis technique to Stokes and Navier-Stokes equations and to problems involving also non-affine mapping (i.e. shape design). See also Grepl [44] and Solodukhov [145]. We are interested in studying a complete coarse configuration for the problem of aorto-coronaric bypass optimization but the procedures can be generalized for other design problems in engineering.

A.8.1 Reduced basis for the study of a complete geometrical configuration

The first step concerns the replacement of the elliptic state equation with Stokes equation (with Dirichlet in-flow condition and Neumann out-flow condition), see also Rovas [134]. The main steps are the following.

- Reformulation of the problem (new state equations, new bilinear forms) with the same geometry (4 subdomains) described in this appendix.
- Test of the bypass model with “macro” parameters in the haemodynamics background (5 macro-parameters) and a preliminary geometrical sensitivity analysis (diameter, stenosis length, bypass angle,...).
- Test on a possible starting configuration to apply the tools of shape optimization and flow control (see Chapter 2-4 and [123]) with a more complex output of interest (e.g. vorticity).
- At the end of the study we get an optimized bridge configuration and useful indication for bypass construction. See Chapter 5.

A.8.2 Reduced basis for pre-process optimization

The problem can be studied at an intermediate level with an approach verified by feedback procedures (a shell model).

- We use reduced-basis techniques to get a preliminary configuration for the bypass problem (coarse configuration) in a pre-process procedure for optimization.
- We apply tools for optimization based on flow control and optimal shape design technique, introduced in Chapter 2-4, using the starting configuration gotten from reduced basis model application. We use steady and unsteady Stokes equations.
- We take the new configuration (the local configuration optimized by shape design tools) for a further feedback using unsteady Navier-Stokes equations and other output of interest (such as unsteady quadratic functionals related to wall shear stress oscillations).

In conclusion at this step we have a “shell” model with three inner feedback procedure (see Figure A.13):
- Reduced-Basis method on a coarse “bridge” configuration.

- Optimal Shape design and Flow Control on fine configuration.

- Unsteady quantities and Navier Stokes equations as a feedback test for the final configuration.

Figure A.13: “Shell” model scheme made up by consecutive steps.

A.8.3 Reduced basis for a shape optimization problem

The last step is the full extension of reduced basis model using more than five parameters and modelling the (curved) wall using non-affine mapping on reference domain. In this case
A.9. THE MODEL-ORDER REDUCTION

Reduced basis can be used not only in the framework of a pre-process for optimization but as a shape optimization tool in itself. This aspect is considered in chapter 6. Then the introduction of non-linearities with the convective term allows us to study Navier-Stokes equation using reduced basis in parametrized domains. At this stage we can also carry out a comparison between the low fidelity fluid model based on the Stokes equation and the high fidelity one based on the Navier-Stokes equations. This aspect is developed in Chapter 7.

A.8.4 Reduced basis for optimal flow control

An other step to complete the present work is the creation of a link between optimal control (based on adjoint formulation) and reduced basis used to solve both state and adjoint problem with geometrical shape parameters. In this last case optimization would be driven by optimal control tools, while approximation by reduced-basis model. See Table A.3. This aspect is developed in Chapter 8.

<table>
<thead>
<tr>
<th>Step</th>
<th>First Option</th>
<th>Second Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem Approximation</td>
<td>Galerkin-Finite Element and Reduced-basis for state equation</td>
<td>Galerkin-Finite Element and Reduced-basis for state and adjoint equation</td>
</tr>
<tr>
<td>Optimization</td>
<td>Sensitivity Analysis</td>
<td>Optimal Control (adjoint formulation)</td>
</tr>
</tbody>
</table>

Table A.3: Third step future developments options.

It is evident that to expand and apply reduced-basis theory on biomechanics problems (i.e. biomedical devices such as bypass) the two most important phases are the use of a great number of geometrical parameters (non-affine mapping) and the use of Navier-Stokes equations to model fluid flow.

A.9 The model-order reduction

We give at the end of this appendix some relevant background (and historical) information on model-order reduction. The issue of reducing complexity while preserving all relevant information, has been a very active research area in many disciplines. A characteristic of systems whose behavior is governed by partial differential equations is that the resulting state models, obtained by a discretization procedure, are of very high-dimension. Therefore some of the existing methods developed, for example in control systems theory, are not directly applicable. In the last period the problem of model-order reduction is becoming more and more relevant due to the versatility that it allows in the parametrization of the problem. See for example the recent review work in *SIAM News* [51] by Gunzburger and Willcox. We
summarize in Section A.9.1 recent developments and relevant approaches. The references provided, although by no means exhaustive, should cover most of the recent work. As model-order reduction methods are by definition pre-asymptotic, validation of the obtained results has been recognized to be a critical ingredient. Even though residual-based error measures have been suggested, rigorous a posteriori error estimation procedures have been developed in the last years. In other contexts, like estimation of the discretization error in finite-element analysis, a plethora of a posteriori error estimation methods had already been existent for years. Some of these methods are relevant for our problems; we discuss in section A.9.2 the connection and differences between them.

A.9.1 Earlier Work

Proper Orthogonal Decomposition

We start our discussion with the proper orthogonal decomposition method, probably the most popular model-order reduction technique. Underlying this method is the solution of the following approximation problem: given a (possibly large) set of vectors, identify the best approximating $N$-dimensional plane (subspace) such that the root-mean square $L_2$-projection error is minimized. A solution to this problem can be obtained using the singular value (or Karhunen-Loève) decomposition [86].

The proper orthogonal decomposition has been applied and (re-)discovered in many different areas: system dynamics, stochastic processes, image processing, to name a few. For reduction of physical systems, it has been extensively applied to time-dependent problems. In this case, time is considered as the varying parameter, and “snapshots” of the field variable (e.g. temperature, displacement) at different times — parameter points — are obtained using numerical or experimental procedures. The optimal $N$-dimensional approximation space (for $N$ small) is constructed by applying the singular-value decomposition to these vectors, and keeping only the $N$ singular vectors corresponding to the largest singular values. As the singular values are related to the total “energy” of the approximation, these modes can be identified as the ones preserving most of the energy. The reduced model is then obtained by using a Galerkin projection to the space spanned by these vectors. More precisely, the POD algorithm is built as follows (see, for example, [50] and [132]):

- we start from a reduced basis $\{\phi_1, ..., \phi_M\}$, a basis of finite element solutions associated to some “snapshots” (chosen parameters), where $\tilde{M}$ is usually $O(100)$, and we compute the mean value

$$\bar{\phi} = \frac{1}{\tilde{M}} \sum_{j=1}^{\tilde{M}} \phi_j.$$ 

We assume that $\phi_i$, $i = 1, ..., \tilde{M}$ has been computed by finite element method: $\phi_i \in \mathbb{R}^{\mathcal{N}}$, dove $\mathcal{N} \gg 1$ is the finite element space dimension.

- We build a correlation matrix $C : \tilde{M} \times \tilde{M}$ (symmetric and positive semidefinite), whose elements are $C_{i,j} = \frac{1}{\tilde{M}} (\phi_i - \bar{\phi})^T (\phi_j - \bar{\phi})$, $i, j = 1, ..., \tilde{M}$. 

We compute eigenvalues \( \{ \lambda_1, ..., \lambda_M \} \) and the corresponding eigenvectors \( \{ v_1, ..., v_M \} \) of \( C \), ordering the eigenvalues by decreasing size.

We choose \( M \):

\[
M = \min_{m \in [1, M]} \frac{\sum_{j=1}^{m} \lambda_j}{\sum_{j=1}^{M} \lambda_j} \geq \gamma,
\]

where \( 0 \leq \gamma \leq 1 \) is a fixed tolerance.

We set \( \Phi_i = \sum_{j=1}^{M} (v_i)_j (\phi_j - \bar{\phi}) \), for \( i = 1, ..., M \). \( (v_i)_j \) is the \( j \)-th component of the \( i \)-th eigenvector.

We normalize \( \hat{\Phi}_i = \frac{\Phi_i}{\| \Phi_i \|} \), where \( |\Phi_i|^2 = \Phi_i^T \Phi_i \).

The set \( B = \text{span}\{ \hat{\Phi}_1, ..., \hat{\Phi}_M \} \) is the POD basis (orthonormal, i.e. \( \Phi_i^T \Phi_j = \delta_{ij} \)); the POD approximation is: find \( u_M(\mu) \in B \) so that

\[
a(u_M(\mu), v; \mu) = f(v), \quad \forall v \in B. \tag{A.9.1}
\]

The optimality property and generality of these ideas, has led to the successful application of the method in many areas: turbulent flows [89], fluid structure-interaction [34], non-linear structural mechanics [73], turbo-machinery flows [166]. Extension of these methods to general multi-parameter problems has been quite limited. The problem is that the singular values are not system invariants as they depend on the choice of “snapshots” and the particular configuration in consideration. It has been observed that reduced-order models obtained for one configuration were not optimal for other configurations; using such models often lead to inaccurate or, even worse, incorrect results. It has been suggested in [24] to give more weight or preselect some of the vectors in the starting basis, leading to “weighted-POD” or “predefined-POD” methods, but the selection of the required weights is not automatic limiting the generality of such approaches.

An analysis of the model-truncation error suggests that the error can be attributed to two sources: first in the inability of the low-order model to reproduce the exact loading; and second, for the approximated loading, in the inability of the low-order model to recover the exact solution [72] (see also [114] for similar ideas). Using terms from control-systems theory, the first error is related to the controllability (primal) and the second to the observability (dual) of the low-order model. In a similar manner, for our methods, we introduce at the end of Chapter 5 a combined primal-dual approach to estimate both of these errors. The notion that a truncation of the model should balance both of these errors, led to balanced-truncation methods [98].

**Reduced Basis Methods**

We turn now to reduced-basis approaches, upon which our method is also based. The reduced-basis method has been proposed in [8, 102] for the non-linear analysis of structures. In these approaches, only single-parameter problems were considered and the method was viewed...
as a continuation procedure. The method has been further investigated and extended by Noor [105, 106], who realized that the method could be applied for general multi-parameter problems. Much of the earlier work focused: first, on the selection and efficient computation of basis functions; and second, on validation of the efficiency and accuracy of reduced-basis approaches in a number of test problems.

As seen in the introduction of this appendix the reduced-basis method recognizes that the field variable is not, in fact, some arbitrary member of the infinite-dimensional solution space associated with the partial differential equation; rather, it resides, or “evolves”, on a much lower-dimensional manifold induced by the parametric dependence. In these earlier approaches, the approximation spaces for the low-dimensional manifold were typically defined “locally” — relative to a particular parameter point. Fink and Rheinboldt [37] placed the method in this geometric setting and carried out an error analysis for a general class of single-parameter problems. Porsching [117] considered Lagrangian, Taylor and discrete least squares approximation spaces, and extended some of the a priori analysis. In [38] a general local error estimation theory for single-parameter problems was developed containing the earlier estimates as special cases. The extension of the error analysis to multi-parameter problems was presented in [133]. Finally, evaluation of the constants that appear on the error bounds was considered in [13]. The a priori theory as developed in the works above concludes that, close to the parameter point selected for the construction of the reduced-basis spaces, the error converges to zero exponentially fast with the number of basis functions used.

Reduced basis approaches have been subsequently developed in many other areas. Peterson [113] applied it to fluid flow problems and the Navier-Stokes equations, and in [65, 64] it was used for control of fluid problems. The reduced-basis approach as earlier articulated was local in parameter space in both practice and theory. As a result, the computational improvements — relative to conventional (say) finite-element approximation — were often quite modest [117]. Balmes [11], was the first to consider general multi-parameter problems. In his approach he suggests choosing the basis functions by sampling globally in parameter space. The importance of error estimation has always been emphasized in the literature and rigorous validation methods have been developed by Patera, Maday et al. in [91], [92] and [134] and since then the range of reduced basis applications has increased a lot.

Other Methods

Krylov-subspace techniques like the Arnoldi or the Lanczos methods and their variants, have traditionally been used for the calculation of a small set of the extremal eigenvalues and eigenvectors for large-scale eigenproblems. But these are precisely the eigenvalues and eigenvectors of interest for model reduction. Many reduction approaches based on Krylov-subspace techniques have been developed; for an overview see [69] and the references contained therein. The iterative nature of the algorithms, makes it difficult to develop error bounds; moreover and the stability of the reduced-order problem is not always guaranteed.
A.9.2 Remarks on \textit{a posteriori} error estimation

The issue of \textit{a posteriori} error estimation and, more generally, validation of the numerical predictions has received considerable attention in the finite-element literature. The problem of interest is related to the choice of mesh to be used for the definition of the finite-element spaces. It is understood that there are certain trade-offs associated with the choice of the finite-element mesh: on one hand, a conservative choice, ensures high accuracy but also the computational costs become formidable; on the other hand, the choice of a relatively coarse mesh ensures efficiency but the accuracy is dubious. More to that, for a specific choice of mesh, the obtained accuracy is not easy to calculate as it depends on the topology of the mesh, the particular problem in consideration, the choice of finite-element spaces, or even the way we choose to measure the error. We can also relate a number of other problems like, for example, the choice of elements to be refined in adaptive refinement or, more generally the choice of “optimal” meshes (i.e. meshes which for a given accuracy minimize computational cost). For all these problems, the ability to estimate and therefore control, the discretization error is critical.

The extensive \textit{a priori} theory can not be used as the provided error bounds depend on norms of the exact solution which, in general, is not known. Rather, the \textit{a posteriori} error estimators give bounds which depend on computable quantities, like residuals. The study of these types of error estimators started in the 70s with the first paper by Babuska and Rheinboldt [10], and since then the literature has grown appreciably; a review can be found in [7]. Most error estimators developed give bounds for abstract norms of the error. Relevant to this thesis are \textit{a posteriori} error estimators directly for outputs of interest; see for example [109, 110] for relevant work.

For the reduced-basis method instead of the finite-element mesh and the discretization error, we have the parameter space “discretization”, and the reduced-basis approximation error; refinement of the mesh, corresponds to adding more basis-functions in the definition of the reduced basis global approximation space $W_N$. But there are also differences, the most important being the parameter-dependence of the operator, consideration of which is not required in the finite-element case. Even though the methodologies are distinctively different; some of the general ideas [110] for \textit{a posteriori} error estimation are common.
Bibliography


Index

a posteriori error estimation, 237
a posteriori error estimator, 122, 211
a priori convergence, 216
adaptive procedure, 218
adjoint correction, 122
adjoint equation, 175
adjoint pressure, 25
adjoint problem, 25, 50, 73, 119, 122
adjoint velocity, 25
advection field, 178
advection-diffusion problem, 173
affine decomposition, 93
affine mapping, 91, 150, 221
anastomosis, 1, 40
anastomotic flow, 36
aorto-coronaric bypass, 1
Armijo rule, 27
arterial diameter, 21, 118, 218
arterial thickness, 21
arterial wall, 135
assembling procedure, 97, 218
backward Euler, 199
basis function, 209
best approximation, 217
blood, 173
blood flow, 41
blood flow perturbation, 117
bound conditioner, 212
bound gap, 213
boundary observation, 23
boundary variations, 29
bypass, 139, 168
bypass bridge, 40
bypass diameter, 118, 218
bypass surgery, 1
cardiovascular system, 2
cavity flow, 100
characteristic function, 43
coercive operator, 207
complete anastomosis, 64
compliant output, 207
computational cost, 109, 161, 225
computational invertibility, 212
conditioning number, 225
continuity equation, 118
continuous operator, 207
control function, 5, 15
control input, 178, 183
control problem, 173
convergence, 217
coronaries, 79
coronary artery, 1
coronary disease, 4
correlation matrix, 234
cost functional, 5, 16, 45, 59, 174
CPU time, 109
curvature, 19, 139, 170
curved wall, 134
data control, 5
Dean number, 171
descent gradient-based method, 26
diastolic phase, 79
diffusivity, 178, 185
distributed observation, 23
distributed vorticity, 84
domain control, 5
domain optimization, 196
down-field zone, 23, 56
drugs, 173
dual error, 122
dual norm, 124
dual problem, 119
dual residual, 121
effectively affine decomposition, 131, 155
effectively affine dependence, 156
effectivity, 214, 217
elastic energy, 24
elasticity, 23
empirical interpolation, 127, 135, 155, 165
end-to-side anastomosis, 117
energy, 208
energy of the system, 31
equivalent inf-sup, 94, 152
error on control, 180
error projection, 96
error residual, 214
exact controllability equation, 45, 69
exact controllability problem, 46, 69, 70
existence results, 51, 74
feedback procedure, 35, 231
Finite Element method, 23
first corrections, 40, 58
first order perturbation, 59
first order problem, 45, 46
flow control, 5, 23, 181, 231
flow disturbances, 34
flow perturbation, 32, 45, 46, 117
flow rate, 17, 79
flow reversal, 81
flow separation, 34
forced flow, 100, 106
Frechet derivative, 149
free-stress condition, 20, 25
Freefem, 33
furrowed channel, 162
Galerkin approximation, 207
Galerkin method, 23
Galerkin orthogonality, 214
Galerkin projection, 234
generalized eigenvalue problem, 98
generalized optimal control problem, 70
generalized Stokes problem, 45
geometrical input, 178, 191
geometrical parameter, 7, 233
geometrical sensitivity, 195
global approximation spaces, 206
gradient method, 26
graft angle, 19, 34, 117, 171, 218
Gram-Schmidt orthogonalization, 110
Green Theorem, 37
grid reconstruction, 29
haemodynamics, 2
Hagen-Poiseuille, 20, 171
heart, 1
heat flux, 219
heat rate, 219
heat source, 222
heat transfer, 219
higher fidelity methods, 18
Hilbert space, 16
hyperplasia, 2
ill conditioning, 218
incoming branch, 18
incompressible flow, 31
inf-sup condition, 93, 131
input-output, 5, 117, 206
inputs, 206
integrated adjoint basis, 120
interpolation error, 129, 136, 155
interpolation points, 128, 155
intimal thickening hyperplasia, 36
iterative processes, 54, 77
Lagrangian basis, 148
Lagrangian function, 174
laminar flow, 18
life sciences, 173
local boundary variations, 59
local shape design, 39
low fidelity methods, 18, 171
low Reynolds, 41
lower bound, 213

macro geometrical optimization, 146
main stream, 37
mean velocity, 96, 117
Mean Wall Shear Stress Gradient MWSSG, 17
mesh stretching, 29
Miller cuff, 2
minimization problem, 45, 69
minimum coefficient bound conditioner, 213
model order reduction, 233
multi-level geometrical approaches, 1
multi-objective geometrical optimization, 24
multi-parameter, 218

Navier-Stokes equations, 35, 147, 149
Newton method, 149
Newtonian fluid, 18
non-affine dependence, 129
non-affine mapping, 154
non-compliant output, 123
non-integrated adjoint basis, 120
non-linear problem, 61
non-stationary Stokes equations, 63
normalizing factor, 222

observation function, 16, 51
occlusion, 218
optimal control, 196
optimal control for unsteady flows, 71
optimal control problem, 16, 46
optimal control theory, 15
optimal shape design, 5, 231
optimality conditions, 27, 50, 73
optimization process, 7
orthogonal basis, 110
orthogonal complement, 110
orthogonal projector, 110
orthonormalization, 110

Oscillatory Flow Index OFI, 17
Oscillatory Shear Index OSI, 17
outflow length, 218
output, 169, 205, 231
output bound, 211
output correction, 122
output sensitivities, 170
oxigen amount, 1
oxygenator, 136

Pécllet number, 199
parameter dependent function, 216
parameter distribution, 225
parameter independent operator, 212, 216
parametrized adjoint equation, 184, 188, 194
parametrized boundary, 63
parametrized domain, 90
parametrized flow control, 173
parametrized operator, 206
parametrized optimal control problem, 195
parametrized PDE, 91
parametrized state equation, 183, 187, 191
parametrized Stokes problem, 90, 129
performance measure, 119
period, 70
perturbation method, 63
perturbed functions, 66
physical input, 178, 187
physiological flow, 173
plateau, 136
pre-process, 231
predictor, 212
preliminary configuration, 231
pressure error projection $L^2$, 97
pressure mass matrix, 97
pressure matrix method, 149
primal problem, 120
primal residual, 121
proper orthogonal decomposition POD, 234
pulsatile flow, 79

radius of curvature, 171
real time solution, 168
recirculation, 19
reduced basis, 173
reduced basis approximation, 132, 151
reduced basis element, 90
reduced basis linear system, 133
reduced basis method, 7, 148, 206, 209
reduced basis pressure space, 94, 132, 152
reduced basis space, 210
reduced basis velocity space, 94, 132, 152
reference domain, 42, 129, 150, 219
regularization of the shape, 58
regularization parameter, 46
residual, 121
retrograde flow, 79
Reynolds, 18
Riesz’ Theorem, 26
rigid rotation, 141
samples, 132, 210
sensitivity analysis, 7, 169, 233
separability assumption, 213
shape deformation, 23
shape development, 41, 65
shape function, 43, 128, 155
shape optimization, 5, 23, 231
shape perturbation, 40, 46
shape polynomial functions, 21
shape sensitivity, 23, 57
shape variation, 81
shear stress, 37
single point conditioner, 214
singularity, 57
small perturbation techniques, 39, 44, 68
smoother, 29
snapshots, 234
source, 178
sparse matrix, 98
spectral condition, 212
stabilized Lagrangian, 177
state equation, 16, 174
state system, 16
steepest descent, 27
steepest descent method, 175
stenosis, 1, 134
stenosis length, 118
stent, 173
step size, 27
Stokes equations, 19, 40
Stokes Theorem, 37
stretching, 141
SUPG stabilization, 23, 80
supremizer, 134
supremizer operator, 93
supremizer option, 113
surgical prosthesis, 37
symmetric operator, 207
systolic phase, 79
T shaped bypass, 104
T shaped region, 218
Taylor expansion, 26
Taylor patch, 2, 34
Taylor-Hood elements, 94, 117
temperature, 219
theoretical upper bound, 123
thermal diffusivity, 219
time dependent problem, 198
time discretization, 80
transformation tensor, 92, 130
transport matrix, 201
truth approximation, 207
uniqueness continuation theorem, 51, 74
uniqueness results, 51, 74
unperturbed domain, 42, 65
unsteady blood flow, 79
unsteady cost functional, 69
unsteady Stokes equations, 63
upper bound, 124, 213
upper stream, 170
variable transformation, 42, 66
velocity error projection $H^1$, 98
velocity mass matrix, 98
velocity stiffness matrix, 98
velocity-divergence matrix, 98
viscous dissipation, 37
viscous drag, 31
viscous stresses, 31
viscous terms, 18
vortices structures, 24
vorticity, 23, 44, 57, 96, 171
vorticity reduction, 23, 81

wall curvature, 33
wall deformation, 24
wall shear stress, 17, 96
Wall Shear Stress Gradient WSSG, 17
wall stiffness, 23
wavy-walled channel, 136
Womersley inflow profile, 35

Young modulus, 24

zero-order problem, 44
Curriculum Vitæ


